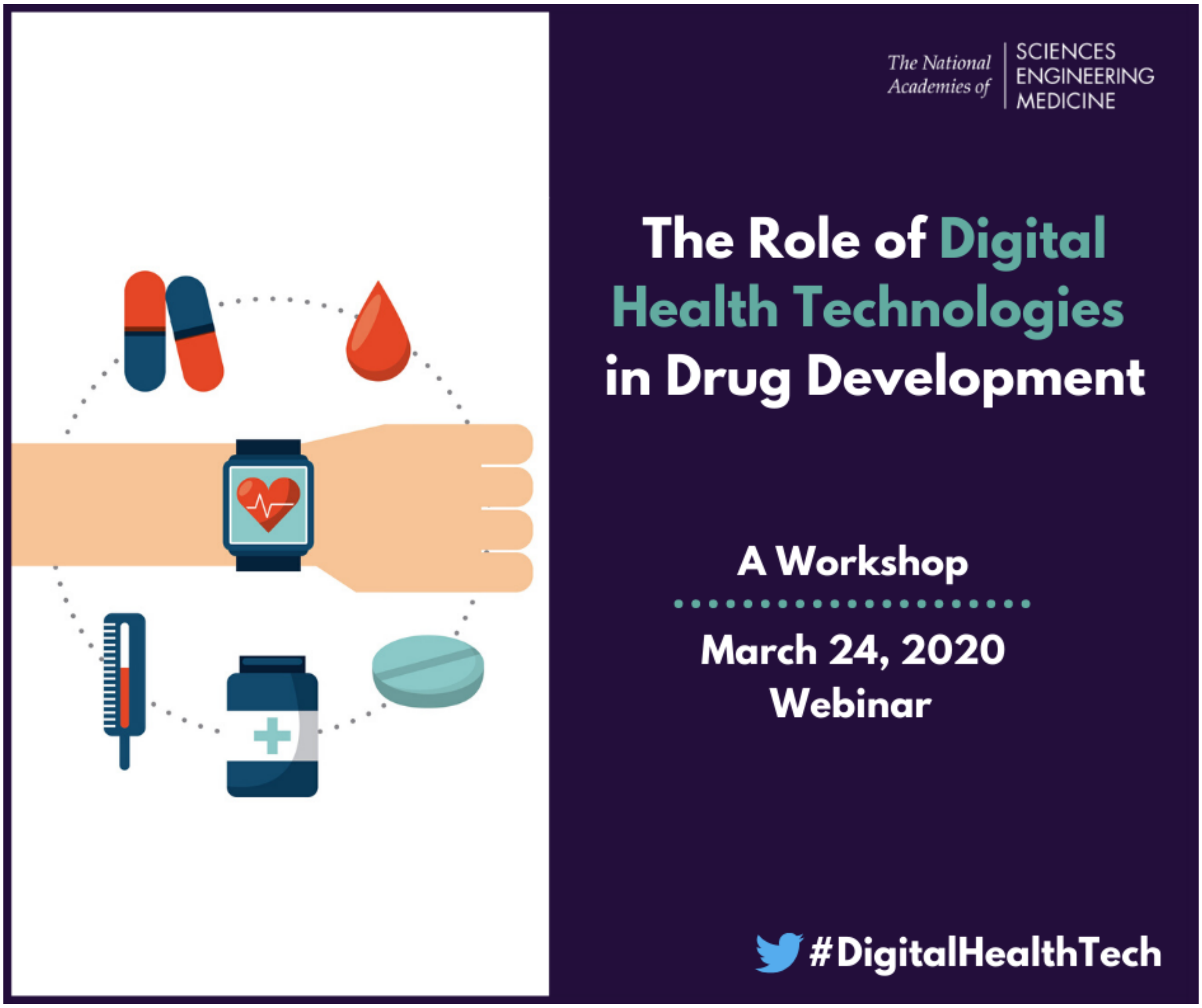
The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health



The National Academies of SCIENCES ENGINEERING MEDICINE

The Role of Digital Health Technologies in Drug Development

A Workshop

March 24, 2020
Webinar

#DigitalHealthTech

Please register on Zoom for the meeting at the following link by no later than March 24, 2020.

https://nasem.zoom.us/webinar/register/WN_8xfTtbfeQrGVh4Y1dAE38A

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The Role of Digital Health Technologies in Drug Development: A Workshop

March 24, 2020
Webinar

Table of Contents

The Role of Digital Health Technologies in Drug Development Workshop Background......................

Forum & Roundtable Information..

Workshop Planning Committee ....

Background Reading...

National Academies Activities of Interest……

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health

The Role of Digital Health Technologies in Drug Development

A Workshop

March 24, 2020 ▪ Zoom Webinar

Digital health technologies (e.g. smartphone apps, wearable sensors, and other remote, sensor-based tools that combine hardware and software) have become increasingly available to consumers, providers, and researchers. They offer new opportunities to address critical challenges or 'pain points,' better connect patients and health care providers, and incorporate patient input throughout the drug research and development (R&D) life cycle. This workshop will provide a venue to discuss challenges and opportunities in using digital health technologies to improve the probability of success in drug development. Workshop participants may consider key components for an evidence-based framework for applying digital health technologies towards drug research and development.

The public workshop will feature invited presentations and discussions to:

- Discuss challenges/questions that digital health technologies may be uniquely suited to address and opportunities for digital health technologies to enable better patient care and more efficient clinical trials;
- Consider strategies for evaluating and selecting digital health technologies that are fit for purpose in drug development (e.g. establishing appropriate evidentiary criteria);
- Discuss privacy, ethical, and regulatory issues related to the use of digital health technologies;
- Examine currently validated and/or FDA-approved drug development applications for digital health technologies;
- Consider opportunities to enable the practical application of digital health technologies for improving drug development (e.g. sharing best practices for the validation and use of digital health technologies, harmonizing guidelines across sectors).

The planning committee will organize the workshop, develop the agenda, select and invite speakers and discussants, and moderate or identify moderators for the discussions. Discussions may incorporate examples and lessons learned from other fields, such as direct-to-consumer genetic testing. A proceedings of the presentations and discussions at the workshop will be prepared by a designated rapporteur in accordance with institutional guidelines.

The National Academies of Sciences, Engineering, and Medicine are private, nonprofit institutions that provide expert advice on some of the most pressing challenges facing the nation and the world. Our work helps shape sound policies, inform public opinion, and advance the pursuit of science, engineering, and medicine. For more information about this workshop, please contact Carolyn Shore (cshore@nas.edu).

1

Planning Committee

Jennifer Goldsack (Co-Chair), Digital Medicine Society

Joseph Menetski (Co-Chair), Foundation for the National Institutes of Health

Linda Brady, Division of Neuroscience and Basic Behavioral Science, NIMH/NIH

Ray Dorsey, University of Rochester Medical Center

Deborah Estrin, Cornell Tech

Geoffrey Ginsburg, Duke University School of Medicine

Husseini Manji, Janssen Research and Development LLC

Deven McGraw, Ciitizen Corporation

Lauren Oliva, Biogen

Bray Patrick-Lake, Evidation Health

Leonard Sacks, Office of Medical Policy, CDER/FDA

Joyce Tung, 23andMe

Effy Vayena, ETH Zurich

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health

ZOOM WEBINAR GUIDANCE

TO: Workshop Attendees

SUBJECT: Zoom Webinar guidance for National Academies virtual workshop, *The Role of Digital Health Technologies in Drug Development* (March 24, 2020)

Please register on Zoom for the meeting at the following link by no later than March 24, 2020.

https://nasem.zoom.us/webinar/register/WN_8xfTtbfeQrGVh4Y1dAE38A

WORKSHOP BACKGROUND

This workshop will provide a venue to discuss challenges and opportunities in using digital health technologies (e.g. smartphone apps, wearable sensors, and other remote, sensor-based tools that combine hardware and software) to improve the probability of success in drug development.

Workshop objectives:

- Highlight critical barriers or 'pain points' along the drug R&D lifecycle for which digital health technologies may be uniquely suited to address.
- Consider lessons learned from currently validated digital health technology applications that could be generalizable for newer digital health technologies.
- Consider opportunities to enable the practical application of digital health technologies for improving drug development (e.g. sharing best practices for the validation and use of digital health technologies, harmonizing guidelines across sectors).
- Consider strategies to determine the evidentiary criteria for selecting digital health technologies that are fit for purpose in drug development (e.g. examining existing frameworks).
- Discuss privacy, ethical, and regulatory issues related to the use of digital health technologies.

ZOOM WEBINAR GUIDANCE

Date/Time: **Tuesday, March 24, 2020 (10:00 AM – 4:15 PM ET)**

Registration: You must register via Zoom in advance of the workshop at the following link:

https://nasem.zoom.us/webinar/register/WN_8xfTtbfeQrGVh4Y1dAE38A

You will receive a confirmation email with a personalized link to attend the webinar, as well as option to add the webinar to your calendar.

Webinar Participation:

All participants will be muted upon entry to the webinar. However, we encourage participation via the chat box feature, which will be located at the bottom of zoom meeting window. If you have questions during the webinar, please submit them through the Q&A box. We will be using the hashtag **#digitalhealthtech** and hope you continue the discussion on social media.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health

The Role of Digital Health Technologies in Drug Development

A Workshop

March 24, 2020
10:00 a.m. - 4:15 p.m. (ET)

ZOOM WEBINAR REGISTRATION:
https://nasem.zoom.us/webinar/register/WN_8xfTtbfeQrGVh4Y1dAE38A

Agenda

Digital health technologies (e.g. smartphone apps, wearable sensors, and other remote, sensor-based tools that combine hardware and software) have become increasingly available to consumers, providers, and researchers. They offer new opportunities to address critical challenges or pain points, better connect patients and health care providers, and incorporate patient input throughout the drug research and development (R&D) life cycle. This workshop will provide a venue to discuss challenges and opportunities in using digital health technologies to improve the probability of success in drug development. Workshop participants may consider key components for an evidence-based framework for applying digital health technologies towards drug research and development.

WORKSHOP OBJECTIVES:

- Highlight critical barriers or "pain points" along the drug R&D lifecycle for which digital health technologies may be uniquely suited to address;
- Consider lessons learned from currently validated digital health technology applications that could be generalizable for newer digital health technologies;
- Consider opportunities to enable the practical application of digital health technologies for improving drug development (e.g. sharing best practices for the validation and use of digital health technologies, harmonizing guidelines across sectors).
- Consider strategies for evaluating and selecting digital health technologies that are fit-for-purpose in drug development (e.g. examining existing frameworks, establishing appropriate evidentiary criteria);
- Discuss privacy, ethical, and regulatory issues related to the use of digital health technologies.

10:00 a.m. ET **Welcome**

ROBERT CALIFF *Forum Co-Chair* Verily Life Sciences	GEOFFREY GINSBURG *Roundtable Co-Chair* Duke University School of Medicine

Openning Remarks

JENNIFER GOLDSACK, *Workshop Co-Chair*
Executive Director
Digital Medicine Society

JOSEPH MENETSKI, *Workshop Co-Chair*
Associate Vice President of Research Partnerships
Foundation for the National Institutes of Health

BRIEFING: ETHICAL CONSIDERATIONS

10:15 a.m. ET ***Ethicist Perspective***
CAMILLE NEBEKER
Director
Research Center for Optimal Digital Ethics
University of California San Diego

SESSION I DIGITAL TOOLS FOR CHARACTERIZING DISEASE

10:45 a.m. ET **Session Moderator**
EFFY VAYENA
Professor
Health Ethics and Policy Lab, ETH Zurich

Non-Profit Perspective/Platform Research Perspective
LARSSON OMBERG
Vice President, Systems Biology
Sage Bionetworks

NIH Perspective
CHRIS LUNT
Chief Technology Officer
All of Us Research Program
National Institutes of Health

Patient Engagement Perspective
ALICIA STALEY
Senior Director, Patient Engagement
Medidata Solutions

Developer Perspective
LUCA FOSCHINI
Chief Data Scientist & Co-founder
Evidation Health, Inc.

11:25 a.m. ET **Panel Discussion with Speakers and Workshop Participants**

11:45 a.m. ET **BREAK**

SESSION II DIGITAL TOOLS FOR RECRUITMENT AND SAFETY TRIALS

12:00 p.m. ET **Session Moderator**
DEVEN MCGRAW
Chief Regulatory Officer
Ciitizen Corporation

Regulatory Perspective
CHRISTOPHER LEPTAK
Director, Regulatory Science Program, Office of New Drugs
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

Industry Perspective
YVONNE YU-FENG CHAN
Senior Director, Medical Affairs for Digital Medicine
Otsuka Pharmaceutical Companies

Developer Perspective
Chris Benko
Chief Executive Officer
Konesksa Health

Academic Perspective
ERIC PERAKSLIS
Rubenstein Fellow
Duke University

12:40 p.m. ET **Panel Discussion with Speakers and Workshop Participants**

1:00 p.m. ET **BREAK**

FIRESIDE CHAT

1:30 p.m. ET **Session Moderator**
JENNIFER GOLDSACK, *Workshop Co-Chair*
Executive Director
Digital Medicine Society

Regulatory Perspective
AMY ABERNETHY
Principal Deputy Commissioner
U.S. Food and Drug Administration

SESSION III DIGITAL TOOLS FOR PIVOTAL TRIALS

2:00 p.m. ET **Session Moderator**
HUSSEINI MANJI
Global Therapeutic Head, Neuroscience
Janssen Research & Development

Regulatory Perspective
LEONARD SACKS
Associate Director of Clinical Methodology, Office of Medical Policy
Center for Drug Evaluation and Research
U.S. Food and Drug Administration

Industry Perspective
SEAN KHOZIN
Global Head of Data Strategy
Janssen Research & Development

Developer Perspective
RITU KAPUR
Head of Biomarkers
Verily Life Sciences

2:30 p.m. ET **Panel Discussion with Speakers and Workshop Participants**

2:50 p.m. ET **BREAK**

SESSION IV DIGITAL TOOLS FOR POSTREGISTRATION SURVEILLANCE

3:00 p.m. ET **Session Moderator**
CHRISTINA SILCOX
Managing Associate
Duke-Margolis Center for Health Policy

Industry Perspective
MICHELLE CROUTHAMEL
Director, Digital Health & Innovation
AbbVie

Patient Engagement Perspective
SALLY OKUN
Policy and Ethics
UnitedHealth Group Research & Development

Clinician/Health System Perspective
EDMONDO ROBINSON
Chief Digital Innovation Officer
Moffitt Cancer Center

3:30 p.m. ET **Panel Discussion with Speakers and Workshop Participants**

KEY REFLECTIONS AND NEXT STEPS

3:45 p.m. ET **Key Reflections and Next Steps**

JENNIFER GOLDSACK, *Workshop Co-Chair*
Executive Director
Digital Medicine Society

JOSEPH MENETSKI, *Workshop Co-Chair*
Associate Vice President of Research Partnerships
Foundation for the National Institutes of Health

4:15 p.m. ET **Adjourn**

ABOUT THE FORUM

Forum on
DRUG DISCOVERY, DEVELOPMENT,
and **TRANSLATION**

The Forum on Drug Discovery, Development, and Translation of the National Academies of Sciences, Engineering, and Medicine was created in 2005 by the Board on Health Sciences Policy to provide a unique platform for dialogue and collaboration among thought leaders and stakeholders in government, academia, industry, foundations, and patient advocacy with an interest in improving the system of drug discovery, development, and translation. The Forum brings together leaders from private sector sponsors of biomedical and clinical research, federal agencies sponsoring and regulating biomedical and clinical research, the academic community, and patients, and in doing so serves to educate the policy community about issues where science and policy intersect. The Forum convenes several times each year to identify, discuss, and act on key problems and strategies in the discovery, development, and translation of drugs. To supplement the perspectives and expertise of its members, the Forum also holds public workshops to engage a wide range of experts, members of the public, and the policy community. The Forum also fosters collaborations among its members and constituencies. The activities of the Forum are determined by its members, focusing on the major themes outlined below.

INNOVATION AND THE DRUG DEVELOPMENT ENTERPRISE

Despite exciting scientific advances, the pathway from basic science to new therapeutics faces challenges on many fronts. New paradigms for discovering and developing drugs are being sought to bridge the ever-widening gap between scientific discoveries and translation of those discoveries into life-changing medications. There is also increasing recognition of the need for new models and methods for drug development and translational science, and "precompetitive collaborations" and other partnerships, including public–private partnerships, are proliferating. The Forum offers a venue to discuss effective collaboration in the drug discovery and development enterprise and also hosts discussions that could help chart a course through the turbulent forces of disruptive innovation in the drug discovery and development "ecosystem."

Key gaps remain in our knowledge about science, technology, and methods needed to support drug discovery and development. Recent rapid advances in innovative drug development science present opportunity for revolutionary developments of new scientific techniques, therapeutic products, and applications. The Forum provides a venue to focus ongoing attention and visibility to these important drug development needs and facilitates exploration of new approaches across the drug development lifecycle. The Forum has held workshops that have contributed to the defining and establishment of regulatory science and have helped inform aspects of drug regulatory evaluation.

CLINICAL TRIALS AND CLINICAL PRODUCT DEVELOPMENT

Clinical research is the critical link between bench and bedside in developing new therapeutics. Significant infrastructural, cultural, and regulatory impediments challenge efforts to integrate clinical trials into the health care delivery system. Collaborative, cross-sector approaches can help articulate and address these key challenges and foster systemic responses. The Forum has convened a multiyear initiative to examine the state of clinical trials in the United States, identify areas of strength and weakness in our current clinical trial enterprise, and consider transformative strategies for enhancing the ways in which clinical trials are organized and conducted. In addition to sponsoring multiple symposia and workshops, under this initiative, the Forum is fostering innovative, collaborative efforts to facilitate needed change in areas such as improvement of clinical trial site performance.

INFRASTRUCTURE AND WORKFORCE FOR DRUG DISCOVERY, DEVELOPMENT, AND TRANSLATION

Considerable opportunities remain for enhancement and improvement of the infrastructure that supports the drug development enterprise. That infrastructure, which includes the organizational structure, framework, systems, and resources that facilitate the conduct of biomedical science for drug development, faces significant challenges. The science of drug discovery and development, and its translation into clinical practice, is cross-cutting and multidisciplinary. Career paths can be opaque or lack incentives such as recognition, career advancement, or financial security. The Forum has considered workforce needs as foundational to the advancement of drug discovery, development, and translation. It has convened workshops examining these issues, including consideration of strategies for developing a discipline of innovative regulatory science through the development of a robust workforce. The Forum will also host an initiative that will address needs for a workforce across the translational science lifecycle.

Forum on Drug Discovery, Development, and Translation

Robert Califf (Co-Chair)
Verily Life Sciences and Google Health; Duke University and Stanford University

Gregory Simon (Co-Chair)
Kaiser Permanente Washington Health Research Institute and University of Washington

Amy Abernethy
Office of the Commissioner, U.S. FDA

Christopher Austin
National Center for Advancing Translational Sciences, NIH

Linda Brady
National Institute of Mental Health, NIH

Rick Bright
Biomedical Advanced Research and Development Authority

Barry Coller
The Rockefeller University

Thomas Curran
Children's Mercy, Kansas City

Richard Davey
National Institute of Allergy and Infectious Diseases, NIH

Katherine Dawson
Biogen

James Doroshow
National Cancer Institute, NIH

Jeffrey Drazen
New England Journal of Medicine

Steven Galson
Amgen Inc.

Carlos Garner
Eli Lilly and Company

Julie Gerberding
Merck & Co., Inc.

Anne Heatherington
Takeda Pharmaceuticals

Deborah Hung
Harvard Medical School

Esther Krofah
FasterCures/Milken Institute

Lisa LaVange
University of North Carolina Gillings School of Global Public Health

Ross McKinney Jr.
Association of American Medical Colleges

Joseph Menetski
Foundation for the NIH

Arti Rai
Duke University School of Law

Kelly Rose
Burroughs Wellcome Fund

Susan Schaeffer
The Patients' Academy for Research Advocacy

Joseph Scheeren
Critical Path Institute

Rob Scott
AbbVie, Inc.

Anantha Shekhar
Indiana University School of Medicine

Jay Siegel
Retired

Ellen Sigal
Friends of Cancer Research

Lana Skirboll
Sanofi

Amir Tamiz
National Institute of Neurological Disorders and Stroke, NIH

Ann Taylor
AstraZeneca

Pamela Tenaerts
Clinical Trials Transformation Initiative

Joanne Waldstreicher
Johnson & Johnson

Jonathan Watanabe
University of California, San Diego

Carrie Wolinetz
Office of Science Policy, NIH

Alastair Wood
Vanderbilt University

Janet Woodcock
Center for Drug Evaluation and Research, U.S. FDA

Project Staff

Carolyn Shore, Ph.D.
Forum Director

Jennifer Hinners, M.D., M.P.H.
Program Officer

Amanda Wagner Gee, M.S.
Program Officer

Eeshan Khandekar, M.Sc.
Associate Program Officer

Melvin Joppy
Senior Program Assistant

For more information, please visit:

NATIONALACADEMIES.ORG/DRUGFORUM

Health and Medicine Division
Board on Health Sciences Policy

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Roundtable on GENOMICS and PRECISION HEALTH

The sequencing of the human genome is rapidly opening new doors to research and progress in biology, medicine, and health care. At the same time, these developments have produced a diversity of new issues to be addressed.

The National Academies of Sciences, Engineering, and Medicine has convened a Roundtable on Genomics and Precision Health (previously the Roundtable on Translating Genomic-Based Research for Health) that brings together leaders from academia, industry, government, foundations and associations, and representatives of patient and consumer interests who have a mutual concern and interest in addressing the issues surrounding the translation of genome-based research for use in maintaining and improving health. The mission of the Roundtable is to advance the field of genomics and improve the translation of research findings to health care, education, and policy. The Roundtable will discuss the translation process, identify challenges at various points in the process, and discuss approaches to address those challenges.

The field of genomics and its translation involves many disciplines, and takes place within different economic, social, and cultural contexts, necessitating a need for increased communication and understanding across these fields. As a convening mechanism for interested parties from diverse perspectives to meet and discuss complex issues of mutual concern in a neutral setting, the Roundtable: fosters dialogue across sectors and institutions; illuminates issues, but does not necessarily resolve them; and fosters collaboration among stakeholders.

To achieve its objectives, the Roundtable conducts structured discussions, workshops, and symposia. Workshop summaries will be published and collaborative efforts among members are encouraged (e.g., journal articles). Specific issues and agenda topics are determined by the Roundtable membership, and span a broad range of issues relevant to the translation process.

Issues may include the integration and coordination of genomic information into health care and public health including encompassing standards for genetic screening and testing, improving information technology for use in clinical decision making, ensuring access while protecting privacy, and using genomic information to reduce health disparities. The patient and family perspective on the use of genomic information for translation includes social and behavioral issues for target populations. There are evolving requirements for the health professional community, and the need to be able to understand and responsibly apply genomics to medicine and public health.

Of increasing importance is the need to identify the economic implications of using genome-based research for health. Such issues include incentives, cost-effectiveness, and sustainability.

Issues related to the developing science base are also important in the translation process. Such issues could include studies of gene-environment interactions, as well as the implications of genomics for complex disorders such as addiction, mental illness, and chronic diseases.

Roundtable sponsors include federal agencies, pharmaceutical companies, medical and scientific associations, foundations, and patient/public representatives. For more information about the Roundtable on Genomics and Precision Health, please visit our website at nationalacademies.org/GenomicsRT or contact Sarah Beachy at 202-334-2217, or by e-mail at sbeachy@nas.edu.

Roundtable on Genomics and Precision Health Membership

Geoffrey Ginsburg, M.D., Ph.D. (Co-Chair) Duke University
Michelle Penny, Ph.D. (Co-Chair) Goldfinch Bio

Naomi Aronson, Ph.D.
BlueCross/BlueShield Association

Aris Baras, M.D., M.B.A.
Regeneron Pharmaceuticals

Karina Bienfait, Ph.D.
Merck and Co., Inc.

Vence Bonham, Jr., J.D.
National Human Genome Research Institute

Robert B. Darnell, M.D. Ph.D.
The Rockefeller University / NY Genome Center

Stephanie Devaney, Ph.D.
All of Us Research Program, NIH

Katherine Donigan, Ph.D.
U.S. Food and Drug Administration

W. Gregory Feero, M.D., Ph.D.
Journal of the American Medical Association

Jessica M. Gill, Ph.D., R.N., FAAN
National Institute of Nursing Research

Marc Grodman, M.D.
Genosity

Richard Hodes, M.D.
National Institute on Aging

Praduman Jain, M.S.
Vibrent Health

Sekar Kathiresan, M.D.
Massachusetts General Hospital

Muin Khoury, M.D., Ph.D.
Centers for Disease Control and Prevention

David Ledbetter, Ph.D.
Geisinger

Charles Lee, Ph.D., FACMG
The Jackson Laboratory for Genomic Medicine

Thomas Lehner, Ph.D., M.P.H.
National Institute of Mental Health

Debra Leonard, M.D., Ph.D.
College of American Pathologists

Patrick Loerch, Ph.D.
Johnson & Johnson

James Lu, M.D., Ph.D.
Helix

Sean McConnell, Ph.D.
American Medical Association

Mona Miller, M.P.P.
American Society of Human Genetics

Jennifer Moser, Ph.D.
U.S. Department of Veterans Affairs

Anna Pettersson, Ph.D.
Pfizer Inc.

Victoria M. Pratt, Ph.D., FACMG
Association for Molecular Pathology

Nadeem Sarwar, Ph.D.
Eisai Inc.

Sheri Schully, Ph.D.
NIH Office of Disease Prevention

Joan A. Scott, M.S., C.G.C.
Health Resources and Services Administration

Nonniekaye Shelburne, C.R.N.P., M.S., A.O.C.N.
National Cancer Institute

Nikoletta Sidiropoulos, M.D.
University of Vermont Health Network Medical Group

Katherine Johansen Taber, Ph.D.
Myriad Women's Health

Ryan Taft, Ph.D.
Illumina

Jacquelyn Taylor, Ph.D.
New York University

Sharon Terry, M.A.
Genetic Alliance

Joyce Tung, Ph.D.
23andMe, Inc.

Jameson Voss, M.D.
Air Force Medical Support Agency

Catherine A. Wicklund, M.S., C.G.C.
National Society of Genetic Counselors

Huntington F. Willard, Ph.D.
Genome Medical

Janet K. Williams, Ph.D., R.N., FAAN
American Academy of Nursing

Sarah Wordsworth, Ph.D.
University of Oxford

Alicia Zhou, Ph.D.
Color Genomics

Member TBD
American College of Medical Genetics and Genomics

Project Staff
Sarah H. Beachy, Ph.D., *Roundtable Director*
Siobhan Addie, Ph.D., *Program Officer*
Meredith Hackmann, *Associate Program Officer*
Kelly Choi, *Senior Program Assistant*

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health

The Role of Digital Health Technologies in Drug Development

A Workshop

Planning Committee Biographies

CO-CHAIRS

JENNIFER GOLDSACK, M.CHEM., M.A., M.B.A., is the Interim Executive Director at the Digital Medicine Society (DiME), a new professional organization promoting the adoption of digital technologies for health. Previously, Ms Goldsack spent several years at the Clinical Trials Transformation Initiative (CTTI) where she led development and implementation several projects within CTTI's Mobile Program and was the operational co-lead on the first randomized clinical trial using FDA's Sentinel System. Ms Goldsack spent five years working in research at the Hospital of the University of Pennsylvania, first in Outcomes Research in the Department of Surgery and later in the Department of Medicine. More recently, Ms Goldsack helped launch the Value Institute, a pragmatic research and innovation center embedded in a large academic medical center in Delaware. Ms Goldsack earned her master's degree in chemistry from the University of Oxford, England, her masters in the history and sociology of medicine from the University of Pennsylvania, and her MBA from the George Washington University. Additionally, she is a certified Lean Six Sigma Green Belt and a Certified Professional in Healthcare Quality. Ms Goldsack is a retired athlete, formerly a Pan American Games Champion, Olympian and World Championship silver medalist.

JOSEPH P. MENETSKI, PH.D., received his Ph.D. from Northwestern University Medical School with Dr. Stephen Kowalczykowski and completed his post-doctoral training at the Laboratory of Molecular Biology, National Institutes of Health (NIH/NIDDK) with Dr. Martin Gellert. He then started his career in industry in 1993 in the Immunopathology Department at Parke-Davis (later Pfizer), where he established a discovery research program in cellular inflammation that eventually transitioned to the molecular study of osteoarthritis. Joseph moved to Merck in 2004. His first position was in the department of Immunology where he was involved in the osteoarthritis new targets and biomarker program. While at Merck he has been a member of the Molecular Profiling group, the Knowledge Discovery and Knowledge Management group and finally a Director in Global Competitive Intelligence. Over the years, he has been a key contributor to many basic research and clinical programs in the areas of arthritis, sarcopenia, osteoporosis and asthma. He has served as a core research team member on several external basic research projects for identification of new targets and molecular biomarkers. His industry research and development experiences include target identification, compound selection, translational biomarker identification, clinical study design and analysis, and external scientific collaborations. In the commercial space, he has been intimately involved in opportunity and asset identification and qualification, and in assessing the competitive landscape of disease areas that he is supporting. During this time, he has been recognized by multiple research and development awards for his contributions

MEMBERS

LINDA BRADY, PH.D., serves as the Director of the Division of Neuroscience and Basic Behavioral Science at the National Institute of Mental Health (NIMH). In this role, she provides scientific, programmatic, and administrative leadership for an extramural research program portfolio in basic neuroscience to support NIMH's mission of transforming the understanding and treatment of mental illnesses. Dr. Brady has directed programs in neuropharmacology, drug discovery, and clinical therapeutics and organized Consortia focused on ways to accelerate the development and clinical application of radiotracers in clinical research. She has provided leadership for many programs, including: Development and Application of PET and SPECT Imaging Ligands as Biomarkers for Drug Discovery and for Pathophysiological Studies of CNS Disorders, the National Cooperative Drug/Device Discovery/Development Groups for the Treatment of Mental Disorders, and First in Human and Early Stage Clinical Trials of Novel Investigational Drugs or Devices for Psychiatric Disorders. Dr. Brady serves as co-chair of the Neuroscience Steering Committee for the Biomarkers Consortium, a public-private research partnership of the Foundation for the National Institutes of Health (FNIH) that focuses on discovery, development, and qualification of biological markers to support drug development, preventive medicine, and medical diagnostics. From 2004-2013, she co-led the Molecular Libraries and Imaging Program, a trans-NIH Common Fund initiative to provide biomedical researchers access to small organic molecules that can be used as chemical probes to study the functions of genes, cells, and biochemical pathways in health and disease. Dr. Brady was trained in pharmacology and neuroscience. She completed her Ph.D. at Emory University School of Medicine, followed by post-doctoral work and research positions at the Uniformed Services University of the Health Sciences and the NIMH Intramural Research Program. She is the author of more than 70 peer reviewed scientific publications and is a member of the Society for Neuroscience and a Fellow in the American College of Neuropsychopharmacology. Dr. Brady has received NIH Director's Awards and NIH Merit Awards in recognition of her activities in biomarker development and drug development for mental disorders.

RAY DORSEY, M.D., M.B.A., is a professor of neurology and director of the Center for Human Experimental Therapeutics at the University of Rochester Medical Center. Dr. Dorsey is investigating new treatments for movement disorders and is working on ways to improve the way care is delivered for individuals with Parkinson's disease (PD) and other neurological disorders. Using simple web-based video conferencing, he and his colleagues are seeking to provide care to individuals with PD and neurological diseases. Dr. Dorsey's research has been published in leading medical and neurology journals and has been featured on National Public Radio and in The New York Times and The Wall Street Journal. He previously directed the movement disorders division at Johns Hopkins and worked as a consultant for McKinsey & Company. He completed his undergraduate studies at Stanford University, business school at the Wharton School and medical school at the University of Pennsylvania.

Deborah Estrin, Ph.D., M.S., is the Robert V. Tishman '37 Professor at Cornell Tech and in the Computer Science Department at Cornell University, and currently serves as Associate Dean for Impact at Cornell Tech. She is founder of the Health Tech Hub and directs the Small Data Lab at Cornell Tech, which develops new personal data APIs and applications for individuals to harvest the small data traces they generate daily. Estrin is also co-founder of the non-profit startup, Open mHealth.

Previously, Estrin was on the UCLA faculty where she was the Founding Director of the NSF Center for Embedded Networked Sensing (CENS), pioneering the development of mobile and wireless systems to collect and analyze real time data about the physical world and the people who occupy it.

Estrin was chosen as a 2018 fellow of the MacArthur Foundation.

GEOFFREY GINSBURG, M.D., PH.D., is the founding director for the Center for Applied Genomics & Precision Medicine at the Duke University Medical Center. He is also the founding director for MEDx, a partnership between the Schools of Medicine and Engineering to spark and translate innovation. He is a professor of Medicine, Pathology, and Biomedical Engineering, and a professor in School of Nursing at Duke University.

While at Duke, Dr. Ginsburg has pioneered translational genomics, the development of novel diagnostics, and precision medicine, initiating programs in genome enabled biomarker discovery, longitudinal registries with linked molecular and clinical data, biomarker-informed clinical trials, and the development of novel practice models and implementation research for the integration of genomic tools and digital health technologies into heath care delivery systems.

In 1990, he was recruited to the faculty of Harvard Medical School, where he was director of Preventive Cardiology at Beth Israel Hospital and led a laboratory in applied genetics of cardiovascular diseases at Children's Hospital. In 1997 he joined Millennium Pharmaceuticals Inc, as senior program director for cardiovascular diseases and was eventually appointed vice president of Molecular and Personalized Medicine, where he was responsible for developing pharmacogenomic strategies for therapeutics, as well as biomarkers for disease and their implementation in the drug development process.

He has received a number of awards, including the Innovator in Medicine Award from Millennium in 2004, the Basic Research Achievement Award in Cardiovascular Medicine from Duke in 2005, and the ILCHUN Molecular Medicine Award from Korean Society for Biochemistry and Molecular Biology in 2014. In 2015 he was an honored speaker at the White House Champions for Change in Precision Medicine. He received Duke's Research Mentoring Award in 2017. He is a founding member and former board member of the Personalized Medicine Coalition, a section editor for The Journal of the American College of Cardiology and an editorial advisor for Science Translational Medicine. In addition, he is the editor of Genomic and Personalized Medicine (Elsevier), 3rd edition published in 2016. He is a member of the Faculty of 1000. He has been a member of the Secretary of Veterans Affairs Advisory Council on Genomic Medicine, a member of the NIGMS External Scientific Panel for the Pharmacogenomics Research Network, the Board of External Experts for the National Heart, Lung and Blood Institute, the National Advisory Council for Human Genome Research at NIH, the Advisory Council for the National Center for Advancing Translational Sciences at NIH (where he was the Vice Chair for the Cures Acceleration Network Board), and the World Economic Forum's Global Agenda Council on The Future of the Health Sector. He is co-chair of the National Academies Roundtable on Genomic and Precision Health and co-chair of the Global Genomic Medicine Collaborative and is a member of the Advisory Committee to the Director of NIH.

HUSSEINI K. MANJI, M.D., FRCPC., is the Global Therapeutic Head for Neuroscience at Janssen Research & Development, LLC, one of the Johnson & Johnson pharmaceutical companies. He is also Visiting Professor at Duke University. Dr. Manji was previously Chief of the Laboratory of Molecular Pathophysiology & Experimental Therapeutics at the National Institutes of Health (NIH) and Director of the NIH Mood and Anxiety Disorders Program, the largest program of its kind in the world.

The major focus of Dr. Manji's research is the investigation of disease- and treatment-induced changes in gene and protein networks that regulate synaptic and neural plasticity in neuropsychiatric disorders. His work has helped to conceptualize these illnesses as genetically-influenced disorders of synaptic and neural plasticity and has led to the investigation of novel therapeutics for refractory patients. Notably, Dr. Manji's research demonstrated that AMPA- and NMDA-mediated synaptic plasticity may underlie the pathogenesis of depression, and that targeting these pathways may produce robust and rapid antidepressant effects. Under his leadership, this has led to the FDA Approval of the first novel antidepressant mechanism (NMDA-antagonism) in decades. Spravato (an NMDA antagonist) was demonstrated to produce robust and rapid antidepressant effects and is approved for treatment resistant depression. Phase 3 studies investigating its efficacy in the treatment of suicidal ideation are underway.

Dr. Manji has received a number of prestigious awards, including the NIMH Director's Career Award for Significant Scientific Achievement, the A. E. Bennett Award for Neuropsychiatric Research, the Ziskind-Somerfeld Award for Neuropsychiatric Research, the NARSAD Mood Disorders Prize, the Mogens Schou Distinguished Research Award, the ACNP's Joel Elkes Award for Distinguished Research, the DBSA Klerman Senior Distinguished Researcher Award, the Briggs Pharmacology Lectureship Award, the American Federation for Aging Research Award of Distinction, the Caring Kind Alzheimer's Disease Leadership Award, and the Global Health & the Arts Award of Recognition. He has received PhRMA's Research & Hope Award for Excellence in Biopharmaceutical Research, and has also been recognized as one of 14 inaugural "Health Heroes" by Oprah magazine.

Dr. Manji has been inducted into the National Academy of Medicine (formerly IOM), the World Economic Forum (WEF) Global Future Councils, and has held numerous leadership positions within the NAM, the FNIH Biomarkers Consortium Executive Committee, the ACNP, and the Society of Biological Psychiatry.

Throughout his career, Dr. Manji has also been committed to undertakings related to medical and neuroscience education and has worked with the National Board of Medical Examiners (NMBE), the Howard Hughes Medical Institute Research Scholars Program, and numerous national curriculum committees. He founded and co-directed the NIH Foundation for the Advanced Education in the Sciences Graduate Course in the Neurobiology of Neuropsychiatric Illness, and has received several teaching and mentoring awards. He has also served as Editor, and on editorial boards of numerous scientific journals.

Dr. Manji has published extensively on the molecular and cellular neurobiology of severe neuropsychiatric disorders and development of novel therapeutics, with over 300 publications in peer-reviewed journals, including Science, Science Translational Medicine, Nature Neuroscience, Nature Reviews Neuroscience, Nature Reviews Drug Discovery , NEJM, J Clinical Investigation, PNAS, J Neuroscience, JAMA Psychiatry, Molecular Psychiatry.

DEVEN MCGRAW, JD, MPH, LLM., is the Chief Regulatory Officer for Ciitizen. Prior to joining Ciitizen, she directed U.S. health privacy and security policy through her roles as Deputy Director for Health Information Privacy at the HHS Office for Civil Rights (the office that oversees HIPAA policy and enforcement) and Chief Privacy Officer (Acting) of the Office of the National Coordinator for Health IT. Deven also advised PCORNet (the Patient Centered Outcomes Research Network), as well as the federal All of Us Research Initiative, on HIPAA and patient-donated data research initiatives.

LAUREN OLIVA, PHARM.D., is the Global Regulatory Policy Lead for New Technologies at Biogen. She oversees the development and execution of R&D's policy roadmap for digital health tools and gene therapy to enable Biogen's neuroscience portfolio. In her time at Biogen she launched a widely used Regulatory Intelligence service and served as a policy lead and regulatory strategy manager. Lauren received her PharmD from Rutgers University, Ernest Mario School of Pharmacy and has previously served as adjunct faculty and lecturer at MCPHS University in Boston, MA.

BRAY PATRICK-LAKE, M.F.S., is the Director of Strategic Partnerships at Evidation Health. She develops collaborations to support the design and implementation of participant-centered studies, and serves on the *All of Us* National Advisory Panel, the Digital Medicine Society Scientific Leadership Board, and the National Academies of Sciences, Engineering, and Medicine (NASEM) Health Science Policy Board. Previously, Bray led engagement for the Duke Clinical Research Institute Project Baseline Study Coordinating Center and served as co-chair on the Advisory Committee to the NIH Director that authored the Precision Medicine Initiative's Cohort Program. Bray holds a BS degree from the University of Georgia, and an MFS degree from National University.

LEONARD SACKS, M.D., received his medical education in South Africa, moving to the USA in 1987, where he completed fellowships in immunopathology and Infectious Diseases. He worked as an attending physician in Washington DC and South Africa and he joined the FDA in 1998 as medical reviewer in the Office of New Drugs. Subsequent positions included acting director of the Office of Critical Path Programs and associate director for clinical methodology in the Office of Medical Policy in the Center for Drug Evaluation and Research. In this capacity he has led efforts to support the use of electronic technology in clinical drug development. Besides his involvement in the design and analysis of clinical trials, he maintains a special interest in tuberculosis and other tropical diseases and has published and presented extensively on these topics. He is board certified in Internal Medicine and Infectious Diseases and holds an academic appointment as Associate Clinical Professor of Medicine at George Washington University.

JOYCE TUNG, PH.D., joined 23andMe in 2007 and manages the 23andMe research team, which is responsible for consumer health and ancestry research and development, academic and industry collaborations, computational analyses for therapeutics, and new research methods and tools development. While a postdoctoral fellow at Stanford University, Joyce studied the genetics of mouse and human pigmentation. She graduated from Stanford with honors and distinction with a B.S. in Biological Sciences and a minor in computer science, and earned her Ph.D. in Genetics from the University of California, San Francisco where she was a National Science Foundation graduate research fellow.

EFFY VAYENA, PH.D., studied Medical History and Bioethics at the University of Minnesota (USA) and completed her habilitation in Bioethics and Health Policy at the University of Zurich. From 2000-2007 she worked at the World Health Organization (WHO). In 2007, she joined the Institute of Biomedical Ethics and History of Medicine at the University of Zurich, with which she remains affiliated. She is a consultant to WHO on several projects, and visiting faculty at the Harvard Center for Bioethics, Harvard Medical School. In 2015 she was named a Swiss National Science Foundation (SNSF) Professor of Health Policy and leads the newly-established Health Ethics and Policy Lab in the Department of Public Health at the EBPI, University of Zurich. Her current research focus is on ethical and policy questions in personalized medicine and digital health. At the intersection of multiple fields, she relies on normative analyses and empirical methods to explore how values such as freedom of choice, participation and privacy are affected by recent developments in personalized medicine and in digital health. She is particularly interested in the issues of ethical oversight of research uses of big data, ethical uses of big data for global health, as well as the ethics of citizen science. Using the ethics lens in innovative ways, her work aims to provide concrete policy recommendations and frameworks that facilitate the use of new technologies for a better and more just health.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health

The Role of Digital Health Technologies in Drug Development

A Workshop

Speaker Biographies

AMY ABERNETHY, M.D., PH.D., is an oncologist and internationally recognized clinical data expert and clinical researcher. As the Principal Deputy Commissioner of Food and Drugs, Dr. Abernethy helps oversee FDA's day-to-day functioning and directs special and high-priority cross-cutting initiatives that impact the regulation of drugs, medical devices, tobacco & food. As acting Chief Information Officer, she oversees FDA's data and technical vision, and its execution. She's held multiple executive roles at Flatiron Health and was professor of medicine at Duke University School of Medicine, where she ran the Center for Learning Health Care and the Duke Cancer Care Research Program. Dr. Abernethy received her M.D. at Duke University, where she did her internal medicine residency, served as chief resident, and completed her hematology/oncology fellowship. She received her Ph.D. from Flinders University, her B.A. from the University of Pennsylvania and is boarded in palliative medicine.

CHRIS BENKO, M.B.A., is the CEO and co-founder of Koneksa Health, the leader in developing and implementing patient-focused digital biomarkers for drug development. By unlocking the potential of real-world data from remote, wearable, and other digital technologies, Koneksa speeds up the time required to understand how a drug is working, requiring fewer patients, and develops real-world evidence for how medicines can impact their daily lives. Prior to founding Koneksa, Chris was a vice president in Merck's corporate strategy office, working with its Global Health Innovation venture capital fund. He began his career at Merck in 1995 and progressed through roles in information technology as well as talent and organizational development, working in R&D, commercial, and at the corporate level as vice president for global talent management.

YVONNE YU-FENG CHAN, M.D., PH.D., F.A.C.E.P., is the Senior Director of Medical Affairs for Digital Medicine at Otsuka America Pharmaceutical, Inc., (Otsuka), a national leader in digital medicine research, and a board-certified Emergency physician. At Otsuka, Dr. Chan develops advanced methodologies, digital tools, and technology platforms to derive real-world clinical and health economics evidence in collaboration with internal and external collaborators. She provides medical input to all aspects of product development at Otsuka's Digital Medicine division. Dr. Chan leverages her 15+ years of medical and digital health experience as a physician-scientist to help lead Otsuka's pioneering work in digital medicine in support of patients, physicians, and caregivers. Previously, Dr. Chan was the Founding Director of the Center for Digital Health at the Icahn School of Medicine at Mount Sinai. The mission of her Digital Health Center was to drive large-scale patient participation in biomedical research and clinical care, by leveraging the latest digital technology and advanced analytic techniques to uncover novel insights and actionable results.

Dr. Chan is an editorial board member of Nature Partner Journals (npj) *Digital Medicine* and *Digital Biomarker* Journals. She is also a member of the Digital Medicine (DiMe) Society's Scientific Leadership Board and the NIH's National Institute of Neurological Disorders and Stroke (NINDS) Common Data Elements for Stroke Group.

Dr. Chan received her B.A. and M.D. from Rutgers University (New Jersey Medical School) and completed her Emergency Medicine training at Albert Einstein School of Medicine, Long Island Jewish Medical Center. At the conclusion of her NINDS T32 Cerebrovascular Research Fellowship at Mount Sinai, she was granted the Mount Sinai Institutional KL2 Clinical and Translational Research Career Development PhD Candidate award.

MICHELLE CROUTHAMEL, D.B.A., is an industry thought leader in digital health with a broad spectrum of R&D experience spanning across drug discovery, clinical development, project management, and digital health initiatives.

Her passion for driving patient-centricity/precision medicine in the pharmaceutical industry led her to become an 'Intrapreneur' at GSK, and the founding member of the Clinical Innovation Unit, which focuses on harnessing digital technologies and platforms to improve clinical evidence generation and optimize trial operation. Michelle is currently the Director of Digital Health & Innovation at AbbVie leading digital health strategy and implementation. She is also involved in the industry-wide consortium, including TransCelerate, IMI and DiMe. Over the past 15+ years, Michelle led many successful programs in discovery and development. She is an inventor who holds multiple patents and has published extensively in the areas of Neuroscience, Oncology, and Digital Health.

Michelle has a Bachelor's degree in Nursing, a Master's degree from the Institute of Neuroscience, and a Doctorate from the Fox Business School, Temple University with a research focus on measuring the firm performance of patient-centricity in the pharmaceutical industry.

LUCA FOSCHINI, M.S., PH.D., is the Co-founder and Chief Data Scientist at Evidation Health, responsible for data analytics and research and development. At Evidation he has driven research collaborations resulting in numerous publications in the fields of machine learning, behavioral economics, and medical informatics. Previously, Luca held research positions in industry and academic institutions, including Ask.com, Google, ETH Zurich, and UC Santa Barbara. He has co-authored several papers and patents on efficient algorithms for partitioning and detecting anomalies in massive networks. Luca holds MS and PhD degrees in Computer Science from UC Santa Barbara, and ME and BE degrees from the Sant'Anna School of Pisa, Italy.

SEAN KHOZIN, M.D., M.P.H., is an oncologist, physician-scientist, and research affiliate at MIT. Dr. Khozin is the Global Head of Data Strategy for Janssen/J&J, focusing on incorporation of data science and advanced quantitative methods (including AI/ML) into R&D activities. He joined the company from US FDA's Oncology Center of Excellence, where he built and led the Center's bioinformatics capabilities and efforts. He was also the founder of Information Exchange and Data Transformation (INFORMED), US FDA's first data science and technology incubator for de-risking solutions through internal R&D and strategic partnerships for improving global biomedical research and advancing national public health priorities. Prior to his tenure in federal government, Dr. Khozin was the cofounder of Hello Health, developing an integrated telemedicine, point-of-care data visualization, and analytical platform for optimizing patient care and clinical research. The company's core technology offerings were first operationalized in a multidisciplinary network of clinics called SKMD that he founded and for which he served as the Chief Medical Officer.

RITA KAPUR, PH.D., is the Head of Digital Biomarkers at Verily Life Sciences (formerly Google Life Sciences), a translational research and engineering organization focused on improving healthcare by applying scientific and technological advances to significant problems in health and biology. She serves as a cross-functional lead across hardware, software, clinical operations and data science to develop and implement initiatives that leverage wearable and passive sensing technology to help better diagnose, monitor, and intervene in disease. Dr. Kapur received a Bachelor's Degree (cum laude) in Human Biology from Stanford University, and a Doctorate in Neuroscience from the University of California, San Francisco, where she specialized in using in vivo awake behaving electrophysiology and signal processing to study the brain systems underlying reward and learning. Prior to joining Verily, she served as a Senior Clinical Research Scientist focused on the analysis of biosensor (electrocorticographic) and clinical trial data to provide support for physicians in selecting, implanting, and optimizing therapy with an implantable closed-loop brain stimulator for the treatment of epilepsy.

CHRISTOPHER LEPTAK, M.D., PH.D., completed his MD and PhD in microbiology/immunology at UCSF. After residency in Emergency Medicine at Harvard's combined Mass General and Brigham program, he joined FDA in 2007 as a primary reviewer in OND's division of gastroenterology products, focusing on immunomodulators for inflammatory bowel diseases. In 2010, he joined OND's Guidance and Policy Team and became OND's Biomarker and Companion Diagnostics Lead. His focus is on biomarker and diagnostic device utility in clinical trials and drug development, both for drug-specific programs. Chris is the Director of CDER's Biomarker Qualification Program which aims to improve regulatory consistency and policy development in areas of emerging science and technology.

CHRIS LUNT, is the Chief Technology Officer for the All of Us Research Program at the National Institutes of Health. All of Us is effort to build a national, large-scale research enterprise with one million or more volunteers to extend precision medicine to all diseases. He has 20+ years of experience designing web services and data platforms. He joined the NIH from GetInsured, where he served as the Vice President (VP) of Government Solutions. There, he worked with the federal government, states, and the vendor community to improve health insurance shopping and enrollment systems. He also worked as an HHS entrepreneur for the Department of Health and Human Services. Earlier in his career he led an IPO, and invented more than 10 social networking patents now owned by Facebook.

CAMILLE NEBEKER, Ed.D, **M.S.,** is an Associate Professor of Behavioral Medicine in the Department of Family Medicine & Public Health, School of Medicine, UC San Diego. Her research and teaching focus on two intersecting area: 1- community research capacity building (e.g., citizen science and community engaged research) and, 2- digital health research ethics (e.g., consent, privacy expectations, data management). She co-founded and directs the Research Center for Optimal Digital Ethics (ReCODE.Health) and leads the Building Research Integrity and Capacity BRIC programs and the Connected and Open Research Ethics CORE initiative. Dr. Nebeker's research has received continuous support from government, foundation and industry sources since 2002.

SALLY OKUN, R.N., M.M.H.S., joined UnitedHealth Group Research and Development (UHG R&D) in 2020 to focus on policy and ethics with emphasis on patient and consumer participation in care, research and policy. Prior to joining UHG R&D Sally was the VP for Policy and Ethics at PatientsLikeMe. In her 12-year tenure she led the development of the company's health data integrity, patient voice taxonomy, drug safety and pharmacovigilance monitoring platform, Research Collaboration Agreement with the FDA and the Ethics and Compliance Advisory Board. Sally advances the science of patient participation and integration of patient perspective into diverse health policy initiatives at the national and global level. She is a member of numerous expert and advisory boards including the National Academy of Medicine's Leadership Consortium on Science and Value Driven Healthcare; Public Responsibility in Medicine and Research (PRIM&R) Public Policy Committee; Duke Margolis Center for Health Policy Real World Evidence Collaborative Advisory Group; and the International Consortium for Health Outcomes Measurement PROMs National Steering Committee. Prior to joining PatientsLikeMe she practiced as a community-based palliative and end of life care specialist. Sally completed her graduate studies at The Heller School for Social Policy & Management at Brandeis University. She was a 2010 Fellow in Biomedical Informatics for the National Library of Medicine and a 2014 Salzburg Global Fellow in New Paradigms for Behavioral and Mental Health.

LARSSON OMBERG, PH.D., as the Vice President of Systems Biology, oversees a research agenda that focuses both on genomics and participant centered research where data is being collected using remote sensors and mobile phones. The group focuses heavily on using open and team based science to get a large number of external partners to collaborate on data intensive problems. Dr. Omberg has a background in computational biology and has been developing computational methods for genomics analysis and disease modeling. Dr Omberg obtained a MSc in Engineering Physics from the Royal Institute of Technology in Stockholm Sweden and a PhD. from the University of Texas at Austin in Physics before performing a postdoctoral fellowship in Computational Biology and Biostatistics at Cornell University.

ERIC PERAKSLIS, PH.D., is a Rubenstein Fellow at Duke University, where his work focuses on collaborative efforts in data science that span medicine, policy, engineering, data science, information technology, privacy and security. Eric is also Lecturer in the Department of Biomedical Informatics at Harvard Medical School, and Strategic Innovation Advisor to Médecins Sans Fontieres. Prior to his current role, Eric served as Chief Science Officer at Datavant and was Senior Vice President and Head of the Takeda R&D Data Science Institute. Prior to Takeda, Eric was the Executive Director of the Center for Biomedical Informatics and the Countway Library of Medicine. Prior to HMS, Eric served as Chief Information Officer and Chief Scientists (informatics) at the U.S. Food and Drug Administration. In this role, Eric, authored the first IT Strategic Plan for FDA and was responsible for modernizing and enhancing the IT capabilities as well as in silico scientific capabilities at FDA. Prior to his time at FDA, Eric was Senior Vice President of R&D Information Technology at Johnson & Johnson Pharmaceuticals R&D. Eric has a Ph.D. in chemical and biochemical engineering from Drexel University. He also holds BSChE and MS degrees in chemical engineering.

EDMONDO ROBINSON, M.D., M.B.A., F.A.C.P., serves as the Senior Vice President and Chief Digital Innovation Officer for Moffitt Cancer Center. Dr. Robinson is responsible for expanding Moffitt's ecosystem from within and outside of health care to deliver on consumer-oriented, real-world solutions for clinical practice, research and administrative processes essential to support growth and competitive advantage. He also oversees Moffitt's portfolio of digital innovation, including the development and commercialization of health products, tools and technology. With this role, Moffitt aims to create and test new services, programs, partnerships and technologies that leverage digital innovations, while challenging the status quo to reduce the cost of care, improve quality, increase access to care, and enhance the patient experience. Previously, Dr. Robinson was the Chief Transformation Officer and Senior Vice President of Consumerism at ChristianaCare, one of the largest health systems in the mid-Atlantic. He was responsible for the transformation of health care delivery to advance population health initiatives and the move from volume-based to value-based care with a special focus on developing and managing ChristianaCare's consumerism and digital strategies. Dr. Robinson is an associate professor of medicine at Thomas Jefferson University's Sidney Kimmel Medical College an adjunct senior fellow in the Leonard Davis Institute of Health Economics at the University of Pennsylvania. He is also a fellow of the American College of Physicians and a senior fellow of the Society of Hospital Medicine. He holds a medical degree from the David Geffen School of Medicine at the University of California, Los Angeles; an MBA with an emphasis in health care management from the Wharton School at the University of Pennsylvania; and a master's degree in health policy research also from the University of Pennsylvania.

CHRISTINA SILCOX, P.H.D., is a Managing Associate at the Duke-Margolis Center for Health Policy, working on policy solutions to advance innovation in health and health care and improve regulation, reimbursement, and long-term evaluation of medical products.

Dr. Silcox’s portfolio includes multiple areas in digital health policy and real-world evidence, with a focus on medical devices. Currently, she is concentrating on challenges to regulating and adopting of artificial intelligence-enabled software as a medical device, using mHealth to collect real-world data, and characterizing real-world data quality and relevancy. Her projects have included the use of patient-generated health data in medical device evaluations, the exploration of value-based payments for medical devices, and the convening the National Evaluation System for health Technology (NEST) Planning Board.

Before she joined Duke-Margolis, Dr. Silcox was a senior fellow at the National Center for Health Research, focused on federal regulation of and policies for medical products. She earned a M.S. from the Massachusetts Institute of Technology (MIT) in Electrical Engineering and a Ph.D. in Medical Engineering and Medical Physics from the Harvard-MIT Division of Health Sciences and Technology (HST).

ALICIA STALEY, M.B.A., M.S.I.S., is the Sr. Director Patient Engagement for mHealth at Medidata. She has over 20 years of experience in software design and information systems management and works to infuse the patient perspective throughout the product development lifecycle and help engage patients in novel ways.Alicia is also a three-time cancer survivor, first diagnosed with Hodgkin’s disease as a sophomore during college. Over the past 10 years, she has applied her engineering background to improve the patient experience for those dealing with cancer. With an extensive network of patient advocates and non-profit organizations, she collaborates with a wide range of stakeholders to improve processes and policies that affect patient care and clinical trials. She has co-led several research studies on how patients share information in online forums and seek out clinical trial opportunities. An early adopter of social media, she co-founded #BCSM, which attracts over 250 global participants each week to its scheduled online discussions. This foundational online social media support channel is recognized as the gold standard for disease-specific social media networks. Prior to joining Medidata, Alicia worked at Cure Forward and Science 37 leading their patient recruitment and engagement initiatives to help advance clinical research. As a champion of patient advocacy and engagement, she understands the critical issues facing patients looking to engage in clinical research. With a keen focus on improving access to clinical trials, Alicia is passionate about making a difference for all patients searching for information about clinical trials.

Nebeker *et al. BMC Medicine* (2019) 17:137
https://doi.org/10.1186/s12916-019-1377-7

BMC Medicine

Beyond Big Data to new Biomedical and Health Data Science: moving to next century precision health

OPINION **Open Access**

Building the case for actionable ethics in digital health research supported by artificial intelligence

Camille Nebeker[1,2*], John Torous[3] and Rebecca J. Bartlett Ellis[4]

Abstract

The digital revolution is disrupting the ways in which health research is conducted, and subsequently, changing healthcare. Direct-to-consumer wellness products and mobile apps, pervasive sensor technologies and access to social network data offer exciting opportunities for researchers to passively observe and/or track patients 'in the wild' and 24/7. The volume of granular personal health data gathered using these technologies is unprecedented, and is increasingly leveraged to inform personalized health promotion and disease treatment interventions. The use of artificial intelligence in the health sector is also increasing. Although rich with potential, the digital health ecosystem presents new ethical challenges for those making decisions about the selection, testing, implementation and evaluation of technologies for use in healthcare. As the 'Wild West' of digital health research unfolds, it is important to recognize who is involved, and identify how each party can and should take responsibility to advance the ethical practices of this work. While not a comprehensive review, we describe the landscape, identify gaps to be addressed, and offer recommendations as to how stakeholders can and should take responsibility to advance socially responsible digital health research.

Keywords: Research ethics, Bioethics, Digital health, Digital medicine, Artificial intelligence, Precision medicine

Background

The digital revolution is disrupting the ways in which health research is conducted, and subsequently, changing healthcare [1–3]. The rise of digital health technologies has resulted in vast quantities of both qualitative and quantitative 'big data', which contain valuable information about user interactions and transactions that may potentially benefit patients and caregivers [4]. Digital data 'exhaust', or the traces of everyday behaviors captured in our digital experiences, are of particular interest because they contain our natural behaviors gathered in real time. No doubt, important societal conversations are needed to shape how these sociotechnical systems influence our lives as individuals, as well as the impact on society [5]. While not a formal review, this opinion essay provides a selective overview of the rapidly changing digital health research landscape, identifies gaps, highlights several efforts that are underway to address these gaps, and concludes with recommendations as to how stakeholders can and should take responsibility to advance socially responsible digital health research.

Direct-to-consumer wellness products and mobile apps (e.g., Fitbit, Strava), wearable research tools (e.g., SenseCam, ActivPAL), and access to social network data offer exciting opportunities for individuals [6], as well as traditional health researchers [7], to passively observe and/or track individual behavior 'in the wild' and 24/7. The volume of granular personal health data gathered using these technologies is unprecedented, and is increasingly leveraged to inform personalized health promotion and disease treatment interventions. The use of artificial intelligence (AI) tools in the health sector is also increasing. For example, electronic health records provide training data for machine learning that inform algorithms, which can detect anomalies more accurately

---

* Correspondence: nebeker@eng.ucsd.edu
[1]Department of Family Medicine and Public Health, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA
[2]Research Center for Optimal Digital Ethics in Health, Qualcomm Institute and School of Medicine, University of California San Diego, La Jolla, CA 92093, USA
Full list of author information is available at the end of the article

than trained humans – particularly in the fields of cancer, cardiology, and retinopathy [8]. The digital therapeutics sector is also seeking to expand and bring products into the healthcare system, with the goal of complementing or providing an alternative to traditional medical treatments [9]. While the digital health revolution brings transformational promise for improving healthcare, we must acknowledge our collective responsibility to recognize and prevent unintended consequences introduced by biased and opaque algorithms that could exacerbate health disparities and jeopardize public trust [10, 11]. Moreover, it is critical that the minimal requirements used to make a digital health technology available to the public are not mistaken for a product that has passed rigorous testing or demonstrated real world therapeutic value [12].

Although rich with potential, the digital health ecosystem presents new ethical challenges for those making decisions about the selection, testing, implementation and evaluation of technologies in healthcare. Researchers began to study related ethical issues over 20 years ago, when electronic health records technology was being conceptualized [13], and as new forms of pervasive information communication technologies produce data, guiding principles and standards are emerging within academic research centers [14–16] and industry sectors [17, 18]. Accepted ethical principles in health research, including respect for persons, beneficence and justice, remain relevant and must be prioritized to ensure that research participants are protected from harms. Applying these principles in practice means that: people will have the information they need to make an informed choice; risks of harm will be evaluated against potential benefits and managed; and no one group of people will bear the burden of testing new health information technologies [19]. However, ethical challenges arise from the combination of new, rapidly evolving technologies; new stakeholders (e.g. technology giants, digital therapeutic start-ups, citizen scientists); data quantity; novel computational and analytic techniques; and a lack of regulatory controls or common standards to guide this convergence in the health ecosystem.

It is of particular importance that these technologies are finding their way into both research and clinical practice without appropriate vetting. For example, we have heard that, "if the product is free, then you're the product." This means that our search terms, swipes, clicks and keyboard interactions produce the data that companies use to inform product improvement. These 'big data' are used to train algorithms to produce, for example, tailored advertisements. Consumers allow this by clicking "I Accept" to confirm their agreement with the Terms and Conditions (T&C), which are not necessarily intended to be easy to read or understand. Why does this matter? When an algorithm is used to serve up a reminder about that yellow jacket you were eyeing, or the summer vacation you mentioned to a friend the other day, it may seem 'creepy', but it might be nice in terms of convenience. Sometimes the AI gets it right, and other times it is not even close. For example, if you were to write something on Facebook that its proprietary AI interprets as putting you at serious risk, it may send the police to your home! Is Facebook getting it right? We do not know: Facebook has claimed that, even though its algorithm is not perfect and makes mistakes, it does not consider its actions to be 'research' [20]. Aside from threats to one's privacy, we should question the process of informed consent, whether there is an objective calculation of risk of harms against potential benefits, and whether people included in the product testing phase are those most likely to benefit.

Governance in the 'wild west'

Those involved in the development, testing and deployment of technologies used in the digital health research sector include technology developers or 'tool makers', funders, researchers, research participants and journal editors. As the 'Wild West' of digital health research moves forward, it is important to recognize who is involved, and to identify how each party can and should take responsibility to advance the ethical practices of this work.

Who is involved?

In the twentieth century, research was carried out by scientists and engineers affiliated with academic institutions in tightly controlled environments. Today, biomedical and behavioral research is still carried out by trained academic researchers; however, they are now joined by technology giants, startup companies, nonprofit organizations, and everyday citizens (e.g. do-it-yourself, quantified self). The biomedical research sector is now very different, but the lines are also blurred because the kind of product research carried out by the technology industry has, historically, not had to follow the same rules to protect research participants. As a result, there is potential for elevated risks of harm. Moreover, how and whether research is carried out to assess a product's effectiveness is variable in terms of standards and methods, and, when the technology has health implications, standards become critically important. In addition, not all persons who initiate research are regulated or professionally trained to design studies. Specific to regulations, academic research environments require the involvement of an ethics board (known as an institutional review board [IRB] in the USA, and a research ethics committee [REC] in the UK and European Union). The IRB review is a federal mandate for entities that receive US federal funding to conduct health research. The ethics review is a peer review process to evaluate proposed research, and identify and reduce potential risks that research participants may experience. Having an objective peer review process is not a requirement for technology giants, startup companies or by those who identify with the citizen science community [10, 21]; however, we have a societal responsibility to get this right.

What questions should be asked?

When using digital health technologies, a first step is to ask whether the tools, be they apps or sensors or AI applied to large data sets, have demonstrated value with respect to outcomes. Are they clinically effective? Do they measure what they purport to measure (validity) consistently (reliability)? For example, a recent review of the predictive validity of models for suicide attempts and death found that most are currently less than 1%; a number at which they are not yet deemed to be clinical viable [22]. Will these innovations also improve access to those at highest risk of health disparities? To answer these questions, it is critical that all involved in the digital health ecosystem do their part to ensure the technologies are designed and scientifically tested in keeping with accepted ethical principles; be considerate of privacy, effectiveness, accessibility, utility; and have sound data management practices. However, government agencies, professional associations, technology developers, academic researchers, technology startups, public organizations and municipalities may be unaware of what questions to ask, including how to evaluate new technologies. In addition, not all tools being used in the digital health ecosystem undergo rigorous testing, which places the public at risk of being exposed to untested and potentially flawed technologies.

Demonstrating value must be a precursor to the use of any technologies that claim to improve clinical treatment or population health. Value is based on the product being valid and reliable, which means that scientific research is needed before a product is deployed within the health sector [12]. We should also not move ahead assuming that privacy and the technology revolution are mutually exclusive. We are in a precarious position in which, without standards to shape acceptable and ethical practices, we collectively run the risk of harming those who stand to benefit most from digital health tools.

Decision-making framework

While there are discussions about the need for regulations and laws, and incremental progress being made on that front, until some consensus is reached, it is essential that stakeholders recognize their obligation to promote the integrity of digital health research [23]. The digital health decision-making domains framework (Fig. 1) was developed to help researchers make sound decisions when selecting digital technologies for use in health research [24, 25]. While originally developed for researchers, this framework is applicable to various stakeholders who might evaluate and select digital technologies for use in health research and healthcare. The framework comprises five domains: 1, Participant Privacy; 2 Risks and Benefits; 3, Access and Usability; 4, Data Management; and 5, Ethical Principles. These five domains are presented as intersecting relationships.

The domains in this framework were developed into a checklist tool to further facilitate decision-making. The checklist was informed via developmental research involving a focus group discussion, and a design exercise with behavioral scientists [25]. To demonstrate how the decision-making domains can be put into practice, we present a use case to illustrate the complexities and nuances that are important for stakeholders to consider.

Use case: MoodFlex for mental health

MoodFlex is a private startup technology company that has developed a mobile app to detect signals of poor mental health by analyzing a person's typing and voice patterns from their smartphones. MoodFlex is negotiating with several municipalities to integrate their product within the public mental healthcare system, with the goal of delivering better services to people with mental illness through predictive analytics. Since MoodFlex does not claim to provide a clinical diagnosis or treatment, approval from the US Food and Drug Administration is not necessary. The vendor claims to have a proven product; however, there are no publications documenting evidence that it is safe, valid or reliable. The only research that is formally acknowledged involves an evaluation of the implementation process and uptake of the product by health providers within the state mental health system. The patient will be invited to download the app after reviewing the vendor's T&C – no other consent process is proposed. The algorithm is proprietary, and therefore, an external body is unable to determine whether the algorithm that resulted from a machine-learning process was trained on representative data, or how decision-making occurs. Data captured about people using the app are owned by the vendor.

Brief analysis

Before introducing MoodFlex into the public healthcare system, decision makers – particularly the funding organization – should evaluate evidence supporting the efficacy of this product. Reproducible evidence is the hallmark of evidence-based practice, and is the first step prior to dissemination and implementation. If a product is supported by evidence, the logical next step is the

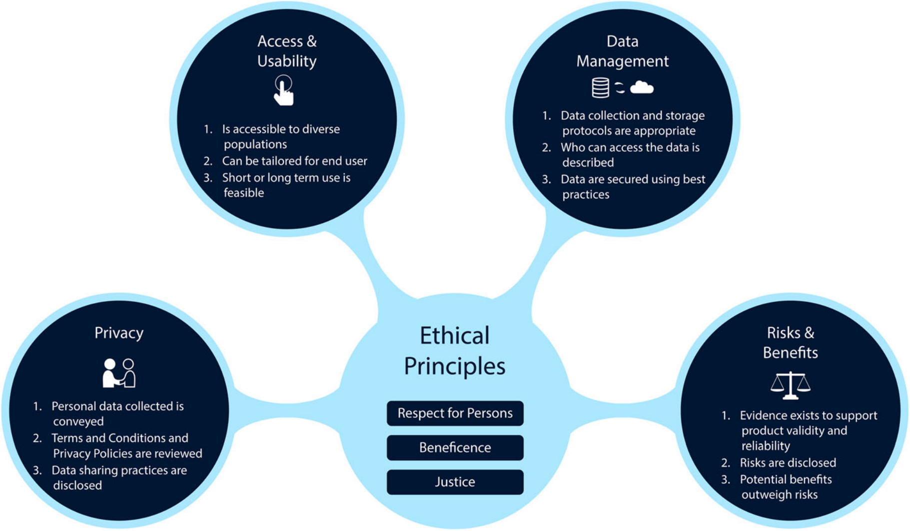


Fig. 1 Digital health decision-making framework and excerpts from the companion checklist designed to support researchers [24]

translational phase, in which a 'dissemination and implementation' (D&I) design is appropriate. Unfortunately, many health apps move straight into a D&I phase before the evidence exists to support that direction.

Lacking evidence that the product is effective, decision-makers should recognize that a testing phase is necessary. As with regulated research involving people, a research plan should be developed and reviewed by an external and objective ethics board (i.e., REC or IRB) that will assess the degree to which people who are invited do not bear an inappropriate burden (justice), potential risks are offset by the benefits (beneficence), and individuals are provided with an ability to make an informed choice to volunteer (respect). At this early stage, it is reasonable for the vendor to provide the sponsor with a robust data management plan, with explicit language regarding data ownership, access, sharing and monitoring. When involving vulnerable populations, such as those with a mental health diagnosis, additional precautions should be considered to ensure that those involved in the study are protected from harms – including stigma, economic and legal implications. In addition, it is important to consider whether some people will be excluded because of access barriers. For example, it may be necessary to adapt the technology to be useful to non-English speakers. Informed consent must also be obtained in a way that results in a person making a choice to participate based on having adequate and accessible information – this demonstrates the principle of 'respect for persons', and is a hallmark of research ethics. Placing consent language for a research study in the T&C is unacceptable. For patients who become research participants, it is particularly important for them to understand the extent to which the technology will support their healthcare needs. Patients might falsely rely on the technology to provide the care they believe they need when, in reality, they may need to see their healthcare provider.

Digital research gaps and opportunities

This use case reflects the shift in health research associated with digital technologies, in that traditional methods of developing an evidence base may be pushed aside in favor of what appears to be exciting innovation. The landscape is unsettled and potentially dangerous, which makes governance important. We have identified three notable gaps: 1, disciplinary/sector challenges; 2, issues of data and technology literacy; and 3, inconsistent or non-extant standards to guide the use of AI and other emerging technologies in the healthcare settings.

Inter/trans/cross-disciplinary and sector challenges

Emerging technologies and AI systems require diverse expertise when applied to digital medicine, which introduces new challenges. Technology makers may not understand patients' needs, and develop tools with limited utility in practice [25, 26]. Computational scientists may train AI using datasets that are not representative of the public, limiting the ability to provide meaningful assessments or predictions [27]. Clinicians may not know how to manage the depth of granular data, nor be confident in decisions produced by AI [28]. Research is needed to examine this disconnect, and identify strategies to reduce gaps and improve meaningful connections between these groups that are integral to digital health research and the use of AI in the health care sector.

Digital/tech-literacy

The idea that keystrokes and voice patterns can be used to aid diagnosis of Parkinson's disease remains impressive, but now it may also be possible to use keystroke dynamics, kinematics and voice patterns to detect mental health problems [29]. Knowing this information may create public concern if not communicated in a way that is useful and contextual, adding to fear, skepticism and mistrust. The 'public' includes policy-makers, educators, regulators, science communicators, and those in our healthcare system, including clinicians, patients, and caregivers. Research is needed to increase our understanding of what these stakeholders know, what they want to know, and how best to increase their technology literacy. This information can then be used to inform educational resources targeting specific stakeholders. For example, when reviewing manuscripts reporting digital health research, reviewers and editors should be aware of how to evaluate new methodologies and computational analytics to verify the accuracy and appropriateness of the research and results.

Ethical and regulatory standards

As new digital tools and AI-enabled technologies are developed for the healthcare market, they will need to be tested with people. As with any research involving human participants, the ethics review process is critical. Yet, our regulatory bodies (e.g., IRB) may not have the experience or knowledge needed to conduct a risk assessment to evaluate the probability or magnitude of potential harms [30]. Technologists and data scientists who are making the tools and training the algorithms may not have received ethics education as part of their formal training, which may lead to a lack of awareness regarding privacy concerns, risks assessment, usability, and societal impact. They may also not be familiar with regulatory requirements to protect research participants [23]. Similarly, the training data used to inform the algorithm development are often not considered to qualify as human subjects research, which – even in a regulated environment – makes a prospective review for safety potentially unavailable.

New initiatives – what resources are available for the digital health/medicine community?

Several initiatives have begun to address the ethical, legal and social implications (ELSI) of the digital revolution in healthcare. Prominent examples of such initiatives concern AI. Specific to AI, the foci are broad, and include autonomous vehicles, facial recognition, city planning, the future of work, and in some cases, health. A few selected examples of current AI efforts appear to be well-funded and collaborative programs (see Table 1).

Across these initiatives are efforts to assess the potential ELSI of AI. Similar to the impact of the European Union (EU)'s General Data Protection Regulation (GDPR) in countries beyond the EU, the intention of groups assessing AI through an ELSI lens is to develop standards that can be applied or adapted globally. In practice, however, most current efforts to integrate ELSI to AI are quite broad, and as a result, may overlap in scope and lack specificity.

While AI has a place in the digital health revolutions, the scope of technologies goes well beyond AI. Other

Table 1 AI initiatives underway to inform broad cross-sector standards

Program	Goal	Collaborators
The Partnership on AI [30]	Develop/test and share best practices	80+ partners in 13 countries
AI-100 [31]	Impact of AI on urban life by 2030 in North America	E. Horvitz, R. Altman
Ethics and Governance of AI Fund [32]	Conduct evidence-based research	Berkman Klein Center, Harvard and MIT Media Lab
AI Now Institute [33]	Conduct evidence-based research	New York University
Initiative on Ethics of Autonomous and Intelligent Systems [34]	Develop standards, certifications, codes	IEEE and ACM
Human Rights, Big Data and Technology Project [35]	Analyze the use of big data, artificial intelligence, associated technologies	University of Essex, United Nations
The Institute for Ethics in Artificial Intelligence [36]	Explore fundamental issues affecting the use and impact of AI	Technical University of Munich partnership with Facebook
High-Level Expert Group on Artificial Intelligence [37]	Recommend ELSI policy development on AI	European Commission
Chinese Association for Artificial Intelligence [38]	Unite artificial intelligence science and technology professionals	Ministry of Civil Affairs, China
AI for Humanity [39]	Create an international group of AI experts to prepare for societal transformation	Future of Life Institute, France

initiatives are looking more specifically at ELSI in mobile apps, social network platforms, and wearable sensors being used in digital research. These include, for example, the Connected and Open Research Ethics (CORE) initiative at the University of California (UC) San Diego Research Center for Optimal Digital Ethics in Health (ReCODE Health), the Pervasive Data Ethics for Computational Research (PERVADE) program at the University of Maryland, and the Mobile Health ELSI (mHealthELSI) project out of Sage Bionetworks and the University of Louisville. What these initiatives have in common is a goal to inform policy and governance in a largely unregulated space. These initiatives are but a few examples, and it is important to note that many laboratories and institutes are working on digital health ELSI.

Conclusion

Being mindful of new health technologies with new actors in the arena, the gap between known and unknown risks fundamentally challenges the degree to which decision-makers can properly evaluate the probability and magnitude of potential harms against benefits. Now is the time to take a step back and develop the infrastructure necessary for vetting new digital health technologies, including AI, before deploying them into our healthcare system. Selecting and implementing technologies in the digital health ecosystem requires consideration of ethical principles, risks and benefits, privacy, access and usability, and data management. New technologies have the potential to add important value; however, without careful vetting, may exacerbate health disparities among those most vulnerable.

Abbreviations

AI: Artificial intelligence; ELSI: Ethical, legal and social implications; IRB: Institutional review board; REC: Research ethics committee

Acknowledgements

Not applicable.

Authors' contributions

CN conceptualized and prepared the original manuscript; JT, RBE and CN wrote, reviewed and edited the manuscript. All authors read and approved the final version of the manuscript.

Authors' information

CN directs the Research Center for Optimal Digital Ethics (ReCODE.health) at the University of California at San Diego. Her appointment is in the Division of Behavioral Medicine, in the Department of Family Medicine and Public Health. JT directs the Digital Psychiatry Division at Beth Israel Deaconess Medical Center, and is a licensed psychiatrist and clinical informaticist. RJBE is a behavioral scientist at Indiana University who uses sensors and mobile technologies to study biobehavioral mechanisms related to medication adherence and treatment efficacy.

Funding

This work was supported, in part, by the Robert Wood Johnson Foundation (grant number 72876, to CN, Principal Investigator [PI] 2015–2019), and by the National Institutes of Mental Health (grant number 1K23MH116130–01, to JT, PI 2018–2022).

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflicts of interest.

Author details

[1]Department of Family Medicine and Public Health, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA. [2]Research Center for Optimal Digital Ethics in Health, Qualcomm Institute and School of Medicine, University of California San Diego, La Jolla, CA 92093, USA. [3]Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave, Boston, MA 02215, USA. [4]School of Nursing, Indiana University, 600 Barnhill Drive, Indianapolis, IN 46202, USA.

Received: 4 April 2019 Accepted: 26 June 2019

References

1. Agarwal R, Gao G, DesRoches C, Jha AK. Research commentary – the digital transformation of healthcare: current status and the road ahead. Inform Syst Res. 2010;21(4):796–809 http://pubsonline.informs.org/doi/abs/10.1287/isre.1100.0327. Accessed 15 May 2019.
2. Herrmann M, Boehme P, Mondritzki T, Ehlers JP, Kavadias S, Truebel H. Digital transformation and disruption of the health care sector: internet-based observational study. J Med Internet Res. 2018;20(3):e104.
3. Wallin A. Transforming healthcare through entrepreneurial innovations. Int J E-Services Mob Appl. 2017;9(1):1–17 https://www.igi-global.com/gateway/article/173033. Accessed 2019 May 15. https://doi.org/10.4018/IJESMA.2017010101
4. Lupton D. The commodification of patient opinion: the digital patient experience economy in the age of big data. Sociol Health Illn. 2014;36(6):856–69.
5. Boyd D, Crawford K. Critical questions for big data. Inform Commun Soc. 2012;15(5):662–79.
6. Lupton D. Self-tracking, health and medicine. Health Sociol Rev. 2017;26(1):1–5.
7. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. J Behav Med. 2019;42(1):67–83.
8. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4):230–43.
9. Joyce M, Leclerc O, Westhues K, Xue H. Digital therapeutics: preparing for takeoff. New York: McKinsey & Company; 2018. https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/digital-therapeutics-preparing-for-takeoff. Accessed 21 Mar 2019
10. Nebeker C, Harlow J, Giacinto-Espinoza R, Linares-Orozco R, Bloss C, Weibel N. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. Am J Bioeth Empir Bioeth AJOB Empir Bioeth. 2017;8(4):266–76.
11. Ferryman K, Pitcan M. Fairness in precision medicine. New York: Data & Society; 2018. https://datasociety.net/wp-content/uploads/2018/02/Data.Society.Fairness.In_.Precision.Medicine.Feb2018.FINAL-2.26.18.pdf. Accessed 13 May 2019
12. Coravos A, Goldsack JC, Karlin DR, Nebeker C, Perakslis E, Zimmerman N, et al. Digital medicine: a primer on measurement. Digit Biomarkers. 2019;3(2):31–71 https://www.karger.com/Article/FullText/500413. Accessed 13 May 2019.
13. Kluge EH. Health information, the fair information principles and ethics. Methods Inf Med. 1994;33(04):336–45.
14. Pimple KD. Emerging pervasive information and communication tTechnologies (PICT): ethical challenges, opportunities and safeguards. Dordrecht: Springer Netherlands; 2013.

15. Vitak J, Proferes N, Shilton K, Ashktorab Z. Ethics regulation in social computing research: examining the role of institutional review boards. J Empir Res Hum Res Ethics. 2017;12(5):372–82.
16. Van Velthoven MH, Smith J, Wells G, Brindley D. Digital health app development standards: a systematic review protocol. BMJ Open. 2018;8(8): e022969.
17. Dittrich D, Kenneally E. The Menlo report: ethical principles guiding information and communication technology research. Washington, DC: US Department of Homeland Security; 2012. https://www.caida.org/publications/papers/2012/menlo_report_actual_formatted/menlo_report_actual_formatted.pdf. Accessed 14 May 2019
18. Jackman M, Kanerva L. Evolving the IRB: building robust review for industry research. Washington Lee Law Rev Online. 2016;72(3) https://scholarlycommons.law.wlu.edu/wlulr-online/vol72/iss3/8. Accessed 14 May 2019.
19. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont report: ethical principles and guidelines for the protection of human subjects of research, vol. 44. Washington, DC: Department of Health, Education, and Welfare, US Department of Health and Human Services; 1979. p. 23192–7. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html. Accessed 14 May 2019
20. Barnett I, Torous J. Ethics, transparency, and public health at the intersection of innovation and Facebook's suicide prevention efforts. Ann Intern Med. 2019;170(8):565.
21. Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, et al. Reimagining human research protections for 21st century science. J Med Internet Res. 2016;18(12):e329.
22. Belsher BE, Smolenski DJ, Pruitt LD, Bush NE, Beech EH, Workman DE, et al. Prediction models for suicide attempts and deaths: a systematic review and simulation. JAMA Psychiatry. 2019. https://doi.org/10.1001/jamapsychiatry.2019.0174.
23. Pagoto S, Nebeker C. How scientists can take the lead in establishing ethical practices for social media research. J Am Med Informatics Assoc. 2019;26(4): 311–3.
24. Nebeker C, Bartlett Ellis R, Torous J. CORE tools. Digital health decision-making framework and checklist designed for researchers. 2018. https://recode.health/dmchecklist/. Accessed 26 Feb 2019.
25. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med. 2019. https://doi.org/10.1093/tbm/ibz074.
26. Wang S, Bolling K, Mao W, Reichstadt J, Jeste D, Kim H-C, et al. Technology to support aging in place: the older adult perspective. Healthcare. 2019. https://doi.org/10.3390/healthcare7020060.
27. Sears M. AI bias and the 'people factor. In: AI development. Jersey City: Forbes Media; 2018. https://www.forbes.com/sites/marksears1/2018/11/13/ai-bias-and-the-people-factor-in-ai-development/#24ddaaf79134. Accessed 26 Feb 2019.
28. Adibuzzaman M, DeLaurentis P, Hill J, Benneyworth BD. Big data in healthcare – the promises, challenges and opportunities from a research perspective: a case study with a model database. AMIA Annu Symp Proc. 2018;2017:384–92.
29. Rabinovitz J. Your computer may know you have Parkinson's. Shall it tell you? Stanford: Stanford Magazine; 2018. https://medium.com/stanford-magazine/your-computer-may-know-you-have-parkinsons-shall-it-tell-you-e8f8907f4595. Accessed 19 Feb 2019
30. The Partnership on AI. https://www.partnershiponai.org/. Accessed 07 June 2019.
31. AI-100. https://ai100.stanford.edu/. Accessed 07 June 2019.
32. Ethics and Governance of AI Fund. https://cyber.harvard.edu/topics/ethics-and-governance-ai. Accessed 07 June 2019.
33. AI Now Institute. https://ainowinstitute.org/. Accessed 07 June 2019.
34. Initiative on Ethics of Autonomous and Intelligent Systems. https://standards.ieee.org/industry-connections/ec/autonomous-systems.html. Accessed 07 June 2019.
35. Human Rights, Big Data and Technology Project. https://hrbdt.ac.uk/. Accessed 07 June 2019.
36. The Institute for Ethics in Artificial Intelligence. 2019. https://www.tum.de/en/about-tum/news/press-releases/details/article/35190. Accessed 07 June 2019.
37. High-Level Expert Group on Artificial Intelligence. https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence. Accessed 07 June 2019.
38. Chinese Association for Artificial Intelligence. https://www.iotone.com/organization/chinese-association-for-artificial-intelligence-caai/o209. Accessed 07 June 2019.
39. AI for Humanity. https://www.aiforhumanity.fr/en/. Accessed 07 June 2019.

Publisher's Note

TBM

ORIGINAL RESEARCH

Development of a decision-making checklist tool to support technology selection in digital health research

Camille Nebeker,[1,2] Rebecca J. Bartlett Ellis,[3] John Torous[4]

[1]Department of Family Medicine and Public Health, School of Medicine, UC San Diego, La Jolla, CA 92093, USA
[2]Center for Wireless and Population Health Systems, UC San Diego, La Jolla, CA 92093, USA
[3]Indiana University School of Nursing, Indianapolis, IN 46202, USA
[4]Beth Israel Deconess Medical Center, Harvard University, Boston, MA 02215, USA

Correspondence to: C Nebeker, nebeker@eng.ucsd.edu

Cite this as: *TBM* 2019;XX:XX–XX
doi: 10.1093/tbm/ibz074

Abstract

Digital technologies offer researchers new approaches to test personalized and adaptive health interventions tailored to an individual. Yet, research leveraging technologies to capture personal health data involve technical and ethical consideration during the study design phase. No guidance exists to facilitate responsible digital technology selection for research purposes. A stakeholder-engaged and iterative approach was used to develop, test, and refine a checklist designed to aid researchers in selecting technologies for their research. First, stakeholders ($n = 7$) discussed and informed key decision-making domains to guide app/device selection derived from the American Psychiatric Association's framework that included safety, evidence, usability, and interoperability. We added "ethical principles" to the APA's hierarchical model and created a checklist that was used by a small group of behavioral scientists ($n = 7$). Findings revealed the "ethical principles" domains of respect, beneficence, and justice cut across each decision-making domains and the checklist questions/prompts were revised accordingly and can be found at thecore.ucsd.edu. The refined checklist contains four decision-making domains with prompts/questions and ethical principles embedded within the domains of privacy, risk/benefit, data management, and access/evidence. This checklist is the first step in leading the narrative of decision-making when selecting digital health technologies for research. Given the dynamic and rapidly evolving nature of digital health technology use in research, this tool will need to be further evaluated for usefulness in technology selection.

Keywords

Mobile health, Digital medicine, Decision-making checklist, Research ethics, Tech ethics

Implications
Practice: Technology-enabled tools and strategies are increasingly used in digital health research.

Policy: We lack guidance to make informed choices about the technologies selected for use in digital research and researchers, institutional review boards, policy makers, and consumers should be involved in shaping policy to ensure safe, effective, and responsible practices follow.

Research: Future research is needed to qualify potential risks of harm and against potential benefits of tech-enabled research.

INTRODUCTION

Personal health data are increasingly more accessible to researchers via mobile apps, wearable sensors, social networks, and other emerging technologies. Given the potential for technology to improve personal health and decrease healthcare costs, ethical and regulatory concerns must be embedded early in the technology selection/development and subsequently, in the research design phase to optimize benefit and reduce potential harm risks. For the researcher, digital technologies allow the testing of personalized and adaptive health interventions, which can potentially be generalized to vulnerable populations. Yet, research studies leveraging technologies raise critical and nuanced ethical challenges specific to informed consent, data collection, bystander rights, risk assessment, privacy, and data management [1,2].

Increase in digital health research and lack of associated guidance

A recent analysis of NIH-supported research revealed a 12-fold increase in studies using pervasive sensing technologies at years: 2005, 2010, and 2015 [3]. While these studies accounted for only about 1.7% of the NIH research budget, it is a clear indication of how future research will be designed and implemented into clinical practice. With the increase in tech-enabled research tools, researchers are challenged to design studies which take into account certain risks that may be unknown, particularly around data management and patient privacy concerns. Due to a lack of familiarity with digital health tools, Institutional Review Boards (IRBs) are also grappling with evaluating risks and determining appropriate risk management strategies [2] for these tools. A key challenge to developing guidance is establishing the probability and potential magnitude of harm that a research participant may be exposed to when using wearable and pervasive sensing technologies. While researchers can identify some potential risks in advance and build in risk management strategies, there are also unknown factors that can only be learned in time [4].

Potential risks of using mobile apps and pervasive sensing devices in health research

Despite encouraging early results, the future of digital technologies is actually in jeopardy due to, in part, the involvement of commercial products being used as research tools. The foundation of all healthcare is based on trust. However, consumer trust in technology and the companies developing these new tools is shockingly weak. One industry study from 2015 noted that just 8% of those surveyed from the general public actually trust technology companies with their healthcare data [5]. For instance, well-publicized cases of mobile apps disclosing sensitive reproductive information [6], and concerns about vulnerability of medical devices like pacemakers to digital hacking [7] threaten public willingness to trust new and innovative technology tools.

While commercial apps offer ready to use technology for health researchers, they present challenges upon closer inspection. Commercial apps come with terms and condition of service agreements, which the user must agree to in order to use the app. When used for research, the participant may be asked to agree with the terms and services as part of enrolling in a study. In some cases, corporate agreements require users waive their rights to litigation, which directly conflicts with federal human research protection requirements [2]. Moreover, companies are not bound to protect participant confidentiality nor obtain informed consent in the same manner that health researchers must. Recent reviews of commercial smartphone app agreements reveal that user data are unprotected and open to sharing, selling, and marketing [8]. A review of dementia apps found that less than half of those that collected personal health information offered any written description of how that data are used and shared, and in those that did, 42% of privacy policies were not relevant to the app or personal data collected [9]. A study reviewing diabetes apps found that up to 80% of these apps lacked a privacy policy and the few that existed offered little protections for app users or their personal health information [10]. The lack of protections for health app users prompted the US Department of Health and Human Services to draft a report on the need to offer more consumer protections across the entire digital health platform space [11].

Emerging initiatives to guide ethical practice

There were several attempts to create guidance yet, these attempts failed or remain in nascent stages. The United Kingdom's National Health Service (NHS) created a curated app library to offer consumers safe and useful health apps, but the service was suspended in Fall 2015 when vetted apps were noted to actually have data security flaws and little evidence for utility [12]. A new iteration of the app library has been online for less than 1 year at the time of this writing. Similarly, in December 2016 the American Medical Association house of delegates voted to adopt a series of principles to promote the safe use of health apps, although offered little actual concrete guidance around evaluating these principles (https://www.ama-assn.org/press-center/press-releases/ama-adopts-principles-promote-safe-effective-mhealth-applications). The U.S. Federal Trade Commission [13] released a website to help app developers learn which regulatory statutes their app may fall under. The US Food and Drug Administration is piloting a new regulatory framework for apps and other technologies it classifies as software as a medical device in its Digital Health Software Precertification (Pre-Cert) Program [14].

Individual medical societies such as the American Psychiatric Association (APA) released app-related guidelines for their clinician members [15] and new initiatives are exploring guidance development. For example, Xcertia has engaged the American Medical Association, the American Heart Association and a few other health organizations and professional societies to develop guidelines [16]; however, it is unclear whether or how these guidelines may support the needs of multiple stakeholders within the digital health research ecosystem. The Connected and Open Research Ethics (CORE) initiative is collecting and curating tech-related research ethics resources and has created a virtual global community where over 600 researchers, ethicists, regulators, and participants can discuss emerging issues [17]. However, no guidance exists to facilitate responsible digital technology selection for research purposes. The purpose of this article is to describe the iterative development of a checklist tool to assist researchers in the decision-making process when selecting digital technologies for their research.

METHODS

A two-phase approach was used to develop and test the checklist. In Phase 1, we engaged an expert panel of key stakeholders to discuss the decision-making domains of the checklist and its hierarchical structure. Phase 2 involved deploying a new version of the checklist in a survey format to a small group of behavioral scientists to identify whether the checklist would prove useful to digital health researchers. This formative research was conducted under an exempt protocol approved by the IRB at the University of California San Diego.

Phase 1

To develop a decision-making checklist, critical domains were identified for selecting digital technology used in research. The APA's App Evaluation model [15] was examined as starting point because it applies to selecting apps, which are a common digital technology increasingly being used in mental health clinical care. Noting that little guidance was available to facilitate app selection in clinical psychiatric healthcare, an APA working group developed a decision-making framework in 2016 for clinicians [18]. The APA pyramid-shaped framework features four domains with safety at the base, followed by evidence, usability, and interoperability. The hierarchical nature of the APA model aligns with medical decision-making of first evaluating risks and then benefits. The hierarchical aspect emerged from a clinician's desire for a tool that was simple to apply in clinical practice.

Utilizing the APA's framework to support decision-making in a research context, we added ethical principles found in the Belmont Report [19] (see Fig. 1) to further ground the framework. Our formative research sought to determine whether the APA framework, modified with the addition of ethical principles, as the fifth and foundational domain, could be a useful tool for the digital health "research" ecosystem. Another aim was to confirm whether this framework was hierarchical, as proposed by the APA working group.

Once the initial decision-making domains were identified, we convened a group of key stakeholders to participate on an expert panel to: (a) validate the domains and (b) determine if any of the identified domains were more important than others, lending to the hierarchical structure that was conceptualized by the APA working group.

Phase 1 recruitment

A 2.5-hr working group was convened in June 2017 in New York City composed of digital research experts ($n = 7$) identified by the lead researcher in collaboration with a member of the HUMAN research project team. Those invited included individuals with digital research expertise including clinician/researchers, human research protections, legal and social perspectives. With the permission of attendees, the panel discussion was audio-recorded and a research assistant took notes on a laptop.

Phase 1 procedures

The discussion began with an introduction to digital health research including related ethical issues of informed consent (i.e., technology literacy, data literacy), bystander rights, and data management [2]. A use-case was then distributed along with an initial draft of the checklist which included the five domains of the framework. The facilitator led the group discussion by focusing on one domain at a time and its relative importance in decision-making when applied to the digital research tailored use-case. The audio-recording was transcribed by the research assistant and augmented by notes. The transcripts were reviewed and coded by two researchers (CN, JT) using traditional qualitative methods [20].

Phase 2

During Phase 2, we incorporated feedback obtained from the expert panel and developed the second iteration of a decision-making checklist for use by researchers and tested it with a group of seven behavioral scientists.

The checklist was expanded to include a set of questions specific to each domain to prompt end user reflection. This initial checklist was refined during discussions among the coauthors and subsequently pre-tested by one of the team members using her own research as a use case scenario. Based on responses and feedback the checklist was further refined and put into an online survey format to add skip logic based on response options (see checklist

Fig 1 | The APA pyramid modified to include ethical principles as a foundation.

Table 1 | Phase 1 checklist domains, items, and implementation opportunities

Domain	Items	Implementation opportunities
Ethical Principles	Informed consent process	
Autonomy	How does the technology work? What personal data are produced? How is privacy managed? How is data stored and shared? Is bystander data captured?	The goal of autonomy is to facilitate good decision-making by the person being recruited for study participation. Dr. Crystobal will need to develop an informed consent process that clearly describes how she is addressing the items in column 2. The terminology used will need to be understandable to those being recruited. Think of the informed consent process as an opportunity to educate.
Beneficence	Favorable risk/benefit analysis? Summary of known risks? Potential of unknown risks? Sound risk management plan? Rigorous study design?	Beneficence is a process of assessing risks and benefits. Dr. Crystobal needs to provide adequate information to a prospective participant that allows them to weight the risks and benefits of participation. To convey the risks and benefits may require some creativity and use of images. For example, conveying the concept of digital dust may require that Dr. Crystobal use pictures that depict metadata and insert questions within the consent process to assess comprehension.
Justice	Sample selection is appropriate?	The principle of Justice is demonstrated by recruitment protocols that result in a fair distribution and the benefits and burdens of research and that are inclusive of participants most likely to benefit from knowledge gained. Dr. Crystobal can include an explanation of why she is recruiting Latina women and how people like those involved in the study may benefit from the knowledge gained.
Privacy and Confidentiality	Are data encrypted? Who has data access? Is data security sound? Does HIPAA apply? What data are shared? Is there a privacy agreement?	It is important to protect participant privacy and the confidentiality of data. When using a commercial app or sensor, Dr. Crystobal will need to review the terms of service and privacy policies to make sure the information collected from participants is not sold or shared without explicit consent of the person she enrolls.
Efficacy	Is there evidence that the product works? Has it been used with the study population?	Not all technologies produce useful data. Dr. Crystobal will need to review the literature to see if the tools she is considering for her study have been tested to ensure the data are accurate. In addition, she should identify whether the technology is culturally appropriate.
Usability	Is the technology customizable, sustainable and accessible? What is needed to use the technology?	Similar to efficacy, the technology needs to be accessible to those enrolled. Dr. Crystobal will need to speak with people who are like those she plans to recruit to identify any barriers to using the technologies. If not usable, she may not obtain the necessary data to answer her research questions.
Interoperability / Data Sharing	Is data available for post research access and sharing? Is there interoperability into electronic health record?	To ensure that data collected is meaningfully integrated into the patients care or progress part of their treatment plan, Dr. Crystobal will need access to information the technology has gathered. This will avoid fragmentation of care and data silos of digital information.

and domains in Table 1). Using feedback obtained during this iterative process, we developed the following descriptions for each of the five domains of: Ethical Principles, Privacy and Confidentiality Evidence, Usability and Interoperability.

Specific to *ethical principles*, we utilized the widely accepted concepts in research ethics including: **Autonomy**—the right for an individual to make his or her own choice once provided with information in a setting conducive to decision making; **Beneficence**—evaluating research risks against potential benefits; and **Justice**—emphasizing fairness and equality among individuals and translation to those who may benefit from the research results [19].

With respect to *privacy*, the focus is on both the participants' expectations about the control of information they hold to be personal and how or whether data they choose to share is managed. The privacy factor prompts one to be familiar with the terms and conditions of use as well as the terms of an associated privacy agreement, should it exist. *Confidentiality* concerns data management and whether data are collected with appropriate consent, securely stored and have clear standards for data sharing. When commercial products are used for research purposes, it is important to evaluate the location of data storage, data transfer protocols as well as ownership and ability to litigate if harm occur.

The evaluation of *evidence* is important as data fidelity is critical to answering research questions. Evidence can take the form of peer-reviewed research using the gold standard of a randomized controlled trial to consumer ratings on a website. Knowing what evidence is used to make an informed decision for app/device selection factors into the risk to benefit assessment. That is, if a device does not produce meaningful data, it is a waste of participant time and resources such that the risk exceeds possible benefit. Given that these tools may not yet be tested for reliability and validity, research may be conducted to generate the evidence.

With *usability*, the focus is shifted on to whether the research participant knows how to use the app/device (i.e., technology literacy), can actually use the device/app and, subsequently adhere to it for the study duration. Usability influences data quality and, if it is not usable for whatever reason, there may be increased missing data or a complete lack of data. This became evident to a researcher working with a community of refugee women who agreed to wear a wrist worn device for tracking physical activity. The researcher later learned, after finding no data on the devices, that a wrist worn device was not suitable for their culture and lifestyle [21]. Lastly, *interoperability* is the ability to access and share data with others. For example, can both the researcher and the participant access raw data and, if appropriate, can the digital data be shared with the research participant's electronic health record, if desired?

To make the function of this checklist more salient to a researcher, examples of how the domains and checklist items can be applied in practice are presented through the following brief use case.

Professor Crystobal is a researcher studying sedentary behavior among Latina women and wants to use a wearable sensor to detect participant's daily activity. Column 1 identifies the domain of interest, column 2 lists items that should be considered, and column 3 provides a few ideas on how this checklist can be used in practice.

Phase 2 recruitment

Once the checklist tool was pre-tested internally, a group of behavioral scientists ($n = 7$) affiliated with the Society of Behavioral Medicine's digital health working group were asked to complete the checklist.

Phase 2 procedures

This group of scientists was asked to reflect on a recent study they designed that used a digital technology. Using that study as their "use case" reference, they were prompted to reflect upon that study while responding to the checklist items. Using SurveyMonkey to deploy the checklist survey, participants provided a narrative description of their use case in response to "Describe the technology and how it was used in your research."

Researchers were asked to evaluate each of the five domains for importance to the decision-making framework and determine for each domain whether it was explicitly considered in their study; and, based on the question prompts in each domain, to identify if something was missing from that particular section of the checklist. Lastly, for each domain, we asked respondents to identify whether additional items were needed and, if so, to describe what would improve the checklist. Researchers were also asked whether they would approach their research differently in the future, having completed the checklist.

RESULTS

Phase 1 expert panel

This group identified their expertise and training as bioethics/law ($n = 1$), bioethics/philosophy ($n = 1$), research/clinical ethics ($n = 1$), bioethics/anthropology ($n = 1$), human research regulatory professionals ($n = 3$) with several participants also identifying as researchers. All were employed by organizations in the greater New York City area when the expert panel convened.

Discussion of each of the five domains resulted in group consensus that: 1—all five were necessary to guide decision-making; 2—there is uncertainty about whether the framework would be useful to all stakeholders in the digital health ecosystem; 3—clinicians should be involved as stakeholders only when their patient is also a research participant and the participant wants to share data with their clinician, and 4—the hierarchical structure weighting one factor as more important than another was not endorsed when applied to the research context.

As the model might be more or less useful depending on represented stakeholders, participants were asked to comment based on their role as a researcher, regulatory/ethics/policy expert or as a potential participant in digital health research.

Participants reflecting on the "researcher" role expressed an obligation to be informed about the technology and potential risks to participants, as noted in the following,

> Being a researcher and having an app and thinking about whether to recommend it to a participant − I would want to be responsible for knowing what this app's privacy policy is and what's happening to the data. ...I'd also feel responsible or compelled to explain to the participant that this policy is there, but privacy is contextual and in some ways the policy doesn't cover.

In the discussion of privacy and confidentiality, participants noted a need for increasing awareness among IRB members. For example, concerns about the potential for data to be subpoenaed or used to profile an individual. The volume and granularity of data captured via apps and devices was voiced as an elevated risk with which neither researchers nor regulatory bodies were sufficiently familiar.

One participant noted:

> I think it would be useful to point out to a potential participant, when it comes to privacy you're asking people to give information that was private and there might be some existential or psychological harm that comes from knowing you're being constantly tracked. All of a sudden, every movement you make is now knowable information and accessible in a way it never was before. Knowing it may help, but maybe there is some self-policing that comes from knowing you're being watched. This is a type of potential harm that people haven't really thought about.

With devices that may be commercially available and are planned for use in health-related research, one participant referenced the Federal Trade Commission's (FTC) role specific to privacy:

> The FTC would demand some sort of privacy agreement and see whether you are conforming to it and then there are state regulatory bodies as well.

While this may be true for studies that are testing a regulated app or device, it is worthy to note that not all technologies are regulated. For example, the health wellness technology sector is not regulated by the FDA.

Discussions about whether a device or app was suitable for research purposes led to questions about the quality of the product or what we called the "Evidence" domain. Participants expressed concern about the lack of trustworthy evidence available to support the use of a particular device or app.

One participant asked if there was an evidence-base that could inform researcher decisions:

> Is there anything that is evidence based that you can choose from... is there something already known that is effective for what you want to use this for? If there isn't anything, then the bar for evidence for that particular tool drops.

As well as concerns about the quality of evidence available from which to form a decision about whether it is the right tool for the project.

> With evidence, it's a little squishy in that we can say yes these are peer reviewed articles, but it's the evaluation of those peer reviewed articles...

Lastly, the hierarchical structure of the model was not deemed an optimal way to communicate the framework. Several times during the discussion of the various domains, participants would comment about the hierarchical nature. When discussing the need to evaluate "evidence" and the "usability" of a product, one commented, "Maybe it's not a pyramid ... not feeling the pyramid thing"; which was followed later by "I don't get the pyramid at all." The group majority voiced that all of the domains discussed were equally important. For example, if a device has not been assessed for usefulness by the researcher in advance of deploying that tool in a study, it will not work. Effectively, if a participant cannot use the tool, then there is no data to protect as they are not likely to use the device or use it correctly. Similarly, a data transfer protocol (presently noted under the "Interoperability" domain) may introduce data security threats, which can impact a participant's privacy and confidentiality of their data.

Specific to the foundation of "ethical principles," while presented as a discrete domain, it was evident that the principles of autonomy, beneficence, and justice cut across the other four factors. For example, if a privacy policy states that individual data may be sold or shared, it naturally increases the risk to a participant if strategies for protecting data are not adequate. Evaluating risks is routinely a process which takes place when determining beneficence—that is, weighing the potential study benefits against potential study risks. Likewise, data management is a process for mitigating data breach and potential compromises to the participant data confidentiality.

Results of our expert panel suggest an appropriate model may actually be closer to a flexible Venn diagram where all domains discussed above are necessary but, of variable importance depending on the unique context and proposed use.

Phase 2 results

Seven behavioral scientists affiliated with the Society of Behavioral Medicine completed the checklist with six completing all questions. The checklist contained 47 items and was dispatched in February 2018 via an online survey tool. Question 1 asked respondents to identify their specialty which included the following options, including behavioral scientist, computer scientist, information scientist, designer/usability expert, computational scientist, IRB/regulatory expert, privacy expert, and ethical/legal expert. While we had specifically targeted behavioral scientists, six responders selected behavioral scientist and one identified as a designer/usability expert. The results of questions specific to the checklist are reported below by domain.

Ethical Principles

Respondents were asked to reflect upon their specific digital health research use case, specific to ethical principles. The initial prompt focused on the informed consent process which represents the application of the principle of autonomy or respect for persons. Using the checklist, respondents anchored their answers via a 5-point Likert scale where 1 = strongly agree to 5 = strongly disagree, to document the extent to which they conveyed information about: 1—how the technology works; 2—the volume and granularity of personal data produced; 3—best practices for protecting participant privacy and storing data; and 4—guidelines if bystander information is obtained. After responding to each item, we asked "Is there something missing from this section of the checklist?" Four indicated "no" and two provided comments with one stating that more information was needed on the term "bystander" rights as noted here:

> not missing, but perhaps more info on the last item (bystander info?).

Another noted their actual practice of informing participants about data management as noted in this passage:

> We made it explicit that participants had full and complete access to their data at any time with written request. We assessed participants' physical activity data using an armband for a week, and many wanted full reports of their activity data. We believe participants should have access to their data, so all requests to produce and securely share their data with them were honored.

The next item focused on the principle of beneficence where responders were prompted to reflect upon "the risks associated with use of the technology" using the following prompts: 1—are known; 2—are potentially unknown; and 3—are well managed; potential benefits outweigh study risks; and study design is rigorous and will result in developing new knowledge. We then asked if there was something missing from this section of the checklist and received one "yes" and three "no" responses with two comments. One comment focused on distinctions of risk with the separate "privacy" section as noted here:

> It was unclear to me how these risks also related to the privacy section below, as the main risks associated with using my technology are loss of privacy/confidentiality.

The other *response*, "items on the benefits (as well as risks)" was inferred as perhaps needing additional focus on benefits within the checklist.

Lastly, we asked about the principle of justice where responders were asked whether the people invited to participate reflected those most likely to benefit from the knowledge gained from the study. When asked if something was missing from the checklist, one reported "no" and two "yes" with three comments. One comment focused on the need to share results with the participants noting that this return of information is not typically considered.

> I wonder if you can reflect something here about the need to SHARE RESULTS with those who are most likely to benefit from study outcomes – this seems like a critical part of the research process that is often neglected.

Two comments focused on the potential for unintended bias related to the time commitment needed for some digital health studies as well as the issue of accessibility to participation.

> It's possible that time commitment will deferentially affect some groups of individuals, within the group of those most likely to benefit. However, I'm not sure this belongs here or whether it could be addressed.
>
> Maybe whether people had equal access to participate?

Privacy and Confidentiality

Within the Privacy and Confidentiality domain, the first prompt asked about the type of data collected in their reference use case using an open-ended format. Results focused on the digital aspects primarily; however, it may not have been clear that digital data were the focus of the prompts. Examples of responses follow:

> self-monitoring of steps and dietary intake; messages sent between parents and children re: physical activity and dietary intake; goals and rewards set for PA and diet; views, clicks, log-ins on website; downloads of study newsletter.
>
> survey, ema [ecological momentary assessment], activity, gps, weather.

> Not sure if this question relates just to the digital component? The Fitbit® collected data on activity (multiple dimensions), heart rate, and sleep.

Responders were asked to reflect upon participant privacy and data confidentiality practices using the following prompts: 1—Are data encrypted to enhance security? 2—Can data be accessed by the participant? 3—Are the data stored in a secure cloud? and 4—Is storage HIPAA compliant? Following this set of questions, we asked who was authorized to access the data and what data are shared and with whom using an open-ended response. We then asked about the use of commercial apps/devices, including whether they had reviewed the privacy policy if, in fact, they were using a commercial product. Specific to the use of commercial products, prompts were included to gauge potential increased risk to participants including selling of participant-generated data to a third party. The section concluded by asking the respondents to identify whether something was missing from this section of the checklist whereby three indicated "no" and two provided comments.

One respondent focused on the distinguishing proprietary from commercial products and the ability to control participant data as noted here:

> this should be separated out for proprietary vs. commercial products used; I know much more about the privacy features of my own website than those of all the commercial products recommended for download in my study.

Another stated that adding an item that focused on the practice of keeping personally identifiable information separate from the data were an important factor, although challenging.

> Maybe a question on whether/how personal info is kept separate from data (user IDs); this is challenging when incentives are delivered via email (e.g., gift card codes).

Evidence

Moving to the domain of "Evidence" we asked responders to reflect on what was known about the technology using an open-ended response option and whether there was evidence to support claims using a forced choice option of "Yes," "No," and "Other." If the respondent selected "Yes," they were asked whether the evidence was peer-reviewed. When asked if something was missing from the checklist, three indicated "*no*" and three commented in lieu of selecting "Yes" or "No." Comments focused on the lack of clarity in the prompts and the need for more guidance and direction in the instruction.

> Other than some clarification of what is meant by efficacy; and evidence. I could imagine it's good to know if there's evidence that the technology is well-conceptualized, even though it's not been tested yet for key outcomes.

One person indicated "This section was difficult to fill out!" and another stated, "I'm not following this anymore."—both being useful comments to apply to the next checklist iteration.

Usability

The domain of "Usability" included the following prompts in response to the question: For the target population: 1—Are the instructions, support, design, and engagement appropriate? 2—Can it be tailored to the end user? 3—Will the target population find long-term use feasible? 4—Is it accessible to diverse populations? 5—Need for internet?, and 6—Need for a smart phone? For these items, a 5-point Likert-scale was used ranging from "Not at All" to "A Great Deal." The final item asked if something was missing from this checklist.

Three responders stated "*no*" that nothing was missing and three added comments, but did not indicate "Yes" as a response. The comments were somewhat diverse with one indicating that a special type of smart phone would be useful and another noting that the concept of tailoring was possibly too large and needed to be more specific as noted below:

> consider adding 'need for particular type of smartphone'? seems to be an issue especially with many research-created apps, etc., where they are only supported by iOS or Android

and

> tailoring is very broad; maybe there are subcategories such as time of use, display would help? I also wonder if it would help to have a checklist option designed specifically for technology that is being developed/improved (Yes, No, In Development) to encourage use of the checklist at all phases of development.

An unanticipated response included the potential for a researcher to feel uncomfortable with past decisions prompted by this reflection process. The following comment captured what a potential lack of preparation can foster in digital health research and supports the notion of using a checklist prospectively during the early technology development and research design phase:

> With these questions, I'm just feeling guilty for doing work; I'm feeling rather helpless in my ability to do good work.

Given a goal of this exercise was to prompt reflection of previous research planning, it should not come as a surprise that scientists participating in this project were concerned that they had not considered all aspects of the checklist prospectively. Moving forward, the checklist will be used during the research planning process to assist researchers in thinking through how the five domains pertain to their specific study plan.

Interoperability

The final domain of "Interoperability" included six prompts to reflect on data sharing including: 1—Will data be shared with others so they can build upon this research? 2—Are data transferrable to an electronic health record? 3—Are data shared with anyone else outside of clinical care or research? 4—Are there limitations to access of data? 5—Can data be accessed by the participant? and 6—Can data be accessed by the investigator? A 4-point Likert scale was used to record responses that ranged from "Not at All" to "To a Great Extent." The final questions asked whether something was missing from the checklist and, as before, three stated "*no*" and three provided comments. One comment focused on an internal practice for data storage and two comments were more tailored to adding items to the checklist that address the connection between different technologies as evidenced by this comment:

> consider adding something about the connection between different technologies within a specific study – for example, participants in my study accessed their own data via the study website but also generated their own data on commercial apps and these didn't "talk" to my website; other interventions are delivered via social media or other tech platforms and ask participants to track on MyFitnessPal, Fitbit, etc. but all tech isn't nec [essarily] connected.

Future research planning

Using an open-ended response option, the final survey questions asked participants if they would approach research planning differently in the future having now completed the checklist. Five of the six responding indicated that the checklist was helpful to think through the potential issues and that they were unaware of the possible risks that they may have introduced by not thinking about interoperability, privacy, and consent as noted here:

> There were various factors around interoperability and privacy that I did not previously consider, especially with the use of commercial apps within my study. Data from these apps weren't collected for my study, but I did request that participants try a variety of apps, which could have put them at risk for loss of confidentiality that wasn't as clear at the time of the study. Will consider this more in the future and also include transparent language in the consent forms. It was very helpful to think through all these questions/issues and apply them to my project. And, yes, I'm sure I would do things differently next time.

> The issues of interoperability and data sharing should be considered. IVR Systems have many systems in place that ensure the safety, security, privacy and usability of the technology, but the technology will advance as there are ways to analyze the sounds of the participants voice which could lead to identification. All tech can be breached so more considerations are necessary today than yesterday.

One participant was fairly confident they had considered prospectively what needed to be addressed yet, acknowledged that this had come from some trial and error experiences.

> I do not believe so. I think it was critical that when using a commercial device, we included the creator as a co-investigator. This helped us think through generalizability and future applicability issues once the software and hardware are updated. I have done other studies in the past where we did not do this, and it was a mess. I highly recommend doing this whenever possible.

Lastly, one commented that this reflection process was fostering an inability to feel confident that they should be conducting this research.

> … I also feel powerless as there is so little I feel I understand in making decisions on use. Truly, I don't know how to handle this at this point as this is making me think that we shouldn't be doing any research anymore; I don't know how to do the proper due diligence needed on the unknowns.

Refined framework and checklist

Our results led to a final step of revising the checklist for future use by digital health researchers. Terminology used in checklist and confusion associated with domain characteristics led to several changes. First, we uncoupled the domain labeled "Privacy and Confidentiality" leaving "Privacy," which pertains to the individual, as its own domain and moving "Confidentiality," which focuses on the data, to the "Interoperability and Data Sharing" domain. We then expanded the latter domain to broadly include: data collection, storage, sharing and interoperability and renamed the "Interoperability and Data Sharing" domain to "Data Management." There was some confusion specific to assessing risks and benefits using the original domains of "Ethical Principles" and subcategory of "Beneficence" with the "Evidence" domain. Because establishing the existence or lack of "Evidence" intersects with a traditional risk to benefit assessment, we renamed the evidence domain to "Risks and Benefits" and folded prompts previously used under "Evidence" within this more comprehensive domain. The domain of "Usability" was expanded to also include "Access." Finally, we observed that while "Ethical Principles" can stand as a separate domain, in practice, the ethical principles cut across all domains. For example, a lack of evidence may elevate the potential risks of harm when compared to potential benefits or, "Beneficence." Likewise, if an app is not available across operating systems, it may limit access to participation, which presents a "Justice" issue. To demonstrate how the Ethical Principles cut across the domains, a column was added so that researchers could connect prompts within each domain to the research plan development and the informed consent process.

As a result of these changes, the final "Digital Research Decision-making Framework" contains five domains. A companion checklist is located on the Connected and Open Research Ethics platform (thecore.ucsd.edu) in the Resource Library under the "Tools" tab. The domains are now labeled: 1—"Privacy," 2—"Risks and Benefits," 3—"Access and Usability," 4—"Data Management," and 5—Ethical Principles and are presented as a Venn Diagram with Ethics at the center (see Fig. 2).

To use the checklist, the researcher is prompted to evaluate the technology, the technology product information, privacy policies and the literature in the cases where the technology has been previously used in research. The checklist response options are *yes*, *no*, and *unsure*. In the cases where the researcher checks no or unsure in evaluating the technology, the researcher must consider whether the technology poses risks to participants and the protection of their rights or threats to the study design that would preclude use of the technology for the intended research purposes. The researcher must also consider if an alternative technology should be selected or if there are different approaches that can be taken in the design and conduct of the study that would overcome any potential risks or barriers. For example, if the selected technology requires internet access, the researcher will need to determine whether internet access is possible and, if not, whether access will be provided as part of the study design.

DISCUSSION

While the need for a framework and a checklist to help inform decision-making in the digital health ecosystem is a growing imperative, the solution is complex. The checklist domains evolved as our process of engaging experts revealed important points for further consideration. However, the linear ordering and hierarchy of the model was deemed suboptimal. Instead, the discussion identified the complex interactions of ethics, privacy, safety, efficacy, usability, and data management that, depending on the situation and stakeholder, assumed different priorities and requirements. One expert panel member suggested a goal of the decision-making framework is to prompt reflection at the point of selecting the digital technology and, is rarely about the tool as an object of study. In the case of using artificial intelligence with a camera and sensors, ideally the researcher would pause and consider whether the participant is adequately informed, that the camera and sensors were vetted for appropriate privacy protections and data management, etc. This comment reveals a potential need to socialize the researchers about their role in the tech-ethics ecosystem as being key actors in advancing responsible research.

Results of this group discussion offer insight into what a more optimal tool may look like, which were further assessed by a group of behavioral scientists. Considering the dynamic and rapidly evolving nature of technology use in health research, a decision-making tool must also be flexible. While the key domains are useful to consider, the checklist questions/prompts and other information is necessary to help users evaluate each of these domains specific to a greater context. Moreover, the framework and checklist must evolve with the technologies themselves. We also recognize that it is important for the framework to inform discussions about technology between different stakeholder groups. For example, a researcher talking to a participant about the use of an app needs to know the participant's level of technology literacy to fully assess usability, which can influence protocol adherence. Likewise, a potential participant needs to understand the possible risks of a data breach in order to fully understand the potential threats to their privacy and related data security merits of an app. The IRB must understand the researcher's evidence for efficacy of a category of apps before deciding if the proposal under review will lead to scientific advancement through the

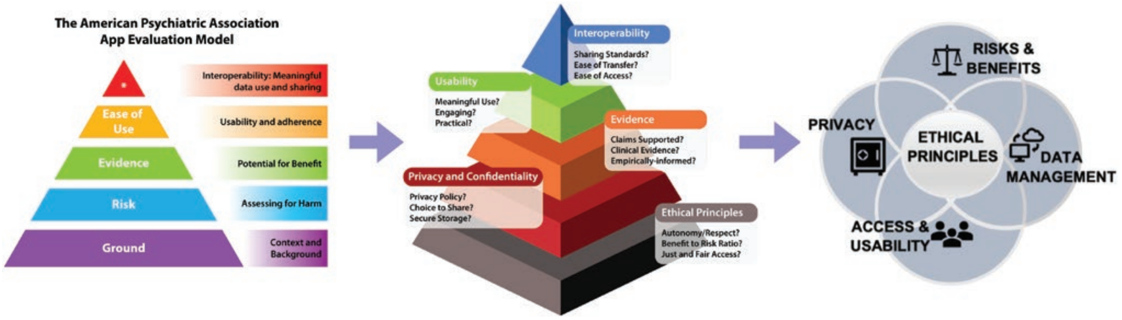


Fig 2 | Decision-making framework progression from hierarchical to interconnected domains.

production of generalizable knowledge. Creating this type of tool is possible yet, presents a unique challenge in untangling the inherent complexity of the possible interactions.

There is a surprising gap in the scholarly literature on the knowledge base, readiness, and current use of health technology for research among all stakeholders and this presents an immediate opportunity for impactful survey research. The discussion on stakeholder adoption identified an urgent need for further education among all stakeholders regarding technology use in health research. Even offering simple use case studies suggested a potentially high-impact intervention. Understanding how to better evaluate the technical aspects of digital security was raised as another area for educational interventions. Consequently, the discussion of ethical principles suggests that the role of ethics in this emerging health technology research space cannot be isolated by a single factor but, may be better represented as an integral element of all factors. In sum, translating current thinking on research ethics into updated and flexible strategies that inform how key factors interact and map to principles of autonomy, beneficence and justice is yet another high impact research area for researchers to investigate.

Since initiating this work to develop a digital research decision-making framework and checklist, we reviewed the publication of other guidelines, including a recent paper to describe key points to consider [22] and the Xcertia initiative advanced with an initial draft. That being known, none of these efforts are tailored to the needs of the research community—specifically those using digital tools for non-clinical purposes. As we demonstrated in our research, the need for clinician–patient developed frameworks are likely not transferable, as the clinician–patient relationship is unique and distinct from the researcher–participant. Regardless, it is encouraging to see new efforts emerging in this important space.

Limitations

To our knowledge, this project is the first study to query researchers, regulatory experts and research/bioethicists to discuss strategies for fostering informed decisions about the selection of apps/devices as tools used in health research. As a result of this work, the framework and checklist tool provides insights from which to build upon for further research. The nature of this developmental work is to gather insight, rather than draw generalized conclusions. This checklist will require further input and the ongoing involvement of researchers, ethicists, legal scholars and regulatory experts. Additional research is needed to obtain perspectives of technology makers and consumers (e.g., research participants) to fully validate the checklist. As this checklist is used by a wide and diverse audience, it will undoubtedly require further updating and refinement.

CONCLUSION

Creating an inclusive and informative tool remains a complex challenge. Translation of clinician–patient tools or simple linear-hierarchical tools are likely insufficient for the multi-stakeholder perspectives of participants, researchers, IRBs and technology makers. We recommend that coordinated efforts to develop best practices and standards that support the responsible selection of technologies used in research remain a priority. Bearing in mind the potential for digital research to impact health and health care, it is critical for those involved to build and choose tools that will provide meaningful research that translates well to clinical practice while protecting research participants in the process.

This study was supported, in part, by the Robert Wood Johnson Foundation (#72876, Nebeker, PI, 2015-2019) and the National Institutes of Mental Health (1K23MH116130-01, Torous, PI, 2018-2022).

COMPLIANCE WITH ETHICAL STANDARDS

Authors' Contributions: C.N. conceptualized the study, contributed to data collection, analysis and interpretation, and wrote/revised the manuscript for intellectual content. R.B.E. contributed to the Phase 2 study design, data collection, analysis and interpretation and manuscript revisions for intellectual content. J.T. contributed to the Phase 1 study design, data analysis and interpretation and manuscript revisions for intellectual content.

Conflict of Interest: The authors declare no conflict of interest.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The UC San Diego IRB verified the work carried out in this study as exempt.

Informed Consent: Informed consent was obtained from all individual participants included in the study via a verbal process.

References

1. Nebeker C, Lagare T, Takemoto M, Crist K, Kerr J. Exploring mobile imaging, sensing and tracking ethics (MIST-E) in behavioral science research: participant perspectives from the iWatch study. *Transl Behav Med.* 2016;6(4):57–7586.
2. Nebeker C, Harlow J, Espinoza Giacinto R, Orozco-Linares R, Bloss CS, Weibel N. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. *AJOB Empir Bioeth.* 2017;8(4):266–276.
3. Dunseath S, Weibel N, Bloss CS, Nebeker C. *NIH Support of Mobile, Imaging, Pervasive Sensing, Social Media and Location Tracking (MISST) Research: Laying the Foundation to Examine Research Ethics in the Digital Age.* 2018;1, Art. no: 20171.
4. Bloss C, Nebeker C, Bietz M, et al. Reimagining human research protections for 21st century science. *J Med Internet Res.* 2016;18(12):e329.
5. Gandhi M, Wang T. Digital health consumer adoption: 2015. Available at https://rockhealth.com/reports/digital-health-consumer-adoption-2015/. Accessed 30 May 2018.
6. Beilinson, J. Glow pregnancy app exposed women to privacy threats, consumer reports finds. 2016. Available at https://www.consumerreports.org/mobile-security-software/glow-pregnancy-app-exposed-women-to-privacy-threats/. Accessed 30 May 2018.
7. Kramer DB, Fu K. Cybersecurity concerns and medical devices: lessons from a pacemaker advisory. *J Am Med Assoc.* 2017;318(21):2077–2078.
8. Glenn T, Monteith S. Privacy in the digital world: medical and health data outside of HIPAA protections. *Curr Psychiatry Rep.* 2014;16(11):494.
9. Rosenfeld L, Torous J, Vahia IV. Data security and privacy in apps for dementia: an analysis of existing privacy policies. *Am J Geriatr Psychiatry.* 2017;25(8):873–877.

10. Blenner SR, Köllmer M, Rouse AJ, Daneshvar N, Williams C, Andrews LB. Privacy policies of android diabetes apps and sharing of health information. *J Am Med Assoc.* 2016;315(10):1051–1052.
11. US Department of Health Human Services. *Examining Oversight of the Privacy & Security of Health Data Collected by Entities not Regulated by HIPAA.* Washington, DC: US Dept of HHS; 2016.
12. Huckvale K, Prieto JT, Tilney M, Benghozi PJ, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med.* 2015;13:214.
13. Federal Trade Commission. App developers: Start with security. 2017. Available at https://www.ftc.gov/tips-advice/business-center/guidance/app-developers-start-security. Accessed 30 May 2018.
14. U.S. Food & Drug Administration. Digital health software precertification (Pre-Cert) program. 2018. Available at https://www.fda.gov/medicaldevices/digitalhealth/digitalhealthprecertprogram/default.htm. Accessed 30 May 2018.
15. Torous J, Chan S, Wang P, Kim J. Introducing a hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatr Serv.* 2018;69(5):498–500. doi:10.1176/appi.ps.201700423. Epub 2018 Feb 15..
16. Xcertia(TM). *Xcertia(TM) mHealth App Guidelines.* 2017. Retrieved from http://xcertia.org. Accessed 30 May 2018.
17. Torous J, Nebeker C. Navigating ethics in the digital age: introducing connected and open research ethics (CORE), a tool for researchers and institutional review boards. *J Med Internet Res.* 2017;19(2):e38.
18. App Evaluation Model. 2017. Available at https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model. Accessed 30 May 2018.
19. U.S. Department of Health and Human Services, 45 CFR 46. Subpart A––basic HHS policy for protection of human research subjects. *Fed Reg.* 1991; 56:28012–28022.
20. Saldaña, J. 2015. *The Coding Manual for Qualitative Researchers.* Thousand Oaks, CA: Sage.
21. Nebeker C, Murray K, Holub C, Haughton J, Arredondo EM. Acceptance of mobile health in communities underrepresented in biomedical research: barriers and ethical considerations for scientists. *JMIR Mhealth Uhealth.* 2017;5(6):e87.
22. Jake-Schoffman DE, Silfee VJ, Waring ME, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. *JMIR Mhealth Uhealth.* 2017;5(12):e190.

Established in 1871

Swiss Medical Weekly

Formerly: Schweizerische Medizinische Wochenschrift

An open access, online journal • www.smw.ch

Review article: Biomedical intellegence | Published 16 January 2018 | doi:10.4414/smw.2018.14571
Cite this as: Swiss Med Wkly. 2018;148:w14571

Digital health: meeting the ethical and policy challenges

Vayena Effy[a], Haeusermann Tobias[b], Adjekum Afua[a], Blasimme Alessandro[a]

[a] Health Ethics and Policy Lab, Department of Health Sciences and Technology (D-HEST), ETH Zurich, Switzerland
[b] Department of Sociology, University of Cambridge, UK

Summary

Digital health encompasses a wide range of novel digital technologies related to health and medicine. Such technologies rely on recent advances in the collection and analysis of ever increasing amounts of data from both patients and healthy citizens. Along with new opportunities, however, come new ethical and policy challenges. These range from the need to adapt current evidence-based standards, to issues of privacy, oversight, accountability and public trust as well as national and international data governance and management. This review illustrates key issues and challenges facing the rapidly unfolding digital health paradigm and reflects on the impact of big data in medical research and clinical practice both internationally and in Switzerland. It concludes by emphasising five conditions that will be crucial to fulfil in order to foster innovation and fair benefit sharing in digital health.

Keywords: *digital health, personalised health, digital ethics, data governance*

Introduction

Digital health is a rapidly expanding medical field premised on the availability of ever increasing amounts of data about people's lifestyles, habits, clinical histories and pathophysiological characteristics. According to the US Food and Drugs Administration (FDA) "[t]he broad scope of digital health includes categories such as mobile health (mHealth), health information technology (IT), wearable devices, telehealth and telemedicine, and personalized medicine" [1]. These categories rely heavily on human health data. Conventionally, the collection of health data is mediated by officially licensed medical devices, such as diagnostic instruments or genome sequencers, operated by health professionals in clinical environments and under strict regulatory conditions. Moreover, clinical data are typically stored in public health registries, at hospitals or in the archives of individual physicians. Digital health, in turn, entails connecting health-related data, including data generated by patients themselves, and harnessing the medical potential of technological tools of common usage, such as smartphones, wellness bands, apps, social media and sensing devices disseminated in our dwelling environment. Most of these tools are not initially conceived for medical use and are not marketed as medical devices. Notably, however, some prominent digital health technologies already cut across the rigid distinction between licensed and ordinary gadgets, and the latter have also started to receive official designation as medical devices (see table 1) [16]. But digital health is not limited to ordinary technology, nor to ordinary-turned-medical technologies. Certain digital health tools present entirely novel features, as in the case of digital pills that, thanks to a microcircuit activated upon contact with liquids in the patient's stomach, can tell an external sensor whether and when a patient has taken his or her medication.

The defining feature of digital health, however, has to do with data rather than technology. What is distinctive about digital health in this respect, is that – typically through wearable, portable, ingestible or otherwise implantable devices – it generates a "seamless flow of critical medical data between patients, their families and their physicians" [17]. The ambition of digital health is therefore aptly described as generating a *circulation of data* from patients (patient-generated data), to devices and/or health professionals (who analyse and make sense of the data), and then back to devices that eventually provide the patient with information regarding their health status and how to manage it.

To this aim, phenotypic and behavioural information, as well as data about socioeconomic status and dwelling environment, need to be collected. Information posted on social media can also turn out to be potentially relevant to both individual and population health [18, 19]. Digital health thus inhabits what has been recently labelled an "evolving health data ecosystem" [20], a space that also includes data gathered by healthcare services, such as electronic health records, genetic or genomic data, diagnostic data, claims data and the like. According to some, given their volume, complexity, variety and propensity to be analysed through data-mining techniques, such data qualify as big data [21] or, more precisely, as biomedical big data [22–24]. This expanded set of health-relevant data is expected to occasion huge progress in medicine, for example by helping people monitor their health status, assisting patients in coping with their conditions, inferring health-related issues earlier on, personalising treatment to individual patients'

Correspondence:
Alessandro Blasimme, PhD, ETH Zurich, CH-8092 Zurich, Switzerland, Alessandro.blasimme[a]hest.ethz.ch

characteristics, improving outcomes, reducing costs and inefficiencies, and also boosting medical discovery and accelerating drug development. Admittedly, there are significant expectations of digital health and there is strong interest on the part of numerous stakeholders in promoting it and seeing it flourish. At the same time, for digital health to materialise several ethical and policy challenges need to be overcome [25].

To review these challenges, a multidisciplinary symposium was held at the University of Zurich (UZH) on 1 December 2016. The symposium, convened by UZH's Health Ethics and Policy Lab (now based at ETH Zurich), brought together different perspectives from national and international experts regarding the challenges that accompany the development of digital health. Participants included scientists, ethicists and lawyers, representative of national research institutions such as the SAMS (Swiss Academy of Medical Sciences) and the SNSF (Swiss National Science Foundation), as well as policy specialists from international organizations such as the OECD (Organization for Economic Co-operation and Development) and the WHO (World Health Organization).

Three key challenges impinging on the development of digital health were identified and discussed:

1. How does digital health fare with respect to the demands of evidence-based medicine?
2. How can public trust in digital health be generated and sustained
3. What policy gaps can and should be addressed through global policy instruments and what instead require specific initiative in the Swiss context?

Here, we provide key considerations on the above three questions, based both on the discussions held at the symposium and further literature review. These considerations are of relevance to scientists, ethicists and public health experts, as well as developers and policy makers interested in assessing the impact of big data in medical research and clinical practice, both internationally and in Switzerland.

Digital health and the quest for evidence

The clinical development of digital health applications is premised on the creation of very large data collections recording sensitive personal data. In the public sector, examples include: the 100K genomes cohort in the UK, which aims to sequence the genome of one hundred thousand NHS cancer patients by 2017; the All of Us cohort of the Precision Medicine Initiative in the US, which will collect samples, and phenotypic and clinical data from one million Americans; or the Million Veteran Program, which currently constitutes the largest genomic database in the world and also includes lifestyle information and access to electronic health records for research purposes [26]. Besides these large-scale public initiatives, the private sector is also collecting huge amounts of phenotypic and genetic data from users of health-related services and products. For example, as of June 2015, the genetic testing compa-

Table 1: Examples of licensed and unlicensed digital health technologies.

Company	Product	Type	Year of license	Licensing body	Description
Licensed					
AdhereTech, Inc.	Smart Pill Bottle [2]	Wireless pill bottle		FDA, CE, ISO	Smart wireless pill bottle capable of alerting patients to missed doses.
Airstrip Technologies	AirStrip ONE® [3, 4]	Mobile app	2014	FDA	Provides an interoperable platform that simplifies clinicians' and patients' access to diverse health data.
AliveCor, Inc.	Kardia Mobile [5, 6]	Wireless pad	2016	FDA	Portable electrocardiogram device displaying results on a smartphone. It can detect cardiac anomalies such as arrhythmia.
Blue Spark Technologies, Inc.	TempTraq® [7, 8]	Disposable patch	Blue Spark Technologies, Inc.	TempTraq® [7, 8]	Disposable patch
Natural Cycles	Natural Cycles [9]	Wireless thermometer	2017	CE, ISO	App to keep track of ovulation and period, associated with a smart thermometer to determine fertile days. It can be used as a contraceptive.
Proteus Digital Health	Proteus Discover [10, 11]	Ingestible sensor, wearable sensor and mobile app	2014	FDA	An ingestible sensor mounted on a pill, which, when swallowed, sends a signal to devices keeping track of compliance with prescriptions.
Not licensed					
Butterfly Network, Inc [12].		Compact ultrasound			Portable ultrasound machines trained through deep learning algorithms.
Fitbit, Inc.	Fitbit Aria™ [13]	Wi-Fi smart scale			Used in conjunction with an app, it tracks body mass index, weight, body fat percentage and lean mass.
Happify, Inc.	Happify™ [14]	Mental health app			Smartphone app aimed at alleviating stress and negative thoughts through techniques in the form of games and exercises.
MyFitnessPal, Inc.	MyFitnessPal [15]	Calorie counter app			Free app to keep track of calorie intake. It is said to help users who want to lose weight.

ny 23&Me had collected and genotyped DNA from more than one million costumers [27]. In June 2016, the US-based healthcare provider and insurer Kaiser Permanente announced the constitution of a research biobank pulling electronic health records, DNA and behavioural and environmental information from 500 000 people [28]. Finally, end-users of digital health devices such as heart monitoring apps or fitness gadgets also contribute vast amounts of data to service providers. Such data can be cross-linked to other existing large-scale repositories both for research purposes and for developing new digital health services to users and professionals alike.

The evidence base for digital health

Mining large-scale data repositories creates challenges regarding data management, privacy protection and oversight mechanisms. Other challenges, however, relate more directly to the composition of such repositories and to the tools employed to mine the data they contain. For instance, the use of convenience samples to populate precision medicine and precision public health cohorts can bias the sample compositions and compromise the representativeness of target populations [29, 30]. Such issues can affect the quality of the evidence derived from digital health research and employed in digital health-based interventions, both at the individual and at the population level. Taking into account ethnicity, age, sex, socioeconomic status and geographical distribution in recruiting research participants thus seems crucial to ensure the generalisability of research findings. Similarly, the representativeness of the datasets employed for product development and the robustness of analytic tools to mine such datasets can affect the development of effective digital health services and devices by private companies.

There seems to be room for precompetitive research in this area in order to at least create standards and possibly reference datasets to enhance reproducibility. Meanwhile, progress in regulatory science should enable better assessments of evidence for safety, efficacy and cost-effectiveness. In both cases, policy stimulus appears crucial to achieve tangible results. As for more user-oriented digital health applications, as with products and services developed outside the realm of licensed devices, there is the need to enhance transparency and accountability by adopting forms of sector-specific self-regulation and adhering to robust corporate responsibility schemes.

Data variety is also a key issue in digital health. For example, although genetics can be extremely informative from a medical point of view, with a few notable exceptions the contribution of genetic variation to most common chronic conditions is either unknown or relatively small. Instead, other types of information, such as levels of physical activity, diet and socioeconomic factors, are better suited for predicting the risk of developing a chronic disease [31]. Therefore, to harness the full potential of data mining and predictive analytics in digital health, genomic data alone are insufficient [32, 33].

Novel modes of evidence generation could take into account multidimensional and unstructured data along with conventional clinical measures. For example, in health outcomes research or assessment of long-term effects of drugs and interventions, pragmatic trial designs are raising considerable interest. Such studies employ less restrictive inclusion criteria than traditional clinical trials and allow for concomitant morbidities and medications. Such models rely on “real-world data” collected from actual patients [34] – data that would simply not be available in randomised controlled trials. Real world data include medical records, data from portable devices and social media, as well as environmental and socioeconomic data. Other than saving on the high costs of randomised controlled trials, pragmatic trials based on reals world data promise to be more representative of real populations. At least when risks are deemed reasonably low, real-world evidence obtained through pragmatic designs could thus be used in support of regulatory decisions about the safety and efficacy of digital health devices and applications. Moreover, real-world evidence could also be employed to retrospectively assess digital health applications that reached the market without being cleared by regulatory agencies.

The technologies that are enabling extensive data collection and the development of digital health can be applied to both individual and population health issues, contributing to the emerging fields of precision medicine and precision public health, respectively [35–38]. Both the former and the latter promise more tailored interventions in their respective domains, progress in the understanding of disease causes and outcomes, along with reduced costs and improved access to effective healthcare. Both precision medicine and precision public health have specific sets of ethical implications [39, 40]. In such areas, larger, more representative and diverse databases are expected to tackle very well-known issues of external validity that afflict randomized controlled trials [41, 42]. Yet this prospect is affected by the challenges discussed above. Moreover, the use of artificial intelligence (AI) and deep learning [43] to mine such large data repositories has led many to think that digital health can dispense with mechanistic explanations and hypothesis-driven research, replacing them with mere algorithm-guided searches for correlations between phenomena in large-scale observational studies [44–47]. It has been noted, however, that even if those methods prove effective in establishing robust correlations, controlled interventional, randomised trials on stratified patient cohorts will still be necessary to establish the safety and clinical utility of novel therapies or public health interventions [48].

Ethical and policy challenges in digital health

Privacy and security

Most of the debate about big data uses for health purposes has focused on privacy. As more data sources become available and advanced analytics can be applied for various purposes, protecting privacy is undoubtedly a complex challenge. What contributes to this complexity is that standard mechanisms of protection such as anonymisation, notice and consent are excessively stretched in this environment of new capabilities. Consent for data uses can hardly include the exhaustive list of all possible future data uses [49]. In turn, anonymisation technologies, even if robust, still leave re-identification in the realm of possibility if enough resources were to be devoted to it. Data security has also been a challenge, with cyber attacks, hacking of databases and data kidnapping being reported frequently. Incidents of data breaches and “kidnapping” (data held by hackers for ransom) are on the rise. According to the Breach Portal of the Health and Human Services (HHS) Office of Civil Rights, millions of healthcare records have been affected to date. In May 2017, healthcare databases in one hundred countries faced a ransomware attack claiming a ransom of $300 in bitcoin to unlock affected machines [50]. The UK’s Information Commissioner’s Office notes that the health sector accounts for most of the data incidents reported to them. These incidents, along with growing public concerns about big data affecting most aspects of contemporary life, have contributed to a bleak picture of the future of privacy [51]. Understandably, such a picture does not create an environment conducive to the demands of digital health, namely easier data circulation between individuals, devices and institutions. Against this background, the public needs to be reassured that robust security measures are mandated and enforced through clearly articulated policies. Concerns can be addressed with the adoption of appropriate technologies, monitoring and evaluation of security systems, transparency and accountability mechanisms such as legal remedies and compensation for privacy harms resulting from security breaches. Security will continue to evolve, but the big data approach will continue to demand more technical skills, responsive policies and regulatory oversight.

Trust

Essentially what is at stake is the creation of a culture of trust that will enable all stakeholders in the big data ecosystem to benefit from the development of digital health [52]. In particular, public trust in health data uses is of paramount importance. The recent case of the care.data in the UK serves as a good example of how mistrust on the part of the public can derail large-scale data initiatives (see table 2). But trustworthy digital health activities require more than privacy protection. Elements of trust include transparency, accountability, benefit sharing and certainly more clarity about data ownership and data control. What is important here is the realisation that trust cannot only be built through achieving just one element, but rather through a concerted effort to promote all of its elements. Therefore, trustworthiness cannot merely be achieved by innovative consent models offering more or less control of data uses. Rather, consent innovation has to also be accompanied by clarity on how individuals and communities will benefit from digital health developments, by oversight mechanisms that protect common interests and by accountability mechanisms that can sustain public scrutiny.

Accountability

With automated data mining for decisions of clinical or public health relevance becoming one of the most promising features of digital health, accountability is of critical importance. In particular, the adoption of these new tools requires relevant adaptations in existing accountability standards. For instance, in the field of digital epidemiology, data mining can be used to analyse free, unstructured text from social networks in order to make predictions about the spread of infectious diseases [34]. Moreover, mobile technologies can be used to target specific populations with health-related information that can help contain the spread of infectious diseases. These new approaches can increase the speed and accuracy of health dynamics monitoring, leading to more targeted and effective interventions. However, premature reliance on such innovative tools could lead to an inappropriate use of public resources, unnecessary public alarm and individual harm from dispensable medications [55]. Similarly, it is anticipated that medical practice will increasingly be aided by AI algorithms for diagnosis, treatment decisions and surgical procedures [55, 56]. Progress in such areas is expected to greatly improve the quality of healthcare provision for individual patients. Such tools can range from simply providing assistance to practitioners, to possibly one day being fully autonomous from human supervision [57]. Indeed, increasing sophistication could lead to more accuracy. However, as more AI-guided tools become autonomous, fewer human operators are able to override their decisions. Hence, AI-guided medical devices have the potential to jeopardise current norms of professional accountability in clinical practice, making it more complicated to trace responsibility back to individual practitioners. It is therefore crucial that *ad hoc*, robust evidence standards are elaborated to guide the adoption of digital health technologies in clinical practice [58, 59].

Governance approaches in the development of digital health

Global perspective

The strong technological component of digital health does not imply that innovation in this area will affect only the most affluent countries. Recent figures published by the Global Observatory on eHealth of the WHO show that health systems in most countries increasingly rely on data [60]. In fact, the decreasing cost of digital technologies is making it possible also for low- and middle-income countries to adopt telehealth, mHealth, eLearning, electronic health records and big data. EHealth initiatives are underway in 83% of WHO Member States, and 90% of them

Table 2: Case study overview: care.data National Health Service (NHS) England [53, 54].

NHS launched care.data in 2013 as an initiative to collect and store patient data from GPs (general practitioners) around the country in the Health and Social Care Information Centre database (HSCIC; now NHS Digital).
HSCIS already collected hospital data. Analysing GPs data as well was supposed to improve outcomes and customer service, as well as to further understanding of diseases and treatments.
Despite initial endorsement by various professional societies, strong public reactions against the initiative were triggered by concerns about privacy, lack of transparency regarding data access and the involvement of commercial entities.
Reports by the National Data Guardian and the Care Quality Commission that highlighted that inadequacies in transparency and privacy led to the discontinuation of care.data.
The reports emphasised that citizens should be able to exercise their “right to know how their data are safeguarded. They should be included in conversations about the potential benefits that responsible use of their information can bring. They must be offered a clear choice about whether they want to allow their information to be part of this.”
Lessons learned: in order to build public trust in the use of health and care data, initiatives need to meet criteria of trustworthiness, transparency, open communication and a clear sense of the distribution of benefits.

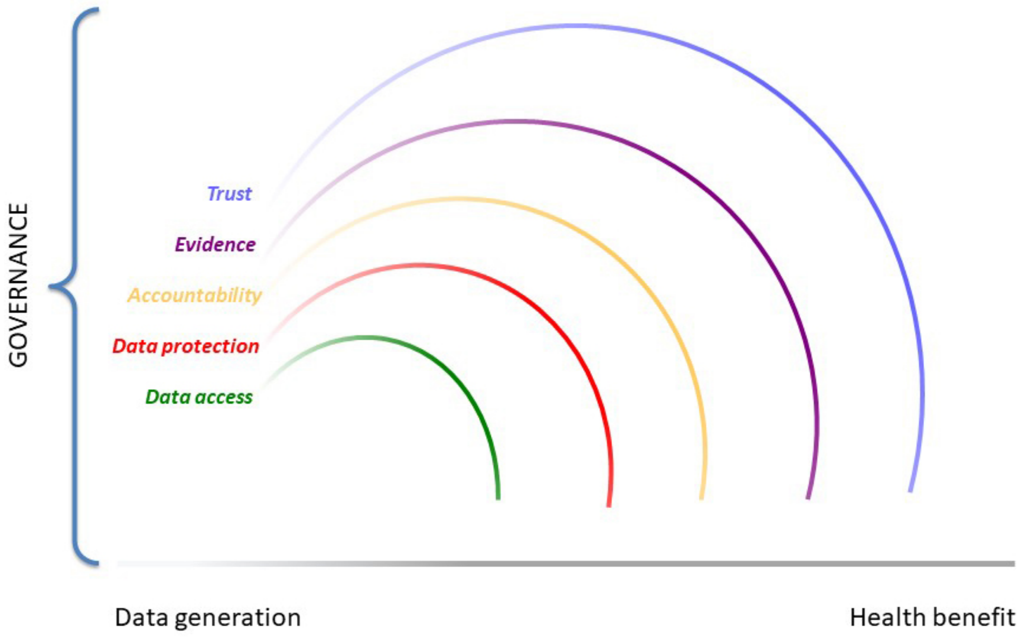
have an eHealth strategy. Different forms of digital health and digital health technology, however, present different patterns of global distribution, with telemedicine being more widely spread than electronic health records, which are more commonly used than big data in healthcare settings. Therefore, despite the fact that digital health represents a global phenomenon, it is adopted and implemented differently across the globe.

Not surprisingly, from a global perspective the governance of health data appears patchy, with only about half of WHO countries having specific privacy protections in place for personal health data. Robust national data governance frameworks tailored to the needs of real populations are thus considered a precondition for digital health to deliver sustained health benefits and to meet global health objectives such as universal health coverage. In addition, the development of international interoperability standards should continue in order to improve the capacity to monitor health needs and to deliver more effective interventions.

International policy organisations have addressed data governance issues for digital health from a global perspective. The OECD, for instance, has published a set of recommendations for health data governance [61]. Besides endorsing the idea that better health information systems and more efficient data use can improve healthcare provision, the OECD focuses on ways to maximise the usability of data for public policy, ensuring that health data processing serves the public interest, and secures public trust in data-driven health systems. To this aim, the OECD highlights several areas of intervention, including: promoting public engagement of a wide array of stakeholders; fostering collaboration to enhance interoperability and data sharing; providing clear information to individual data subjects; ensuring appropriate informed consent procedures; pursuing accurate review of data access and data processing requests; promoting transparency through public information about data use; and adopting effective control and safeguard mechanisms to protect personal data.

At the European level, the recently promulgated General Data Protection Regulation [62], which replaced the Data Protection Directive of 1995, aims at creating a more homogeneous legal framework in European Union Member States for the governance of personal data, including personal health data. This new framework stresses the importance of explicit consent to data processing, but recognises that explicit consent is not always possible in the domain of scientific research, in which data originally collected for one project are likely to be re-used by multiple researchers for purposes unrelated to the initial one. The GDPR also recognises that data processing can take place without consent if there is a pressing public health need to be ad-

Figure 1: Conditions of innovation in digital health. This graph describes the conditions for innovation in digital health, for both licensed and non-licensed products and applications. Along the continuum from data generation to health impact, several conditions need to be fulfilled for digital health applications to have a tangible effect on individual and public health. To begin with, sufficient amounts of health data about individuals, as well as other types of data helpful to the detection, treatment and monitoring of health conditions in peoples and populations, need to be accessible to developers. Secondly, digital health products need to comply with data protection and privacy requirements in the countries in which they operate. Third, accountability mechanisms should be in place to trace responsibility for data uses and their consequences on individuals, families and communities. Accountability also ensures transparent communication of health relevant information to data subjects. Fourth, solid evidence of safety and efficacy should back medical claims of digital health products. More rigidly enforced evidentiary standards – including cost-effectiveness requirements – will foreseeably apply to digital health products seeking license from national regulatory agencies (such as the FDA or EMA). Yet, also non-licensed products can and should have sufficient evidentiary bases. Only the fulfilment of all such conditions creates trust in developers and regulators of digital health products and is conducive to fair benefit sharing of digital health innovation.



dressed. Similarly, certain informational rights such as the right to have one's data erased can be limited in the name of public health emergencies, while certain sensitive data – like genetic data, for instance – can enjoy special protections set by individual member states. At any rate, the governance of data processing for research purposes and the processing of data from health registries remain subject to national rules. In terms of governance, the GDPR puts the burden of demonstrating compliance with its provisions entirely on the shoulders of data controllers, thus considerably raising the bar of accountability demands in comparison with the previous data protection directive.

Governance should enable digital health innovation to address the challenges discussed above, which include not only accountability but also privacy, quality of evidence, data access and sharing, and ultimately trust. Essentially, these are five key conditions that can determine whether digital health innovation can lead to health benefit (fig. 1). It remains to be seen whether, and how, a global governance approach can achieve this. For any approach it will be crucial to ensure that all stakeholders are involved and engaged. In this respect, the emphasis that the WHO puts on public participation and engagement of broad arrays of stakeholders aptly recognises the need to ensure that digital health serves the public interest and facilitates patients' engagement in health-related decisions.

The Swiss context

The development of digital health faces similar challenges in most developed countries. However, individual countries face these challenges to different degrees depending on the quality of their IT infrastructure, regulatory frameworks, healthcare systems and so on. Currently, a number of significant developments mark a turning point for digital health in Switzerland. First, the enactment of the Swiss electronic patient dossier legislation [63], on 15 April 2017, is an important step toward further digitalisation in the country's healthcare sector. The dossier, a voluntary electronic collection of personal medical documents, is designed to provide healthcare professionals with easier access to patient information, thus improving the safety and accuracy of diagnosis, with the ultimate goal of a positive impact on patient treatment and care. Whereas some Swiss regions have already put digital patient dossiers in place (see for instance the Geneva health information exchange *e-toile* [64], or the project *dossier patient partagé - Infomed* in the canton of Valais [65]), no provider has been officially certified to date, and both the legal and organisational prerequisites are being gradually implemented this year with a view to have the system running by mid-2018. Even though Switzerland benefited from the insights of major ongoing eHealth projects in Europe [66], the process towards more centralisation of national digital health policy-making has been slow and non-linear [67, 68]. Nevertheless, the electronic patient dossier has overcome various political and organisational hurdles and can help advance other digital health services and initiatives, such as the cross-border harmonisation of e-medication records [69].

One crucial factor for the development of digital health is data accessibility. Ideally, data should be made available for further research uses that promise progress in individual or population health, and research and clinical institutions should be willing to open up their patients' data for that aim. Despite repeated appeals on the importance of data access, however, this practice is still implemented to an insufficient degree. Some barriers to data sharing are more regulatory in nature, such as the inability of data subjects to truly consent to uses that are not foreseeable at the moment of data collection. Some others are more organisational, as in the case of institutions that are reluctant to share data for liability issues. Currently existing patient data are collected through diverse technological systems and with variations in the consent that authorises further uses.

The second important development in Switzerland aims to address this issue through the proposal of a national broad consent template. Spearheaded by the Swiss Academy of Medical Sciences, a so-called "general consent" has been developed after extensive consultation with various stakeholders. The aim of this broad consent is to harmonise the conditions under which further data uses can take place. The model of broad consent has been highly debated in the bioethics literature, however, and commentaries range from full approval to complete rejection [70–73]. Broad consent may not be the ultimate solution to conducting ethical secondary uses of data. However, if accompanied by robust oversight and accountability systems it can be a pragmatic solution that facilitates ethical digital health research [74].

The third relevant development in Switzerland is the launch of the Swiss Personalized Health Network (SPHN) [75] – a national initiative designed to build the necessary infrastructure to improve the utilisation of health-related data for research and innovation. The development of digital health, as that of other data-driven activities, depends on the development of appropriate technical standards to make data securely exchangeable and efficiently computable. Accordingly, the SPHN aims to develop interoperability standards that will enhance data accessibility for research uses in Switzerland. The SPHN's vision on data governance is based on an ethics framework including four principles: respect for persons, data fairness, privacy, and accountability. Such a soft law instrument, while indicating the direction for improving data sharing, is also flexible enough to adapt to stakeholders' organisational needs.

Public engagement

Citizens and patients are increasingly becoming the driving forces behind digital health developments [76, 77]. The extensive adoption and sustainability of health data exchange thus depend upon information technology that facilitates patient engagement and the earning of public trust [78]. To build on the support of the public, it should be made clear that digital health is a tool for citizens and professionals alike [79–81]. This is a condition for fostering trust around digital health [82]. Furthermore, public policy needs take into account the digital divide and the capacity of citizens to engage with e-health [83–86]. And whereas it is certainly important to promote collaboration among healthcare professionals and institutions, other agents, such as startups and the industry in general, ought to be included in the country's digital health transformation with mechanisms to incentivise partnership, investments and data sharing [87–89]. This can take the form of public/private partnerships [90], such as the Digital Switzerland Initiative [91] and the Opendata.ch Foundation [92].

Other innovative models to leverage private initiatives and foster public engagement are emerging. In Switzerland, the MIDATA cooperative is a case in point [93]. MIDATA offers data subjects the possibility of storing health data from different sources and leaves it to the data subjects to decide collectively on data access requests [94, 95]. All data contributors are equal shareholders of the cooperative, which is a not-for-profit entity and will re-invest any potential income generated by granting access to its data. This unique model is already active in digital health-related projects in Switzerland and will promote the inclusion of patient-generated data that are needed to develop digital health into clinical applications.

Conclusion

Innovation in digital health faces several ethical and policy challenges. We have argued that, for digital health products and applications to produce tangible innovation and health impacts, be it at the individual or at the population level, five conditions need to be met. First, data are of paramount importance for digital health: access to sufficient amounts of data is thus a primary condition for the development of innovative diagnostic, therapeutic and monitoring tools is this area. Second, alignment with existing legal provisions regarding data protection, data security and privacy are key to digital health innovation. Legal frameworks can thus have a major impact in facilitating or hindering progress in this field. Nonetheless, legal provisions do not address the full range of ethical issues in data processing. Nor do they cover the full spectrum of legitimate concerns of data subjects. Third, robust and transparent accountability mechanisms should ensure the precise identification of responsibility for data uses and their consequences on individuals, families and communities. What is more, accountability also sets up mechanisms for communicating health relevant information to data subjects. Fourth, evidence of safety and efficacy is a significant condition for the success of digital health. Licensed digital health products and applications will have to go through extensive assessment processes and will have to meet cost-effectiveness requirement before they can be reimbursed by insurers and public healthcare systems. This does not, however, mean that unlicensed products and applications can lack some form of evidence to back up their claims. Fulfilling these requirements will foster the fifth condition for digital health innovation, that is, trust in both developers and regulators, which in turn will facilitate the uptake of digital health by healthcare providers and lead to fair benefit sharing of digital health innovation.

Disclosure statement

No financial support and no other potential conflict of interest relevant to this article was reported.

References

1 U.S. Food and Drug Administration. [Internet]. Silver Spring: U.S. Food and Drug Administration; c1995-2017 [cited 2017 May 3]. U.S. Department of Health and Human Services; [about 2 screens]. Available from: https://www.fda.gov/medicaldevices/digitalhealth/#mobileapp.

2 Adheretech.com. [Internet]. New York: AdhereTech Inc.; c2017 [cited 2017 July 21]. Available from: https://adheretech.com/

3 Airstrip.com. [Internet]. Texas: Airstrip Technologies; c2017 [cited 2017 July 21]. Available from: http://www.airstrip.com/

4 Food and Drug Administration. Section 5_510(k) Summary revised v3 [Internet]. Maryland: FDA; 2014 [cited 2017 July 21]. Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf13/k133450.pdf

5 Alivecor.com. [Internet]. San Francisco: AliveCor Inc.; c2017 [cited 2017 July 21]. Available from: https://www.alivecor.com/

6 Food and Drug Administration. Section 5_510(k) Summary revised v3 [Internet]. Maryland: FDA; 2014 [cited 2017 July 21]. Available from: https://www.fda.gov/cdrh/510k/k122356.pdf

7 Bluesparktechnologies.com. [Internet]. Ohio: Blue Spark Technologies Inc.; c2017 [cited 2017 July 21]. Available from: http://bluesparktechnologies.com/index.php/products-and-services/temptraq

8 Food and Drug Administration. Section 5_510(k) Summary revised v3 [Internet]. Maryland: FDA; 2015 [cited 2017 July 21]. Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf14/k143267.pdf

9 Natural Cycles.com [Internet]. Stockholm: Natural Cycles [cited 2017 July 21]. Available from: https://www.naturalcycles.com/en

10 Proteus.com. [Internet]. Redwood City: Proteus Digital Health; c2017 [cited 2017 July 21]. Available from: http://www.proteus.com/.

11 Food and Drug Administration. Ingestible Event Marker [Internet]. Maryland: FDA; 2014 [cited 2017 July 21]. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn_template.cfm?id=k133263

12 Butterflynetinc.com. [Internet]. New York: Butterfly Network Inc. [cited 2017 July 21]. Available from: https://www.butterflynetinc.com/

13 Fitbit.com. [Internet]. San Francisco: Fitbit; c2017 [cited 2017 July 21]. Available from: https://www.fitbit.com/aria

14 Happify.com. [Internet]. New York: Happify Inc.; c2017 [cited 2017 July 21]. Available from: https://my.happify.com/public/contact/

15 Myfitnesspal.com. [Internet]. MyFitnessPal Inc. c2005-2017 [cited 2017 July 21]. Available from: https://www.myfitnesspal.com/.

16 Elenko E, Speier A, Zohar D. A regulatory framework emerges for digital medicine. Nat Biotechnol. 2015;33(7):697–702. doi: http://dx.doi.org/10.1038/nbt.3284. PubMed.

17 Eisenstein M. Miniature wireless sensors presage smart phone medicine. Nat Biotechnol. 2012;30(11):1013–4. doi: http://dx.doi.org/10.1038/nbt1112-1013. PubMed.

18 Flahault A, Geissbuhler A, Guessous I, Guérin P, Bolon I, Salathé M, et al. Precision global health in the digital age. Swiss Med Wkly. 2017;147:w14423. PubMed.

19 Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. PLOS Comput Biol. 2012;8(7):e1002616. doi: http://dx.doi.org/10.1371/journal.pcbi.1002616. PubMed.

20 Vayena E, Dzenowagis J. Langfeld M. [Internet]. Geneva: World Health Organization; c1948-2017 [cited 2017 May 3]. U.S. United Nations; [about 1 screen]. Available from: http://www.who.int/ehealth/resources/ecosystem/en/.

21 Groves P, Kayyali B, Knott D, Van Kuiken S. The 'big data'revolution in healthcare. New York: McKinsey & Company; 2013.

22 Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. JAMA. 2014;311(24):2479–80. doi: http://dx.doi.org/10.1001/jama.2014.4228. PubMed.

23 Blasimme A. Healthcare meets big data: the science and politics of precision medicine. In: Blanchard A, Strand R, editors. Social, ethical and economic aspects of cancer biomarkers. Kokstad: Megaloceros Press; 2017. p. 95-110.

24 Vayena E, Blasimme A. Biomedical big data: new models of control over access, use and governance. J Bioeth Inq. 2017;14(4):501–13. doi: http://dx.doi.org/10.1007/s11673-017-9809-6. PubMed.

25 Vayena E, Gasser U, Wood A, O'Brian DR, Altman M. Elements of a new ethical framework for big data research. Wash Lee Law Rev. 2016;72(3):420–41. http://lawreview.journals.wlu.io/elements-of-a-new-ethical-framework-for-big-data-research/

26 US Department of Veterans Affairs [Internet]. Washington DC: The Department; c1930-2017 [cited 2017 Jul 20]. Office of Public and Intergovernmental Affairs; [about 2 screens]. Available from: https://www.va.gov/opa/pressrel/pressrelease.cfm?id=2806

27 Wojcicki A. Power of One Million. 2015 June 18 [cited 2017 Jul 20]. In: 23andMe. Blog [Internet]. Mountain View: 23andMe, Inc. c2007-2017. [about 2 screens]. Available from: https://blog.23andme.com/news/one-in-a-million/.

28 Kaiser Permanente [Internet]. Oakland: Kaiser Foundation Health Plan, Inc.; c2017 [cited 2017 Jul 20]. [about 2 screens]. Available from: https://share.kaiserpermanente.org/article/kaiser-permanente-launches-research-biobank-aims-to-transform-health/.

29 Khoury MJ, Evans JP. A public health perspective on a national precision medicine cohort: balancing long-term knowledge generation with early health benefit. JAMA. 2015;313(21):2117–8. doi: http://dx.doi.org/10.1001/jama.2015.3382. PubMed.

30 Blasimme A, Vayena E. “Tailored-to-you” - public engagement and the political legitimation of precision medicine. Perspect Biol Med. 2016;59(2):172–88. doi: http://dx.doi.org/10.1353/pbm.2017.0002.

31 Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. Am J Prev Med. 2016;50(3):398–401. doi: http://dx.doi.org/10.1016/j.amepre.2015.08.031. PubMed.

32 Rubin R. Precision medicine: the future or simply politics? JAMA. 2015;313(11):1089–91. doi: http://dx.doi.org/10.1001/jama.2015.0957. PubMed.

33 Coote JH, Joyner MJ. Is precision medicine the route to a healthy world? Lancet. 2015;385(9978):1617. doi: http://dx.doi.org/10.1016/S0140-6736(15)60786-3. PubMed.

34 Jarow JP, LaVange L, Woodcock J. Multidimensional evidence generation and FDA regulatory decision making: defining and using “real-world” data. JAMA. 2017;318(8):703–4. doi: http://dx.doi.org/10.1001/jama.2017.9991. PubMed.

35 Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015;372(9):793–5. doi: http://dx.doi.org/10.1056/NEJMp1500523. PubMed.

36 Khoury MJ, Evans JP. A public health perspective on a national precision medicine cohort: balancing long-term knowledge generation with early health benefit. JAMA. 2015;313(21):2117–8. doi: http://dx.doi.org/10.1001/jama.2015.3382. PubMed.

37 Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. Am J Prev Med. 2016;50(3):398–401. doi: http://dx.doi.org/10.1016/j.amepre.2015.08.031. PubMed.

38 Flahault A, Geissbuhler A, Guessous I, Guérin P, Bolon I, Salathé M, et al. Precision global health in the digital age. Swiss Med Wkly. 2017;147:w14423. PubMed.

39 Sankar PL, Parker LS. The Precision Medicine Initiative’s All of Us Research Program: an agenda for research on its ethical, legal, and social issues. Genet Med. 2017;19(7):743–50. doi: http://dx.doi.org/10.1038/gim.2016.183. PubMed.

40 Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLOS Comput Biol. 2015;11(2):e1003904. doi: http://dx.doi.org/10.1371/journal.pcbi.1003904. PubMed.

41 Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?”. Lancet. 2005;365(9453):82–93. doi: http://dx.doi.org/10.1016/S0140-6736(04)17670-8. PubMed.

42 Frueh FW. Back to the future: why randomized controlled trials cannot be the answer to pharmacogenomics and personalized medicine. Pharmacogenomics. 2009;10(7):1077–81. doi: http://dx.doi.org/10.2217/pgs.09.62. PubMed.

43 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. doi: http://dx.doi.org/10.1038/nature14539. PubMed.

44 Anderson C. The end of theory: the data deluge makes the scientific method obsolete. WIRED [Internet]. 2008 Jun [cited 2017 May 8]; [about 3 p.]. Available from: https://www.wired.com/2008/06/pb-theory/.

45 Pigliucci M. The end of theory in science? EMBO Rep. 2009;10(6):534. doi: http://dx.doi.org/10.1038/embor.2009.111. PubMed.

46 Price WN. Black-Box Medicine. Harv. J. L. & Tech. 2015;28(2):419.

47 Mazzocchi F. Could Big Data be the end of theory in science? A few remarks on the epistemology of data-driven science. EMBO Rep. 2015;16(10):1250–5. doi: http://dx.doi.org/10.15252/embr.201541001. PubMed.

48 Khoury MJ, Ioannidis JPA. Big data meets public health. Science. 2014;346(6213):1054–5. doi: http://dx.doi.org/10.1126/science.aaa2709. PubMed.

49 Vayena E, Mastroianni A, Kahn J. Caught in the web: informed consent for online health research. Sci Transl Med. 2013;5(173):173fs6. doi: http://dx.doi.org/10.1126/scitranslmed.3004798. PubMed.

50 Gayle D, Topping A, Sample I, Marsh S, Dodd V. NHS seeks to recover from global cyber-attack as security concerns resurface. The Guardian 13 May 2017. Available from https://www.theguardian.com/society/2017/may/12/hospitals-across-england-hit-by-large-scale-cyber-attack

51 Information Commissioner’s Office [Internet]. London: Information Commissioner’s Office; c2017 [cited 2017 Jul 20]; [about 3 screens]. Available from: https://ico.org.uk/action-weve-taken/data-security-incident-trends/

52 Tasioulas J, Vayena E. The place of human rights and the common good in global health policy. Theor Med Bioeth. 2016;37(4):365–82. doi: http://dx.doi.org/10.1007/s11017-016-9372-x. PubMed.

53 Department of Health. Freeman G [Internet]. London: Gov.UK; c2016 [cited 2017 Jul 20]; Available from: https://www.gov.uk/government/speeches/review-of-health-and-care-data-security-and-consent

54 Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. J Med Ethics. 2015;41(5):404–9. doi: http://dx.doi.org/10.1136/medethics-2014-102374. PubMed.

55 Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005;330(7494):765. doi: http://dx.doi.org/10.1136/bmj.38398.500764.8F. PubMed.

56 Hsu W, Markey MK, Wang MD. Biomedical imaging informatics in the era of precision medicine: progress, challenges, and opportunities. J Am Med Inform Assoc. 2013;20(6):1010–3. doi: http://dx.doi.org/10.1136/amiajnl-2013-002315. PubMed.

57 Yang GZ, Cambias J, Cleary K, Daimler E, Drake J, Dupont PE, et al. Medical robotics - regulatory, ethical, and legal considerations for increasing levels of autonomy. Sci Robot. 2017;2(4):eaam8638. doi: http://dx.doi.org/10.1126/scirobotics.aam8638.

58 Elenko E, Underwood L, Zohar D. Defining digital medicine. Nat Biotechnol. 2015;33(5):456–61. doi: http://dx.doi.org/10.1038/nbt.3222. PubMed.

59 Elenko E, Speier A, Zohar D. A regulatory framework emerges for digital medicine. Nat Biotechnol. 2015;33(7):697–702. doi: http://dx.doi.org/10.1038/nbt.3284. PubMed.

60 World Health Organization. Global diffusion of eHealth: making universal health coverage achievable. Report of the third global survey on eHealth [Internet]. Geneva: WHO Document Production Services; 2016 [cited 2017 Jul 20]. Available from: http://apps.who.int/iris/bitstream/10665/252529/1/9789241511780-eng.pdf?ua=1.

61 Organization of Economic Cooperation and Development. Recommendations of OECD Council on Health Data Governance [cited 21 July 2017] Available from http://www.oecd.org/health/health-systems/Recommendation-of-OECD-Council-on-Health-Data-Governance-Booklet.pdf

62 General Data Protection Regulation [Internet]. Brussels: European Union; 2016 [cited 2017 July 21]. Available from: http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf

63 e-health-suisse.ch. [Internet]. Bern: eHealth Suisse; c2007-2017 [cited 2017 Jun 1]. Kompetenz- und Koordinationsstelle von Bund und Kantonen; [about 2 screens]. Available from: https://www.e-health-suisse.ch/elektronisches-patientendossier.html [available in German, French and Italian only].

64 Rosemberg A, Schmid A, Plaut O. MonDossierMedical.ch - the personal health record for every Geneva citizen. Stud Health Technol Inform. 2016;225:700–2. doi: http://dx.doi.org/10.3233/978-1-61499-658-3-700. PubMed.

65 Gnaegi A, Michelet C. Dossier patient partagé Infomed, qu’en pensent les médecins. Swiss Medical Informatics. 2016;32. doi: http://dx.doi.org/10.4414/smi.32.361.

66 Lovis C, Looser H, Schmid A, Wagner J, Wyss S. eHealth in Switzerland - building consensus, awareness and architecture. Stud Health Technol Inform. 2011;165:57–62. doi: http://dx.doi.org/10.3233/978-1-60750-735-2-57. PubMed.

67 De Pietro C, Camenzind P, Sturny I, Crivelli L, Edwards-Garavoglia S, Spranger A, et al. Switzerland: health system review. Health Syst Transit. 2015;17(4):1–288, xix. PubMed.

68 Dossier électronique du patient: la Suisse romande avance en ordre disperse [Electronic patient records: Switzerland is moving forward haphazardly]. Rev Med Suisse. 2015;11(499):2411. Article in French.

69 Gall W, Aly AF, Sojer R, Spahni S, Ammenwerth E. The national e-medication approaches in Germany, Switzerland and Austria: A structured comparison. Int J Med Inform. 2016;93:14–25. doi: http://dx.doi.org/10.1016/j.ijmedinf.2016.05.009. PubMed.

70 Grady C, Eckstein L, Berkman B, Brock D, Cook-Deegan R, Fullerton SM, et al. Broad consent for research with biological samples: Workshop conclusions. Am J Bioeth. 2015;15(9):34–42. doi: http://dx.doi.org/10.1080/15265161.2015.1062162. PubMed.

71 Helgesson G. In defense of broad consent. Camb Q Healthc Ethics. 2012;21(1):40–50. doi: http://dx.doi.org/10.1017/S096318011100048X. PubMed.

72 Hofmann B. Broadening consent--and diluting ethics? J Med Ethics. 2009;35(2):125–9. doi: http://dx.doi.org/10.1136/jme.2008.024851. PubMed.

73 Sheehan M. Can broad consent be informed consent? Public Health Ethics. 2011;4(3):226–35. Published online August 3, 2011. doi: http://dx.doi.org/10.1093/phe/phr020. PubMed.

74 Vayena E, Gasser U. Between openness and privacy in genomics. PLoS Med. 2016;13(1):e1001937. doi: http://dx.doi.org/10.1371/journal.pmed.1001937. PubMed.

75 Swiss Personalized Health Network [Internet]. Bern: Swiss Academy of Medical Sciences (SAMS); c1943-2017 [cited 2017 June 1]. Swiss

Academy of Medical Sciences; [about 2 screens]. Available from: http://www.samw.ch/en/Projects/SPHN.html.

76 Haeusermann T, Greshake B, Blasimme A, Irdam D, Richards M, Vayena E. Open sharing of genomic data: Who does it and why? PLoS One. 2017;12(5):e0177158. doi: http://dx.doi.org/10.1371/journal.pone.0177158. PubMed.

77 Blasimme A, Vayena E. Becoming partners, retaining autonomy: ethical considerations on the development of precision medicine. BMC Med Ethics. 2016;17(1):67. doi: http://dx.doi.org/10.1186/s12910-016-0149-6. PubMed.

78 Walker DM, Sieck CJ, Menser T, Huerta TR, Scheck McAlearney A. Information technology to support patient engagement: where do we stand and where can we go? J Am Med Inform Assoc. 2017;24(6):1088–94; Epub ahead of print. doi: http://dx.doi.org/10.1093/jamia/ocx043. PubMed.

79 Moen A, Hackl WO, Hofdijk J, Van Gemert-Pijnen L, Ammenwerth E, Nykänen P, et al. eHealth in Europe - status and challenges. Yearb Med Inform. 2013;8:59–63. PubMed.

80 Patil S, Lu H, Saunders CL, Potoglou D, Robinson N. Public preferences for electronic health data storage, access, and sharing - evidence from a pan-European survey. J Am Med Inform Assoc. 2016;23(6):1096–106. doi: http://dx.doi.org/10.1093/jamia/ocw012. PubMed.

81 Leuthold M. Patients as partners for improving safety. World Hosp Health Serv. 2014;50(3):20–2. PubMed.

82 Tripathi M, Delano D, Lund B, Rudolph L. Engaging patients for health information exchange. Health Aff (Millwood). 2009;28(2):435–43. doi: http://dx.doi.org/10.1377/hlthaff.28.2.435. PubMed.

83 Mantwill S, Monestel-Umaña S, Schulz PJ. The relationship between health literacy and health disparities: a systematic review. PLoS One. 2015;10(12):e0145455. doi: http://dx.doi.org/10.1371/journal.pone.0145455. PubMed.

84 Nohr C, Wong MC, Turner P, Almond H, Parv L, Gilstad H, et al. Citizens' access to their digital health data in eleven countries - a comparative study. Stud Health Technol Inform. 2016;228:685–9. doi: http://dx.doi.org/10.3233/978-1-61499-678-1-685. PubMed.

85 Sak G, Rothenfluh F, Schulz PJ. Assessing the predictive power of psychological empowerment and health literacy for older patients' participation in health care: a cross-sectional population-based study. BMC Geriatr. 2017;17(1):59. doi: http://dx.doi.org/10.1186/s12877-017-0448-x. PubMed.

86 Romano MF, Sardella MV, Alboni F, Russo L, Mariotti R, Nicastro I, et al. Is the digital divide an obstacle to e-health? An analysis of the situation in Europe and in Italy. Telemed J E Health. 2015;21(1):24–35. doi: http://dx.doi.org/10.1089/tmj.2014.0010. PubMed.

87 Feldman HH, Haas M, Gandy S, Schoepp DD, Cross AJ, Mayeux R, et al.; One Mind for Research and the New York Academy of Sciences. Alzheimer's disease research and development: a call for a new research roadmap. Ann N Y Acad Sci. 2014;1313(1):1–16. doi: http://dx.doi.org/10.1111/nyas.12424. PubMed.

88 Salter H, Holland R. Biomarkers: refining diagnosis and expediting drug development - reality, aspiration and the role of open innovation. J Intern Med. 2014;276(3):215–28. doi: http://dx.doi.org/10.1111/joim.12234. PubMed.

89 Leyens L, Reumann M, Malats N, Brand A. Use of big data for drug development and for public and personal health and care. Genet Epidemiol. 2017;41(1):51–60. doi: http://dx.doi.org/10.1002/gepi.22012. PubMed.

90 Chataway J, Fry C, Marjanovic S, Yaqub O. Public-private collaborations and partnerships in stratified medicine: making sense of new interactions. N Biotechnol. 2012;29(6):732–40. doi: http://dx.doi.org/10.1016/j.nbt.2012.03.006. PubMed.

91 digitalswitzerland.com. [Internet]. Zürich: digitalswitzerland; c2015 [cited 2017 June 1]. Available from: http://digitalswitzerland.com.

92 opendata.ch. [Internet]. Zurich: Verein Opendata.ch, the Swiss Chapter of the Open Knowledge Foundation; c2012 [cited 2017 June 1]. Available from: https://opendata.ch.

93 www.midata.coop. [Internet]. Zürich: MIDATA Genossenschaft; c2017 [cited 2017 Jul 20]. Available from: https://www.midata.coop/.

94 Hafen E, Kossmann D, Brand A. Health data cooperatives - citizen empowerment. Methods Inf Med. 2014;53(2):82–6. doi: http://dx.doi.org/10.3414/ME13-02-0051. PubMed.

95 Hafen E. Midata Cooperatives - Citizen-Controlled Use of Health Data Is a Pre-Requiste for Big Data Analysis, Economic Success and a Democratization of the Personal Data Economy. Tropical Medicine & International Health [Internet]. 2015 Sep [cited 2017 July 20]; 20 (Supplement 1):129. [about 1 p.] Available from: https://insights.ovid.com/tropical-medicine-international-health/tmih/2015/09/001/midata-cooperatives-citizen-controlled-use-health/313/00060771.

Viewpoint

Reimagining Human Research Protections for 21st Century Science

Cinnamon Bloss[1,2,3*], PhD; Camille Nebeker[1,3*], MS, EdD; Matthew Bietz[4], PhD; Deborah Bae[5], MPA, MBA; Barbara Bigby[6], CCRC, MA; Mary Devereaux[7], PhD; James Fowler[8,9], PhD; Ann Waldo[10], JD, CIPP; Nadir Weibel[1,11], PhD; Kevin Patrick[1,3], MS, MD; Scott Klemmer[11,12], PhD; Lori Melichar[5], PhD

[1]Center for Wireless and Population Health Systems, The Qualcomm Institute, Calit2, University of California, San Diego, La Jolla, CA, United States
[2]Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States
[3]Department of Family Medicine and Public Health, School of Medicine, University of California, San Diego, La Jolla, CA, United States
[4]Department of Informatics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, Irvine, CA, United States
[5]Robert Wood Johnson Foundation, Princeton, NJ, United States
[6]Scripps Human Research Protections Program, La Jolla, CA, United States
[7]Department of Pathology, School of Medicine, University of California, San Diego, La Jolla, CA, United States
[8]Department of Political Science, University of California, San Diego, La Jolla, CA, United States
[9]Department of Medicine, School of Medicine, University of California, San Diego, La Jolla, CA, United States
[10]Waldo Law Offices, PLLC, Washington, DC, United States
[11]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, United States
[12]Department of Cognitive Science, University of California, San Diego, La Jolla, CA, United States
[*]these authors contributed equally

Corresponding Author:
Cinnamon Bloss, PhD
Center for Wireless and Population Health Systems, The Qualcomm Institute, Calit2
University of California, San Diego
9500 Gilman Drive
Atkinson Hall (mail code 0811, office 6103)
La Jolla, CA, 92093-0811
United States
Phone: 1 858 534 9595
Fax: 1 858 534 9404
Email: cbloss@ucsd.edu

Abstract

Background: Evolving research practices and new forms of research enabled by technological advances require a redesigned research oversight system that respects and protects human research participants.

Objective: Our objective was to generate creative ideas for redesigning our current human research oversight system.

Methods: A total of 11 researchers and institutional review board (IRB) professionals participated in a January 2015 design thinking workshop to develop ideas for redesigning the IRB system.

Results: Ideas in 5 major domains were generated. The areas of focus were (1) improving the consent form and process, (2) empowering researchers to protect their participants, (3) creating a system to learn from mistakes, (4) improving IRB efficiency, and (5) facilitating review of research that leverages technological advances.

Conclusions: We describe the impetus for and results of a design thinking workshop to reimagine a human research protections system that is responsive to 21st century science.

(J Med Internet Res 2016;18(12):e329) doi:10.2196/jmir.6634

KEYWORDS

ethics committees, research; biomedical research; telemedicine; informed consent; behavioral research

Introduction

Over half a century ago, in response to egregious cases of research participant mistreatment, the US government proposed prospective review of research involving human participants. This prospective review process is what we now know as the institutional review board (IRB) system. Today, IRBs are firmly entrenched within the fabric of academic research institution, with estimates putting the number of IRBs in the United States at around 6000 (I Prichard, Senior Advisor to the Director of the Office for Human Research Protections, Department of Health and Human Services; oral communication, September 2014).

IRBs have aimed to serve an important function, which is to protect human research participants. While IRBs have helped address this critical need, the IRB system has not kept pace with the evolution of research methods and practices or current and emerging trends in science and technology. The fact that the system has become antiquated calls into question whether the IRB continues to foster the protection of human research participants per the principles originally put forth in the Belmont Report [1]. New forms of research enabled by technological advances in information technology and data science appear to be particularly challenging to IRBs [2], yet clear standards to guide best practices are not well established [3-5]. We propose that the time has come to reimagine and ultimately work toward redesigning our human research protections system so that it is responsive to both the evolution of general research practices and new forms of research enabled by technological advances—what we refer to here as 21st century science. This is critical for the proper protection of research participants, ethical and efficient use of research resources, and facilitation of research insights important for human health specifically and knowledge production more generally.

A Changing Research Landscape

The IRB model was created when research was typically conducted by a single principal investigator in a single academic institution, and when data were both scarce and expensive to collect. Today, multiple principal investigator, multi-institution, and even multicountry studies are common, and such studies have resulted in unprecedented insights regarding human health. Researchers now need, or are even expected to share, data between different universities, across entities in different sectors (eg, universities, corporations, and nonprofits) and frequently across international borders. It also used to be that the scale of research was closely linked to the research methods. For instance, intervention studies were conducted with small numbers of participants in tightly controlled environments, and large-N surveys tended to collect data in ways that limited the possibility of individual identification and promoted easy anonymization. Today, expanding computational capabilities, social media, and broad research networks allow us to conduct an intervention study on Facebook with millions of participants [6], engage patients using mobile phone technology [7], study the whole genomes of thousands of individuals [8], or collect digital traces of human activity [9] at such granular levels that reidentification of individuals is possible if one possesses the right tools and expertise [10]. While traditional approaches to research require collecting only as much data as is necessary to test a hypothesis, data mining and other big data techniques derive their power from large data sets, where it may be impossible to determine, a priori, which variables will be of interest.

A Static Regulatory Environment

In contrast to the evolving research practices landscape, procedures for research oversight have been markedly static. The Common Rule, which refers to a set of regulations that specify the procedures for establishing and operating IRBs, was adopted in 1991, and the Belmont Report and the Common Rule remain the primary sources for guiding review of human research. In 2011, recognizing that these regulations had not kept pace with the evolving human research enterprise, the US Department of Health and Human Services issued an Advance Notice of Proposed Rulemaking (ANPRM) aimed at "enhancing protections for research subjects and reducing burden, delay, and ambiguity for investigators" [11]. In 2015, the ANPRM transitioned to a Notice of Proposed Rulemaking (NPRM), the next step in the process to update federal regulations [12]. The NPRM updates include, for example, new consent requirements for biological specimens, use of a central IRB for multisite studies, and changes to procedures for determining exempt versus expedited study review categories. Use of a central IRB is particularly contentious, with concerns focusing on whether protection of participants may be compromised for the increased efficiency of a single IRB [13]. Regardless, in June 2016, the US National Institutes of Health published a policy requiring single IRB review for multisite studies [14]. While the NPRM reflects important and potentially promising activity toward IRB system improvements, many have questioned or objected outright to some of the proposed changes, and even supporters have suggested they are not ideal [15]. Most recently, a report by the National Academies of Science, Engineering, and Medicine Committee on Federal Research Regulations and Reporting Requirements issued a report criticizing the NPRM, citing that the proposed changes would be detrimental to advancing research [16]. This committee recommended that the US Congress authorize the presidential appointment of a national commission to examine and update the ethical and regulatory frameworks governing human research protections. Regardless of whether and to what extent the Common Rule or principles of the Belmont Report are revised, the extent to which IRBs can keep abreast of changes in the research landscape and be responsive to studies that leverage emerging technologies remains questionable at best.

A Flawed Institutional Review Board System

There is increasing evidence that the IRB system is deeply and inherently flawed [17]. Lidz and colleagues captured the tip of the proverbial iceberg in their study of 20 IRB panels at 10 large medical institutions, where they documented 104 protocol reviews [18]. They found that IRBs consistently discussed the informed consent document, one of the Common Rule's central mandates, and requested changes to the consent document in 88% of those cases. They also documented a disturbing observation, which was that other elements of the Common

Rule (eg, data monitoring and protection of vulnerable populations) that are intended to promote research ethics were rarely discussed. Furthermore, in studies that exceeded minimal risk of harm, 21% of reviews did not address the inclusion of strategies to minimize risk. Likewise, they noted that 50% of reviews did not compare risks and benefits, and 60% of the protocols that excluded groups of potential research subjects without explicit justification were not discussed. They also found that critical review of the research design and methodology was not uncommon, and that IRBs often requested that investigators make changes to their research design, which is typically considered outside the purview and mission of IRBs.

Taken together, these observations call into question whether IRB members are sufficiently familiar with the standards intended to guide their review of research. Furthermore, while this is the case with respect to studies that leverage traditional research methods (eg, clinical trials), these concerns are magnified when the studies under review involve emerging technologies and nontraditional methods that the IRB was not originally designed to handle and that IRB members often do not understand. Examples of such studies are those that use smartphone capabilities to measure physical activity, social media to assess adverse drug reactions, or N-of-1 genome sequencing studies for diagnosis of rare disease. Such studies raise new and nuanced ethical issues regarding participant privacy, informed consent, and data security. Some of these novel methods also inadvertently include nonparticipants [19] or “bystanders” [20] in the research record, raising potential concerns that further challenge IRB processes.

Methods

How Might We Redesign the Institutional Review Board?

In light of these issues, in January 2015 we assembled a multidisciplinary group of 11 researchers and IRB professionals drawn from academic and research institutions in San Diego, California, to consider how we might reimagine and redesign human research protections for 21st century science. The half-day workshop was set up as a brainstorming session to generate ideas for addressing IRB challenges related to review of human studies, with a particular focus on studies that leverage emerging technologies and methods. The aim was not only to stimulate creative thinking about how the existing IRB structure and process could be modified to meet the often cited challenges of the current system, but also to generate ideas for exploring entirely new ways of evaluating research to ensure that research participants are informed and protected.

Design Thinking

A central feature of this workshop was the use of design thinking strategies in the brainstorming process. Design thinking is a formal method for practical and creative resolution of problems [21] that emphasizes a phase during which the group or team focuses on generating as many ideas as possible using thoughtful prompts (eg, How might we advise as opposed to restrict? How might we simplify IRB review?). Design thinking is also considered particularly useful when the problem itself, in addition to the solution, may be unknown or ill defined at the outset of the problem-solving exercise.

Workshop Description and Stages

Workshop participants included a facilitator (SK), a cofacilitator (CB), and 9 participants (the remaining authors). A high-level goal of the session was to generate ideas for how we might reimagine and ultimately redesign the human research protections system to foster the ethical conduct of research in the changing landscape of 21st century science.

The design thinking protocol consisted of 3 primary stages. During the first stage, we asked participants to brainstorm ideas using 6 categories as prompts: (1) settings, scenarios, and steps; (2) stakeholders and extreme users; (3) utopia and dystopia; (4) change levers; (5) change agents and obstacles; and (6) things to find out. During the second stage, we asked participants to consider the ideas generated in stage 1 and to complete the sentence “How might we...?” using the stage-1 ideas as prompts. A total of 22 “How might we…?” statements were generated (see Textbox 1). From the full list of “How might we…?” statements generated in stage 2, we asked participants to select 3 ideas that they were most interested in pursuing further.

In stage 3, participants were broken into groups based on overlapping interests to further discuss and expand on specific ideas. The 5 refined “How might we…?” statements that received the most votes were (1) How might we redesign the consent form and process? (2) How might we empower researchers to protect their participants? (3) How might we learn from our efforts to protect participants? (4) How might we make the IRB system more efficient? and (5) How might we help IRBs review new forms of research enabled by technological advances? The group discussions related to each of these ideas are presented below.

Textbox 1. "How might we…?" statements.

- Start a learning health system experiment?
- Share all of our data?
- Prevent those interested in profit from taking advantage of those interested in science?
- Conduct bold experiments? (Incentivize and facilitate)
- Expedite institutional review board (IRB) review? (More appropriately classify)
- Make consent actually informed?
- Increase transparency of IRB processes and outcomes?
- Set up an appropriate surveillance system to monitor ethical violations?
- Simplify IRB review?
- Abolish IRBs?
- Reframe the IRB as a research partner rather than a research barrier?
- Increase confidence in anonymization?
- Create a learning system where IRB evolves along with research practices?
- Engage the public in research and in helping IRBs?
- Assess the true cost of the IRB system? (Direct and indirect; What are we not doing [that we should be] because of the current IRB process?)
- Collect more empirical research on the current state of the IRB?
- Create a movement around IRB?
- Influence current legislation wisely? (Start at the state level to guide national policy; eg, California Embryonic Stem Cell Research Oversight [ESCRO] committee)
- Seek an IRB waiver process?
- Include topic experts in IRB decision-making processes?
- Advise as opposed to restrict?
- Move from permission to forgiveness?

Results

Redesigning the Consent Form and Process

The ethical principle of respect for persons implies that individuals should be informed about and voluntarily consent to participate in research. How do we ensure that consent is actually informed? How do we ensure that research participants from diverse backgrounds truly understand research study risks and opportunities? In regard to the first question, one idea may be to establish mechanisms through which participants can provide real-time feedback about their experiences to researchers. These mechanisms could serve to collect empirical data regarding the clarity of consent forms and potential participants' perceptions of risks and benefits. These data could inform and drive potential revisions to the consent form and other aspects of the research protocol. Relatedly, it is often the case that investigators write their consent forms to adhere to institutional templates, which may prompt the inclusion of content that is not relevant to or appropriate for a study. Thus, accurate and understandable descriptions of research should be encouraged in consent forms and processes, and inappropriate adherence to templates should be discouraged.

In addition, to make the informed consent process more accessible, one idea may be to think of the Creative Commons licenses [22] as a model. Similarly to the "three layers of licenses" used by Creative Commons, research studies could create three consent forms: one that contains all the legalese and scientific exposition; one in plain English that presents the facts; and a third that is simplified even further and presents risks in bullet point format. To make the process of obtaining consent culturally appropriate for underserved and underrepresented populations, community leaders, such as a *Promotor/a* in a Latino community, could be asked to help design the consent form and facilitate its use in ways that address community-specific concerns that researchers might not anticipate. Researchers could work with the community leader to help communicate these risks in a way that resonates with the community.

Empowering Researchers to Protect Participants

It may be worthwhile to consider how to construct a system of human research protections that fosters the ethical conduct of research without relying on an institution like the IRB. How might we start anew and reimagine and redesign research oversight without the traditional IRB in mind? What would an alternative system look like? One idea is to place responsibility for participant protection on the researcher rather than on the IRB. Researchers intending to engage in human-participant research could produce a document that lays out plans and risks of the research. They could then offer those documents, along with an outline of the proposed consent process, for review by their peers. Peers would be researchers in the field of relevance for the research. These documents could be posted on the Web in the same way clinical trials are registered; not to get approval but to create a public record of the research. Peers who review the documents might be accredited with some type of certification in human research protections, although an open question would be what entity would design and provide such certification (and how such an entity would look different from a traditional IRB). Obtaining this certification and participating in this process could be incentivized for researchers by considering these activities to be professional service required for career advancement and academic promotion. In this scenario, responsibility for ethical conduct during the study would be shared by both the researchers and the peers who agreed that the plan would adequately protect participants.

To make it easier to create high-quality plans, researchers could consult a Web-based resource similar to Stack Overflow [23], a resource that software developers often use to obtain quick answers from experts about specific technical issues. With this resource, the median response time is 11 minutes [24], and the people responding are rated, which provides information pertaining to their credibility and expertise. Using this Web-based resource, within a few hours, researchers posing questions such as “How do I ensure that I won’t cause harm by asking this interview or survey question?” would receive answers from researchers who have been rated in terms of experience and expertise in human research protections. Elements of the plans could ultimately become like “protection modules” that could be swapped in and out of consent forms and research protocols, drawing attention to highly ranked modules. We note, however, that such a solution would require an active community with a critical mass of users, which may or may not be realistic depending on whether the IRB process ever became truly standardized. Importantly, if such a system were found to be feasible, it is an approach that could be coupled with a system that punishes offenders (see below).

Reinforcement and Learning From Experience

This notion also begins with the premise that the burden to protect participants be shifted to the researcher rather than remain with an IRB or other regulatory body. How might we simultaneously reduce the bureaucratic burden of IRBs for researchers, particularly those conducting low-risk studies, and, at the same time, improve protection for research participants? In addition, how might we transform universities into learning ethics institutions that continuously improve their capacity to conduct ethical research [25]? One model for doing this could be the US Federal Aviation Administration’s Aviation Safety Reporting System [26]. Pilots who have a “bad” landing or make another safety-related error who self-report their mistake are spared from punishment, but those who do not report it themselves are penalized if someone elects to report [27]. Analogously, as an alternative to an IRB, in this system, researchers who create a protocol they believe to be safe, who then observe a harm during the research and who report that harm to their university or institution, present an opportunity for the research institution and community to learn how to prevent future harm. This expectation would be reinforced because, if the harm were to be reported by anyone else, including research staff or the research participant, the researcher would be sanctioned. That being said, there are clearly potential risks of supplanting the traditional IRB with a system entirely driven by researcher self-regulation. There could be conflicts between the researchers’ mandate to conduct studies and publish them and their mandate to protect participants, thus creating the opportunity for bias, the perception of bias, or, in extreme cases, maleficence. A system of researcher self-regulation would need to carefully consider and guard against these potential threats.

Increasing Efficiency of the Institutional Review Board

We suggest that, in order to improve the IRB process, it is essential to understand its costs, both direct and indirect. How might we collect and analyze empirical data on costs of the IRB system? Obvious tangible costs associated with the IRB system include salaries of personnel, IRB fees, space and infrastructure costs, and fees paid for training, education, and accreditation. In addition, for researchers, costs include the amount of time required for study staff to prepare and process a study protocol through the IRB. Depending on the institution and the type of protocol, IRB submissions can be extremely time intensive to prepare, which is an opportunity cost in terms of other ways in which that time could be spent. For research participants, costs include the time and cognitive effort needed to understand increasingly complex and bureaucratic consent forms. There are also less-tangible costs related to the broader public health caused by unnecessary delays to research imposed by IRBs.

One idea to increase efficiency may be to use the “Cooperative Research” process (see 45 CFR 46.114 [28]) to reduce the multiple IRB review of multisite studies and to use the “exempt” category to a greater degree, as it was intended. The exempt category is frequently appropriate for the vast majority of social and behavioral science studies, yet it is underused, which leads to delays in review and approval [29] and, thus, wasted resources. In addition, IRBs could take care to ensure that the process of review for exempt studies is reasonable and truly reflects their low-risk nature. Interestingly, exempt research, according to US federal regulations, does not need to be verified or reviewed by IRB staff. If institutions permitted, determining exempt status could be made the responsibility of the researcher. Overall, the idea that the bureaucracy of the IRB creates a significant burden to the research enterprise while producing unclear or intangible benefits to research participants is consistent with the purported rationale cited for the development of the proposed revisions to the Common Rule in the form of the NPRM [11] discussed above. We suggest that the IRB may benefit from an analysis of costs and benefits of its own activities, much like it does with the studies it oversees.

Review of Research That Leverages Technological Advances

New forms of research enabled by technological advances in information technology, data science, and other fields appear to be particularly challenging to IRBs. How might we develop resources that would facilitate appropriate review of 21st century science? The California Institute for Regenerative Medicine (CIRM) research oversight process could serve as a model. In CIRM 1.0, a committee separate from the IRB called the Embryonic Stem Cell Research Oversight (ESCRO) committee was formed to review stem cell research. Recognizing that few IRB members would have sufficient expertise to provide a meaningful review, the ESCRO committee, which is composed of scientists and a community representative, serves in an advisory capacity to the IRB. Such a model could be replicated for studies using emerging technologies about which IRBs may be similarly unfamiliar to ensure that experts are involved in the review.

For example, mobile, visual imaging, pervasive sensing, and geolocation tracking technologies present new ethical and regulatory challenges [20]. For instance, visual imaging using wearable sensors have made it possible for researchers to measure physical activity, diet, travel, and the settings in which these behaviors occur using a first-person point-of-view wearable camera. Given the increasing interest in these methods for studying behavior "in the wild," we anticipate increased research using visual methods, which raises privacy concerns and issues related to the rights of bystanders. Likewise, with wearable sensors, mobile phone transmission, and analytics in the cloud, health information can be captured continuously in real time. Location tracking technologies provide spatial data and the opportunity for assessing the context in which behavior is occurring, as well as identifying underlying spatial relationships such as clustering or transmission pathways. These data are fine grained and specific down to the exact longitude and latitude at a given point in time. Standards for how these data are transmitted, stored, and shared are necessary, since the introduction of the US Health Insurance Portability and Accountability Act, in most cases (at least at present), does not apply. A virtual network composed of researchers, technologists, and bioinformatics experts may prove to be a workable solution to augment or replace the traditional IRB review process resulting in an informed and meaningful human protections review of 21st century science.

Discussion

In this paper we imagine, and offer some ideas for the design of, a progressive, responsive, and nimble human research protections system. By encouraging broad and innovative ideas, the design thinking method not only opens up new avenues for exploration, but also provides clarity about some of the shortcomings of our current IRB system. The workshop described here aimed to stimulate creative thinking about how the existing IRB structure could be improved, while also generating ideas for entirely new ways of protecting research participants. Clearly, some of the ideas presented here are more feasible than others. For example, it may be more realistic to encourage IRBs to exploit the current regulations and use the "exempt" category more appropriately and frequently. Alternatively, imagining a completely different review process that would replace the traditional IRB entirely may be less acceptable and would likely create new problems. If the proposed NPRM is adopted, we note that some research may be excluded from the traditional IRB review yet may benefit from an "ethics consultation" process to avoid making mistakes that an IRB may have detected. It would be valuable to estimate the cost savings realized by implementing the new practices that reduce the burden on both IRBs and researchers without compromising human research protections. Clearly, further debate by stakeholders is necessary to develop these and other ideas into concrete recommendations to advance applied human research ethics.

Sometimes, systems are so entrenched in ways of doing things that change from within is not possible and disruptive external approaches are required (eg, Uber as an alternative to transportation via taxis, or specialty charter schools as an alternative to traditional public schools). The IRB may be a system in need of disruption. Using the design thinking method fostered the development of "outside the box" ideas that may improve research participant protections and the IRB structure. As such, we initiated the exercise described in this paper and seek to share the process and results with the greater research community. The IRB system will need to be updated or possibly even reinvented in order to be responsive to technological advances of recent decades that have enabled new forms of research. These advances have created challenges to our current system that the NPRM will not likely solve. Pilot research programs that test-drive the reform ideas presented here, or perhaps other ideas, would be worthwhile and informative as the research community considers how to intervene and make healthy what many believe is an ailing human research protections system.

Acknowledgments

We thank the Robert Wood Johnson Foundation for providing support for the workshop that was the source of the ideas presented. We also thank The Qualcomm Institute, the University of California, San Diego Division of the California Institute for Telecommunications and Information Technology (Calit2), for providing the space in which to hold the meeting. We also thank Ms Katrina Ansert for taking detailed notes during the workshop, Mr Johnny Nguyen for administrative support, and Ms Cynthia Cheung for assistance with manuscript preparation.

This work was supported by two grants from the Robert Wood Johnson Foundation ("The Health Data Exploration [HDE] project," principal investigator: KP, #71693, 2013-2017, and "Connected and Open Research Ethics [CORE]," principal investigator: CN, #72876, 2015-2017) and a grant from the National Human Genome Research Institute ("Impact of Privacy Environments for Personal Health Data on Patients", principal investigator: CB, R01 HG008753, 2015-2018).

Conflicts of Interest

None declared.

References

1. Zucker B. The Belmont Report. In: Wiley Encyclopedia of Clinical Trials. New York, NY: John Wiley & Sons; Dec 14, 2007.
2. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med 2016 Dec;6(4):577-586. [doi: 10.1007/s13142-016-0426-4] [Medline: 27688250]
3. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. JMIR Mhealth Uhealth 2015;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
4. Eysenbach G. Towards ethical guidelines for e-health: JMIR theme issue on eHealth ethics. J Med Internet Res 2000;2(1):E7 [FREE Full text] [doi: 10.2196/jmir.2.1.e7] [Medline: 11720926]
5. Conway M. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. J Med Internet Res 2014;16(12):e290 [FREE Full text] [doi: 10.2196/jmir.3617] [Medline: 25533619]
6. Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. Nature 2012 Sep 13;489(7415):295-298 [FREE Full text] [doi: 10.1038/nature11421] [Medline: 22972300]
7. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for schizophrenia: patient engagement with a mobile phone intervention following hospital discharge. JMIR Ment Health 2016;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
8. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012 Nov 1;491(7422):56-65 [FREE Full text] [doi: 10.1038/nature11632] [Medline: 23128226]
9. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May 12;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
10. de Montjoye MY, Radaelli L, Singh VK, Pentland AS. Identity and privacy. Unique in the shopping mall: on the reidentifiability of credit card metadata. Science 2015 Jan 30;347(6221):536-539. [doi: 10.1126/science.1256297] [Medline: 25635097]
11. [No authors listed]. Human subjects research protections: enhancing protections for research subjects and reducing burden, delay, and ambiguity for investigators. Federal Register 2011 Jul 26;76(143):44512-44531 [FREE Full text]
12. [no authors listed]. Notice of proposed rulemaking: federal policy for the protection of human subjects. Federal Register 2015.
13. Klitzman R. How local IRBs view central IRBs in the US. BMC Med Ethics Jun 23 2011:12. [doi: 10.1186/1472-6939-12-13] [Medline: 21699725]
14. National Institutes of Health. Final NIH Policy on the use of a single institutional review board for multi-site research. 2016. URL: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-094.html) [accessed 2016-10-31] [WebCite Cache ID 6ltef8okD]
15. Emanuel E. Reform of clinical research regulations, finally. N Engl J Med 2015 Oct 10;373(24):2296-2299. [doi: 10.1056/NEJMp1512463] [Medline: 26536071]
16. Committee on Federal Research Regulations and Reporting Requirements: A New Framework for Research Universities in the 21st Century, Committee on Science, Technology, and Law, Board on Higher Education and Workforce, Policy and Global Affairs, National Academies of Sciences, Engineering, and Medicine. Optimizing the Nation’s Investment in Academic Research: A New Regulatory Framework for the 21st Century. Washington, DC: National Academies Press; Jul 2016.
17. Schneider C. The Censor's Hand. Cambridge, MA: MIT Press; 2015.
18. Lidz C, Garverich S. What the ANPRM missed: additional needs for IRB reform. J Law Med Ethics Summer 2013;41(2):390-396. [doi: 10.1111/jlme.12050] [Medline: 23802892]
19. Bloss C. Does family always matter? Public genomes and their effect on relatives. Genome Med 2013;5(12):107. [doi: 10.1186/gm511] [Medline: 24342550]
20. Nebeker C, Linares-Orozco R, Crist K. A multi-case study of research using mobile imaging, sensing and tracking technologies to objectively measure behaviorthical issues and insights to guide responsible research practice. J Res Admin 2015;46(1):118-137.
21. Brown T. Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation. New York, NY: HarperCollins; 2009.
22. Creative Commons. About the licenses: what our licenses do. Mountain View, CA: Creative Commons URL: https://creativecommons.org/licenses/ [accessed 2016-10-19] [WebCite Cache ID 6lNd6mMfm]
23. Stack Overflow. New York, NY: Stack Exchange Inc; 2016. URL: http://stackoverflow.com/ [accessed 2016-12-13] [WebCite Cache ID 6mjRnc3eP]

24. Mamykina L, Manoim B, Mittal M, Hripcsak G, Hartmann B. Design lessons from the fastest q&a site in the west. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC, Canada p. 2857-2866. [doi: 10.1145/1978942.1979366]
25. Senge P. The Fifth Discipline: The Art and Practice of the Learning Organization. New York, NY: Broadway Business; 2006.
26. NASA. Aviation Safety Reporting System. URL: https://asrs.arc.nasa.gov/ [accessed 2016-10-19] [WebCite Cache ID 6lNcRUWUZ]
27. Norman D. The Design of Everyday Things. New York, NY: Doubleday; 1988.
28. Office for Human Research Protections. Code of Federal Regulations, Title 45: Public Welfare, Department of Health and Human Services. Part 46: protection of human subjects. Rockville, MD: Office for Human Research Protections; 2010 Jan 15. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html# [accessed 2016-12-13] [WebCite Cache ID 6mjSREPpp]
29. Schrag Z. Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965-2009. Baltimore, MD: Johns Hopkins University Press; 2010.

Abbreviations

ANPRM: Advance Notice of Proposed Rulemaking
CIRM: California Institute for Regenerative Medicine
ESCRO: Embryonic Stem Cell Research Oversight
IRB: institutional review board
NPRM: Notice of Proposed Rulemaking

Edited by G Eysenbach; submitted 08.09.16; peer-reviewed by J Hawkins, D Maslove; comments to author 19.10.16; revised version received 09.11.16; accepted 23.11.16; published 22.12.16

Please cite as:
Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, Fowler J, Waldo A, Weibel N, Patrick K, Klemmer S, Melichar L
Reimagining Human Research Protections for 21st Century Science
J Med Internet Res 2016;18(12):e329
URL: http://www.jmir.org/2016/12/e329/
doi:10.2196/jmir.6634
PMID:28007687

Review

Elements of Trust in Digital Health Systems: Scoping Review

Afua Adjekum, MSc; Alessandro Blasimme, MA, PhD; Effy Vayena, MSc, PhD

Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

Corresponding Author:
Effy Vayena, MSc, PhD
Department of Health Sciences and Technology
ETH Zurich
Auf der Mauer 17
Zurich,
Switzerland
Phone: 41 44 632 8301
Email: effy.vayena@hest.ethz.ch

Abstract

Background: Information and communication technologies have long become prominent components of health systems. Rapid advances in digital technologies and data science over the last few years are predicted to have a vast impact on health care services, configuring a paradigm shift into what is now commonly referred to as digital health. Forecasted to curb rising health costs as well as to improve health system efficiency and safety, digital health success heavily relies on trust from professional end users, administrators, and patients. Yet, what counts as the building blocks of trust in digital health systems has so far remained underexplored.

Objective: The objective of this study was to analyze what relevant stakeholders consider as enablers and impediments of trust in digital health.

Methods: We performed a scoping review to map out trust in digital health. To identify relevant digital health studies, we searched 5 electronic databases. Using keywords and Medical Subject Headings, we targeted all relevant studies and set no boundaries for publication year to allow a broad range of studies to be identified. The studies were screened by 2 reviewers after which a predefined data extraction strategy was employed and relevant themes documented.

Results: Overall, 278 qualitative, quantitative, mixed-methods, and intervention studies in English, published between 1998 and 2017 and conducted in 40 countries were included in this review. Patients and health care professionals were the two most prominent stakeholders of trust in digital health; a third—health administrators—was substantially less prominent. Our analysis identified cross-cutting personal, institutional, and technological elements of trust that broadly cluster into 16 enablers (altruism, fair data access, ease of use, self-efficacy, sociodemographic factors, recommendation by other users, usefulness, customizable design features, interoperability, privacy, initial face-to-face contact, guidelines for standardized use, stakeholder engagement, improved communication, decreased workloads, and service provider reputation) and 10 impediments (excessive costs, limited accessibility, sociodemographic factors, fear of data exploitation, insufficient training, defective technology, poor information quality, inadequate publicity, time-consuming, and service provider reputation) to trust in digital health.

Conclusions: Trust in digital health technologies and services depends on the interplay of a complex set of enablers and impediments. This study is a contribution to ongoing efforts to understand what determines trust in digital health according to different stakeholders. Therefore, it offers valuable points of reference for the implementation of innovative digital health services. Building on insights from this study, actionable metrics can be developed to assess the trustworthiness of digital technologies in health care.

(J Med Internet Res 2018;20(12):e11254) doi: 10.2196/11254

KEYWORDS

digital health; digital health technologies; health care; health systems; trust

Introduction

Background

Digital health broadly refers to the use of information and communication technologies to improve human health, health care services, and wellness for both individuals and populations [1,2]. It has been argued that the capacity to collect, store, and analyze extensive amounts of health data is the chief driving force of digital health [3]. The accessibility of such data is rejuvenating the process involved in diagnosing, managing, and treating disease, thus exceeding the conventional boundaries of how health care institutions and providers operate. A case in point is the myriad number of smartphone apps that allow patients to seamlessly monitor various aspects of their health care beyond the confines of a health care institution [1].

There is currently no consensus on a definition for digital health. The term "digital medicine" for instance, resembles digital health, as it also refers to the use of digital technologies such as biosensors and smartphones to refine and individualize medicine [4]. Given how they are often described, electronic health, mobile health (mHealth), telecare, and telehealth could also be used interchangeably with digital health [5]. This ambiguity calls for a need to generate an inclusive definition that captures the different terms that may be used to portray digital health.

The US Food and Drug Administration (FDA) depicts digital health as comprising of mHealth, wearable devices, telehealth, telemedicine, personalized medicine, electronic health records (EHRs), and health information technology (IT) [6]. In this review, we adopt this as our working definition of digital health. Throughout this paper, the term "digital health" refers to all of the aforementioned categories. So far, there has been a prolific development of digital health technologies, and the value of such ventures continues to rise at a steady pace. In 2017 alone, the global net worth of the digital health industry was estimated at US $25 billion (£19 billion; €21 billion). Some estimates even project that digital health could cut back up to US $7 billion of US health care expenditure annually [7].

Beyond economic gains, improved safety and efficacy are among the anticipated benefits of digital health [7-10]. Current evidence supports the notion that digital health does indeed bolster safety within health systems [11]. In the domain of health care delivery, digital health promises to abate mortality, shorten hospital admissions, and decrease medication errors [11]. Despite these advances, there are privacy and data protection concerns associated with the pace of development of digital health products [7,12]. Moreover, as data from digital health tools such as mHealth apps increasingly inform medical decision making, the issue of medical liability comes to the fore [13,14]. The considerations about privacy and data protection highlight the ethical challenges that bear directly on the trustworthiness of digital health. While numerous studies have analyzed such ethical issues [15-19], the determinants of trust in digital health are yet to receive comparable levels of attention [1,3,20-22].

What is Trust?

Trust is an elusive concept that is difficult to pin down in operational terms. Relationships of trust can exist between individuals, between individuals and the organizations they come into contact with, or between 2 organizations of any given nature [23]. Trust is oftentimes illustrated as a relationship between one party (a trustor) and another (a trustee) with optimistic anticipation that the trustee will fulfill the trustor's expectations [23,24]. Trust relationships often lack enforceable obligations and are thus vulnerable to deception [25]. Consequently, different sets of reasons encourage trust relationships. Chief among them are the trustee's reliability (possessing a good reputation), competence (having the technical skills to perform the task at hand), and integrity (generally acting in an honest way) [26].

Within health systems, trust is a prominent component of doctor-patient relationships [27-29]. It improves not only health care access but also treatment outcomes and patient satisfaction [30,31]. However, whether or not it is appropriate to talk about trust between people and inanimate objects—such as technological products—remains an open question in the literature [21,32]. Indeed, the inclination of individuals to purchase or use products that are derived from "expert systems"—those structures that rely on either technical know—how or professional expertise and whose outcomes are consequently pervasive, opaque, or easily taken for granted—has been described as a tangible component of trust [33].

Some experts suggest that trust is propelled by contingency rather than risk, while others maintain that the ability to weigh risks and to choose between different actions drives trust [34]. Despite the risk of deception within any trust relationship, it is disputable whether one chooses to trust solely by weighing risks or actively by evaluating alternative options. Be that as it may, in the case of medical technologies, institutional trust and technical reliability are deeply intertwined [35]. In terms of digital health technologies, we hypothesize that trust is likely to develop if the risks and uncertainties associated with their use can be minimized.

As health care becomes increasingly dependent on digital technologies, exploring what determines and what foregoes trust in digital health is of paramount importance. Identifying the factors pertinent to trust can inform the development of novel health care services as well as meet the needs and expectations of users and patients. In addition, such factors can be taken into account for the assessment of both new and existing digital health services. Thus, this study seeks to contribute to this discourse by analyzing what the relevant stakeholders in digital health consider as the enablers and impediments of trust in digital health.

Methods

Overview

This review aimed to summarize the enabling and impeding factors of trust in digital health. To this end, we conducted a scoping review using Arksey and O'Malley's proposed framework on scoping reviews [36]. A scoping review methodology was chosen, as it appropriately captures broad and ambiguous topics, like digital health, that may involve a myriad of study designs. We searched for studies that reported on the perspectives of different digital health stakeholders. From these perspectives, we discerned views on what was reported to facilitate trust and what hindered it. Often, some of these same factors were recognized as relevant for the acceptance of a particular technology. By acceptance, we mean adoption and use grounded in or at least co-occurring with trust on the part of users. This understanding of trust as a potential determinant of acceptance reflects some credited models of technology acceptance in the health care sector [37].

Information Sources

We searched 5 databases: MEDLINE, EMBASE, the Cumulative Index to Nursing and Allied Health Literature, PsycINFO, and Web of Science for peer-reviewed studies as well as gray literature. We worked with a research librarian at the University of Zurich, Switzerland, to identify relevant bibliographic databases and to construct a search strategy that would ensure comprehensive results.

Search Strategy

The search strategy involved formulating keywords and Medical Subject Headings around the 2 main themes of this study, namely, trust and digital health. Since the concept of trust can be ill-defined within the literature [35], we set out to include synonyms such as expectation, mistrust, confidence, and experience to capture the heterogeneity of trust descriptions within the literature (Multimedia Appendix 1). Digital health, on the other hand, was disaggregated into its distinctive components as described by the FDA: mHealth, wearable devices, telehealth, telemedicine, personalized medicine, and health IT. The searches were restricted to publications available in English, French, German, Italian, and Spanish with no publication date restrictions, to allow the search results to encompass a broad range of relevant studies. The searches commenced on July 20, 2017, and concluded on August 18, 2017. The recovered studies were then exported into the Endnote X8.2 reference software.

Eligibility Criteria of Included Studies

To capture the wide array of studies that may be relevant to this topic, we did not predefine the study designs of included studies. This allowed for the inclusion of qualitative, quantitative, intervention, and mixed-methods studies. We assessed the relevance of the retrieved studies to ensure that they related to either of the abovementioned digital health technologies. Moreover, each study was required to meet at least 1 of the following criteria: (1) investigate stakeholder perceptions, attitudes, expectations, and perspectives toward digital health or (2) highlight some potential enablers and impediments to trust in digital health technologies and services.

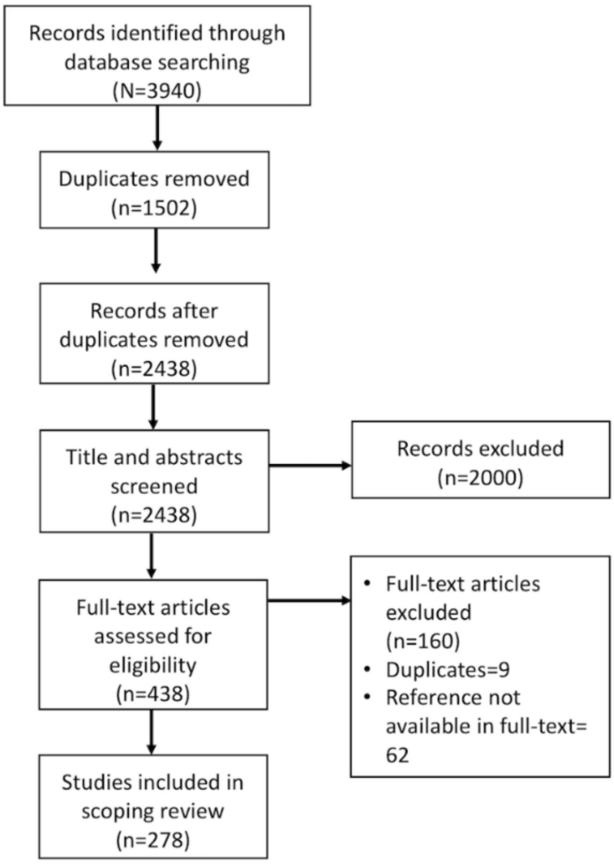
Study Selection, Categorization, and Data Extraction

As is customary in scoping reviews, we employed an iterative approach to select, categorize, and extract data from the recovered studies [36]. We used a 2-step process to select relevant articles. At first, 1 author (AA) reviewed all of the titles and abstracts derived from the search. In order to reduce sampling bias [38], a second author (AB) reviewed a random sample of 243 titles along with their associated abstracts (constituting 10% of the total sample after duplicates had been removed). To assess the level of agreement between the 2 reviewers, an interrater reliability score using Cohen kappa was computed along with its corresponding CI and *P* value. The Cohen kappa score for the 2 coders (AA and AB) was .661 (95% CI 0.465-0.857; $P<.001$). According to McHugh (2012), a kappa of.661 signifies a moderate agreement between the coders [39].

Overall, we retrieved a total of 3940 search results from the 5 databases. Of these, 1474 were identified as duplicates and discarded. However, during the screening process, we discovered an extra 28 duplicates, increasing the total number discarded to 1502. This led to screening the titles and abstracts of 2438 articles of which 438 were eligible for full-text screening. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram below (Figure 1) lays out these procedures in more detail [40]. The final number of articles included in the review was 278.

From each article, we documented the author's name, year of publication, country of origin, sample size, study design (eg, qualitative or quantitative), digital health type as well as the relevant stakeholders. A descriptive, analytical approach was used to summarize the outcomes of the studies. We identified the trust elements (enablers and impediments) by charting the key themes and issues identified from each study [36]. To develop these themes, the results section of each study was scrutinized to identify various stakeholder priorities, perspectives, expectations, perceptions, and attitudes toward a particular digital health technology or service. Multimedia Appendix 2 shows the studies from which each element was derived. Since either an enabler or impediment could be derived from the same study, we reported the overall number of studies that support each element rather than percentages. Simultaneously, we compiled a list of recurring terminologies that were used to represent or describe the various digital health technologies, which we termed "health technology types."

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.



Results

Characteristics of Articles

Of the 278 articles included in this review, 51 (51/278, 18.3%) related to telemedicine and telehealth, 24 (24/278, 8.6%) to personalized medicine, 47 (47/278, 16.9%) to mHealth, 73 (73/278, 26.3%) to health IT, 73 (73/278, 26.3%) to EHRs, and 4 (4/278, 1.4%) to wearable devices, while 6 (6/278, 2.2%) concerned 2 or more digital health technologies. Most of the studies were conducted in 2015 (50/278, 18.0%), and the median year was 2014. The oldest study was conducted in 1998, while the most recent study was from 2017. There were 98 qualitative studies, 133 quantitative studies, 45 mixed method studies, and 2 intervention studies. Data from Web-based sources were collected in 7 studies. Overall, the studies were conducted in 40 countries; the United States was the most represented (101/278, 36.3%). The United Kingdom had the second highest number of studies (47/278, 16.9%) followed by Australia (16/278, 5.8%) and Canada (15/278, 5.4%; see Multimedia Appendix 3).

Digital Health Technologies and Services

For each digital health technology, we uncovered several health technology types employed to provide digital health services. Within each digital health category, there appear to be multiple terminologies to describe identical or variable technologies or services. In many instances, there were only slight variations differentiating one service from the other. For example, electronic patient records, electronic medical records, and electronic health care records were variable forms of EHRs, while Web-based consultations, online support groups, and Web-based health information were some examples of health IT. Multimedia Appendix 4 provides a list of the variable terminologies identified from the included studies.

Stakeholders

In our analysis, we identified 2 major stakeholders: *patients* or the *public* (187 studies) and *health care professionals* (HCPs; 101 studies). A third less predominant group—*health administrators* (HAs; 20 studies)—was also identified. For the sake of clarity, HCPs refer to a broad range of health care specializations that include pharmacists, occupational therapists, physical therapists, physicians, and nurses. Other stakeholders that were considerably less represented in the analyzed studies included medical and nursing students, consumer groups, health policy makers, data controllers, academic researchers, social workers, counselors, and IT technicians.

Trust Enablers and Impediments

Our findings indicate that trust in digital health technologies and services is affected by a variety of elements. In this study, trust enablers refer to those factors that encourage stakeholders' trust in digital health, while trust impediments denote the factors that can potentially hinder trust. These trust enablers and impediments, therefore, underscore the elements that influence stakeholder decisions on whether or not to place their trust in digital health technologies.

Personal Elements

By personal elements, we designate factors that influence trust in digital health at the individual level. The higher the likelihood of a digital health technology or service to enhance job performance, the more likely stakeholders are to trust it due to convenience and *usefulness* (110 studies). Moreover, *sociodemographic factors* (84 studies) such as ethnicity, income, and educational status affected an individual's trust in digital health either positively or negatively, thereby acting simultaneously as enablers and impediments. *Ease of use* (53 studies)—the propensity for systems to require minimal effort for use—also influenced trust positively. Other personal elements include *fair data access* (21 studies), *recommendations* (17 studies) from family members, acquaintances and colleagues as well as *self-efficacy* (15 studies). The latter denotes a refined acumen to manage one's own health [41]. *Altruism* (9 studies) also contributed to stakeholder involvement in digital health enterprises and was driven by the prospect of contributing to novel and beneficial therapies that would benefit society.

A number of studies reported *excessive costs* (34 studies) and *limited accessibility* (55 studies) as potential barriers to trust and, therefore, acceptance. *Fear of data exploitation* (25 studies) from third parties such as insurance and pharmaceutical companies was another palpable impediment to trusting digital health systems.

Technological Elements

The technological elements refer to the technical components of digital health technologies that make them appealing to accept and use. In terms of sensitive personal data such as genetic data, robust systems that delivered on safety and *privacy* (73 studies) were crucial to trust. There was a high affinity for *customizable design features* (28 studies) that allowed stakeholders to tailor devices to their specific needs. Since HCPs were often required to utilize disparate software programs, they requested *interoperable* (10 studies) systems that ensured that newer systems are compatible with currently existing ones. Relating to trust impediments, *defective technology* (32 studies) was a culprit for the minimal use of digital health technologies or services.

Institutional Elements

The institutional elements denote the strategies that are implemented within establishments that influence stakeholder trust in digital health. Several studies highlighted that various stakeholders had suggestions, expectations, or feedback to provide on how best to improve digital health services. Consequently, *stakeholder engagement* (71 studies), which involves taking stakeholders' opinions into account, emerged as a relevant condition to increase trust in digital health. *Improved communication* (46 studies) was a cross-cutting expectation from digital health technologies. Both patients and HCPs valued the many communication avenues that digital health provided. In 40 studies, it appeared that there was a need for *initial face-to-face* interactions prior to the introduction of digital health services. Generally, stakeholders expected digital health technologies to build upon and improve on existing systems. Hence, they preferred technologies that *decreased workloads* (82 studies).

The *reputation of service providers* (71 studies), however, served as either an enabler or impediment to trust in digital health. A good reputation encouraged trust and vice versa. *Time-consuming* (42 studies) technologies as well as those that provided *information of poor quality* (51 studies) impeded trust. Other impediments identified included *insufficient training* (54 studies) and uncertainties originating from *inadequate publicity* (44 studies) about the capabilities, existence, and risks involved in using digital health. Finally, trust was also hindered by the absence of *guidelines for standardized use* (22 studies).

In Table 1, we provide a summary of these findings and highlight the stakeholders for whom these elements appeared pertinent. In the table, found in parenthesis next to each element are the total number of studies (n). A checkmark is also used to illustrate the respective trust elements that each stakeholder is associated with.

Table 1. Trust enablers and impediments alongside their corresponding stakeholders.

Element classification	Enablers of trust	Impediments to trust	Stakeholders		
			Patients	HCPs[a]	HAs[b]
Personal elements	Altruism (n=9)	N/A[c]	✓[d]	N/A	N/A
	Ease of use (n=30)	N/A	✓	✓	✓
	N/A	Excessive costs (n=34)	✓	✓	✓
	Fair data access (n=21)	N/A	✓	✓	N/A
	N/A	Fear of data exploitation (n=25)	✓	N/A	N/A
	Recommendation by others (n=17)	N/A	✓	✓	N/A
	Self-efficacy (n=15)	N/A	✓	✓	N/A
	N/A	Limited accessibility (n=55)	✓	✓	N/A
	Sociodemographic factors (n=84)[e]	Sociodemographic factors (n=84)[e]	✓	✓	N/A
	Usefulness (n=110)	N/A	✓	✓	N/A
Technological elements	Customizable design features (n=28)	N/A	✓	✓	N/A
	N/A	Defective technology (n=32)	✓	✓	✓
	Interoperability (n=10)	N/A	N/A	✓	N/A
	Privacy (n=73)	N/A	✓	✓	N/A
Institutional elements	Decreased workloads (n=83)	N/A	N/A	✓	✓
	Guidelines for standardized use (n 22)	N/A	N/A	✓	✓
	Improved communication (n=46)	N/A	✓	✓	✓
	N/A	Inadequate publicity (n=44)	✓	✓	✓
	Initial face-to-face contact (n=40)	N/A	✓	✓	N/A
	N/A	Insufficient training (n=54)	✓	✓	✓
	N/A	Poor information quality (n=51)	✓	✓	✓
	Service provider reputation (n=71)[e]	Service provider reputation (n=71)[e]	✓	✓	N/A
	Stakeholder engagement (n=71)	N/A	✓	✓	N/A
	N/A	Time-consuming (n=42)	N/A	✓	✓

[a] HCP: health care professional.

[b] HA: health administrator.

[c] N/A: not applicable.

[d] Check mark indicates respective trust elements that each stakeholder is associated with.

[e] These elements (sociodemographic factors and service provider reputation) are simultaneously trust enablers and impediments.

Discussion

Principal Findings

This study highlights the enablers of and impediments to trust in digital health technologies and services. Our results show that digital health encompasses a wide variety of health technology types and their respective services. Altogether, we identified 3 primary stakeholders: *patients*, *HCPs*, and *HAs*. Moreover, our findings map out cross-cutting *personal*, *technological*, and *institutional* trust elements in the form of enablers and impediments to trust in digital health technologies. Of these elements, sociodemographic factors and service provider reputation acted simultaneously as enablers and impediments.

A possible interpretation of the ambivalent nature of sociodemographic factors may lie in the fact that a lack of resources, be them material or educational, render people in a vulnerable state. Within health care settings, individuals often compensate for their vulnerability by perceiving health workers as potential threats [42]. The level of risk involved in instances of unfulfilled or broken trust impacts the willingness of vulnerable people to entrust individuals, institutions, or technologies with various tasks. In a similar fashion, those sitting at the high end of the socioeconomic spectrum may be prone to trust new technologies because of their perceived ability to control them. Alternatively, they may have higher expectations with regards to health care services and, thus, set the bar of trustworthiness much higher than the more disadvantaged strata of the population.

The ambiguity that we uncovered in this study reflects what other studies on trust vis-à-vis sociodemographic status have highlighted. Available evidence on the role of sociodemographic factors (eg, ethnicity, gender, and educational status) within the health care context is mixed. For instance, 1 study, has shown that patient characteristics (with the exception of age) rarely predict trust in patient-doctor relationships [43]. Conversely, others have identified patient characteristics such as age, ethnicity, income status, educational level, and literacy levels as crucial factors affecting the use of electronic health [20,44]. In light of these discrepant findings, further research is needed to clarify the underlying effects of sociodemographic factors in digital health.

A prevalent theme throughout this review was that stakeholders appear to trust profit-making entities such as insurance and pharmaceutical companies much less than they do public institutions like universities. This is a widespread phenomenon that reflects greater public assumptions about the private sector's interests and profits [45]. Our findings support the importance of reputation to trust even though *service provider reputation* was identified as both a trust impediment and enabler. On the one hand, when a service provider embodies high ethical standards and is proficient at providing required services, they attain the advantage of shaping the expectations of stakeholders positively. In contrast, negative performance statistics of a service provider stand to give rise to negative expectations about their proficiency.

Despite stakeholder optimism about digital health tools, there are notable concerns about the accuracy of digital information exacerbated by the absence of uniform quality controls and standards [23]. Onora O'Neill has underscored the importance of enacting policies that address these challenges [26]. Based on the studies concerning Web-based health information included in this review, it was observed that patients and HCPs struggled to establish the quality of digital information. Consequently, in order to gauge the authenticity, veracity, and usefulness of digital health technologies or services, they relied quite significantly on *recommendations* from family members, colleagues, or acquaintances.

The FDA definition that we adopted for this review features personalized medicine as one of the components of digital health. Domains such as personalized medicine rely on the creation of large cohorts of deeply characterized individuals, as is the case with the 1 million participant research cohort being built for the Precision Medicine Initiative in the United States [3,46,47]. Success in this area will crucially depend on trust [48,49]. How to gain the degree of public support and personal commitment that is needed to build such infrastructures is far from obvious. In such cases, the ability to measure trustworthiness against a validated set of criteria will greatly increase the odds of success for such initiatives. Our study can be considered as a vital step in this direction, laying the conceptual groundwork for the development of such tools.

As we have shown, trust in digital health technologies and services depends on the interplay of a complex set of enablers and impediments. This study sheds light on what determines trust in digital health according to different stakeholders. More specifically, our findings can be of help in the implementation of innovative digital health technologies and services as well as in the management of existing digital health infrastructures. Building on insights from this study, actionable metrics such as the patient trust in telemedicine services tool can be developed to assess the trustworthiness of digital technologies in health care [50]. Each metric would need to undergo a validation process before being deployed in practice by HAs charged with monitoring or developing digital health services.

Overall, engaging with efforts to investigate the different dimensions of trust is particularly urgent given the growing attention from entities such as governments. This heightened level of attention is warranted due to the potential impacts of ever more innovative forms of digital health. Some approaches to digital health, in particular, those relying on big data, predictive analytics, and artificial intelligence [51-53] will require dedicated governance models in order to deliver on their promises while meeting the expectations of their users [54]. Reliable ways of measuring trustworthiness will, thus, be a key tool in such a rapidly evolving scenario.

Limitations

A drawback to this study is the unequal number of studies in each digital health category. Although this was unlikely to have skewed our findings, there were relatively fewer studies on the newer forms of digital health such as wearable devices. Despite suggestions for reviews to be screened by 2 individuals, the volume and the complicated 2-step process involved in gleaning relevant information meant that only 1 author (AA) could fully screen all of the publications. Nevertheless, a second author (AB) screened 10% of the total publications for which a kappa statistic was calculated to ensure a minimal level of bias. Even though there was a moderate interrater agreement score (kappa=.661; 95% CI 0.465-0.857; P<.001), our kappa statistic is well above the .60 value that represents an inadequate agreement threshold [39]. Lastly, we acknowledge that scoping reviews can have several shortcomings [55]. However, the poorly-defined nature of both digital health and trust within the literature required a method that could map out the discourse and, thus, pave the way for a systematic review.

Conclusion

Rapid advances in digital technologies and data science over the last few years are predicted to have a tangible impact on health care services, configuring a paradigm shift into what is now commonly referred to as digital health. Digital health, however, relies heavily on trust to succeed. What counts as the building blocks of trust in digital health systems has so far remained underexplored. In this study via a scoping review approach, we seek to fill this gap by analyzing what relevant stakeholders consider as the constitutive elements of trust in digital health. Overall, 278 qualitative, quantitative, mixed-methods, and intervention studies in English were included in this review. *Patients* and *HCPs* were the 2 most prominent stakeholders to trust, while *HAs* were a third and substantially less prominent stakeholder. Altogether, the trust elements that either enabled or hindered trust in digital health clustered into *personal*, *technological*, and *institutional* factors. This study paves the way for the implementation of the criteria necessary to measure and anticipate trust in emerging health care technologies.

Acknowledgments

This project was funded by grants from the Swiss National Science Foundation to EV (PP000P3-157556; 4075210-167223). The authors wish to thank Martina Gosteli for her help with the search strategy and Felix Gille for valuable comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search queries.

[PDF File (Adobe PDF File), 27KB-Multimedia Appendix 1]

Multimedia Appendix 2

Studies illustrating trust enablers and impediments.

[PDF File (Adobe PDF File), 135KB-Multimedia Appendix 2]

Multimedia Appendix 3

List of study countries.

[PDF File (Adobe PDF File), 24KB-Multimedia Appendix 3]

Multimedia Appendix 4

Health technology types.

[PDF File (Adobe PDF File), 33KB-Multimedia Appendix 4]

References

1. Kostkova P. Grand challenges in digital health. Front Public Health 2015;3:134 [FREE Full text] [doi: 10.3389/fpubh.2015.00134] [Medline: 26000272]
2. World Health Organization. 2018. eHealth URL:http://www.who.int/ehealth/en/ [accessed 2018-06-07] [WebCite Cache ID 6zzwR7Y69]
3. Vayena E, Haeusermann T, Adjekum A, Blasimme A. Digital health: meeting the ethical and policy challenges. Swiss Med Wkly 2018 Dec 29;148:w14571 [FREE Full text] [doi: 10.4414/smw.2018.14571] [Medline: 29376547]
4. Steinhubl SR, Topol EJ. Digital medicine, on its way to being just plain medicine. npj Digital Med 2018 Jan 15;1(1):1-1. [doi: 10.1038/s41746-017-0005-1]
5. Boogerd EA, Arts T, Engelen LJ, van de Belt TH. “What Is eHealth”: Time for An Update? JMIR Res Protoc 2015 Mar 12;4(1):e29 [FREE Full text] [doi: 10.2196/resprot.4065] [Medline: 25768939]
6. United States Food and Drug Administration. 2018. Digital Health URL:https://www.fda.gov/medicaldevices/digitalhealth/ [accessed 2018-06-07] [WebCite Cache ID 6zzwKIrhs]
7. Duggal R, Brindle I, Bagenal J. Digital healthcare: regulating the revolution. BMJ 2018 Dec 15;360:k6. [Medline: 29335296]
8. Krenn L, Schlossman D. Have Electronic Health Records Improved the Quality of Patient Care? PM R 2017 May;9(5S):S41-S50. [doi: 10.1016/j.pmrj.2017.04.001] [Medline: 28527503]
9. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood) 2005;24(5):1103-1117 [FREE Full text] [doi: 10.1377/hlthaff.24.5.1103] [Medline: 16162551]
10. McKenna RM, Dwyer D, Rizzo JA. Is HIT a hit? The impact of health information technology on inpatient hospital outcomes. Applied Economics 2017 Dec 22;50(27):3016-3028. [doi: 10.1080/00036846.2017.1414934]
11. Agboola SO, Bates DW, Kvedar JC. Digital Health and Patient Safety. JAMA 2016 Apr 26;315(16):1697-1698. [doi: 10.1001/jama.2016.2402] [Medline: 27115372]
12. Horvitz E, Mulligan D. Policy forum. Data, privacy, and the greater good. Science 2015 Jul 17;349(6245):253-255. [doi: 10.1126/science.aac4520] [Medline: 26185242]
13. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. J Med Syst 2012 Oct;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7] [Medline: 22052129]

14. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3. [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
15. Yusif S, Soar J, Hafeez-Baig A. Older people, assistive technologies, and the barriers to adoption: A systematic review. Int J Med Inform 2016 Oct;94:112-116. [doi: 10.1016/j.ijmedinf.2016.07.004] [Medline: 27573318]
16. Mittelstadt B. Ethical Issues of Personal Health Monitoring: A Literature Review. 2011 Presented at: ETHICOMP; 2011; Sheffield.
17. Chan M, Estève D, Fourniols J, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012 Nov;56(3):137-156. [doi: 10.1016/j.artmed.2012.09.003] [Medline: 23122689]
18. Khoja S, Durrani H, Nayani P, Fahim A. Scope of policy issues in eHealth: results from a structured literature review. J Med Internet Res 2012 Feb 17;14(1):e34 [FREE Full text] [doi: 10.2196/jmir.1633] [Medline: 22343270]
19. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]
20. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
21. Torous J, Roberts LW. Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
22. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
23. Kelton K, Fleischmann KR, Wallace WA. Trust in digital information. J. Am. Soc. Inf. Sci 2008 Feb 01;59(3):363-374. [doi: 10.1002/asi.20722]
24. Adjekum A, Ienca M, Vayena E. What Is Trust? Ethics and Risk Governance in Precision Medicine and Predictive Analytics. OMICS 2017 Dec;21(12):704-710 [FREE Full text] [doi: 10.1089/omi.2017.0156] [Medline: 29257733]
25. O'Neill O. A Question of Trust : The BBC Reith Lectures 2002. Cambridge: Cambridge University Press; 2002:4-8.
26. O'Neill O. Intelligent Trust in a Digital World. New Perspectives Quarterly 2017;34(4):6-31 [FREE Full text] [doi: 10.1111/npqu.12105]
27. Dorr GS, Lipkin M. The doctor-patient relationship: challenges, opportunities, and strategies. J Gen Intern Med 1999 Jan;14 Suppl 1:S26-S33 [FREE Full text] [Medline: 9933492]
28. Bending ZJ. Reconceptualising the doctor-patient relationship: recognising the role of trust in contemporary health care. J Bioeth Inq 2015 Jun;12(2):189-202. [doi: 10.1007/s11673-014-9570-z] [Medline: 25124984]
29. O'Neill O. Autonomy and Trust in Bioethics. Cambridge: Cambridge University Press; 2002:16-21.
30. Gilson L. Trust and the development of health care as a social institution. Soc Sci Med 2003 Apr;56(7):1453-1468. [Medline: 12614697]
31. Rowe R, Calnan M. Trust relations in health care: developing a theoretical framework for the “new” NHS. J Health Organ Manag 2006;20(5):376-396. [doi: 10.1108/14777260610701777] [Medline: 17087401]
32. Hawley K. Trust, Distrust and Commitment. Noûs 2012 Oct 25;48(1):1-20. [doi: 10.1111/nous.12000]
33. Sztompka P. Trust: a Sociological Theory. Cambridge: Cambridge University Press; 1999:44-46.
34. Giddens A, Anthony. The consequences of modernity. Stanford: Cambridge, UK : Polity Press in association with Basil Blackwell, Oxford, UK, 1990; 1990:25-36.
35. Montague EN, Kleiner BM, Winchester WW. Empirically understanding trust in medical technology. International Journal of Industrial Ergonomics 2009 Jul;39(4):628-634. [doi: 10.1016/j.ergon.2009.01.004]
36. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
37. Tung F, Chang S, Chou C. An extension of trust and TAM model with IDT in the adoption of the electronic logistics information system in HIS in the medical industry. Int J Med Inform 2008 May;77(5):324-335. [doi: 10.1016/j.ijmedinf.2007.06.006] [Medline: 17644029]
38. Buscemi N, Hartling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. J Clin Epidemiol 2006 Jul;59(7):697-703. [doi: 10.1016/j.jclinepi.2005.11.010] [Medline: 16765272]
39. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
40. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535 [FREE Full text] [Medline: 19622551]
41. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
42. Plomp HN, Ballast N. Trust and vulnerability in doctor-patient relations in occupational health. Occup Med (Lond) 2010 Jun;60(4):261-269. [doi: 10.1093/occmed/kqq067] [Medline: 20511267]

43. Hall MA, Dugan E, Zheng B, Mishra AK. Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? Milbank Q 2001;79(4):613-39, v [FREE Full text] [Medline: 11789119]
44. Hardiker NR, Grant MJ. Factors that influence public engagement with eHealth: A literature review. Int J Med Inform 2011 Jan;80(1):1-12. [doi: 10.1016/j.ijmedinf.2010.10.017] [Medline: 21112244]
45. Abelson J, Miller FA, Giacomini M. What does it mean to trust a health system? A qualitative study of Canadian health care values. Health Policy 2009 Jun;91(1):63-70. [doi: 10.1016/j.healthpol.2008.11.006] [Medline: 19117635]
46. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795. [doi: 10.1056/NEJMp1500523] [Medline: 25635347]
47. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. Digital Med 2018 Mar 14;1(1):1-4. [doi: 10.1038/s41746-017-0012-2]
48. Blasimme A, Vayena E. Becoming partners, retaining autonomy: ethical considerations on the development of precision medicine. BMC Med Ethics 2016 Dec 04;17(1):67 [FREE Full text] [doi: 10.1186/s12910-016-0149-6] [Medline: 27809825]
49. Blasimme A, Vayena E. “Tailored-to-You”: Public Engagement and the Political Legitimation of Precision Medicine. Perspectives in Biology and Medicine 2016;59(2):172-188. [doi: 10.1353/pbm.2017.0002]
50. Velsen LV, Tabak M, Hermens H. Measuring patient trust in telemedicine services: Development of a survey instrument and its validation for an anticoagulation web-service. Int J Med Inform 2017 Dec;97:52-58. [doi: 10.1016/j.ijmedinf.2016.09.009] [Medline: 27919395]
51. Chen JH, Asch SM. Machine Learning and Prediction in Medicine - Beyond the Peak of Inflated Expectations. N Engl J Med 2017 Jun 29;376(26):2507-2509. [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
52. Shapshay SM. Artificial intelligence: the future of medicine? JAMA Otolaryngol Head Neck Surg 2014 Mar;140(3):191. [doi: 10.1001/jamaoto.2013.6465] [Medline: 24457970]
53. Beam AL, Kohane IS. Translating Artificial Intelligence Into Clinical Care. JAMA 2016 Dec 13;316(22):2368-2369. [doi: 10.1001/jama.2016.17217] [Medline: 27898974]
54. Vayena E, Blasimme A. Health Research with Big Data: Time for Systemic Oversight. J Law Med Ethics 2018 Mar 27;46(1):119-129. [doi: 10.1177/1073110518766026] [Medline: 30034208]
55. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009 Jun;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]

Abbreviations

EHR: electronic health record
FDA: Food and Drug Administration
HA: health administrator
HCP: health care professional
IT: information technology
mHealth: mobile health

Edited by G Eysenbach; submitted 09.06.18; peer-reviewed by L van Velsen, C Fry; comments to author 17.08.18; revised version received 28.09.18; accepted 29.09.18; published 13.12.18

Please cite as:
Adjekum A, Blasimme A, Vayena E
Elements of Trust in Digital Health Systems: Scoping Review
J Med Internet Res 2018;20(12):e11254
URL: https://www.jmir.org/2018/12/e11254/
doi: 10.2196/11254
PMID: 30545807

Digit Biomark 2019;3:92–102

DOI: 10.1159/000502000
Received: May 2, 2019
Accepted after revision: July 8, 2019
Published online: August 16, 2019

Viewpoint – Review Article

Traditional and Digital Biomarkers: Two Worlds Apart?

Lmar M. Babrak[a] Joseph Menetski[b] Michael Rebhan[c, d]
Giovanni Nisato[e] Marc Zinggeler[f] Noé Brasier[g] Katja Baerenfaller[h]
Thomas Brenzikofer[d] Laurenz Baltzer[i] Christian Vogler[j] Leo Gschwind[j]
Cornelia Schneider[k, l] Fabian Streiff[d] Peter M.A. Groenen[d, m]
Enkelejda Miho[a, d, n]

[a] FHNW University of Applied Sciences Northwestern Switzerland, Muttenz, Switzerland; [b] Foundation for the National Institutes of Health, North Bethesda, MD, USA; [c] Novartis Institutes for Biomedical Research, Basel, Switzerland; [d] DayOne, BaselArea.Swiss, Basel, Switzerland; [e] Health Hacking Lab, Basel, Switzerland; [f] CSEM, Muttenz, Switzerland; [g] CMIO Research Group, University Hospital Basel, Basel, Switzerland; [h] Swiss Institute of Allergy and Asthma Research (SIAF), University of Zurich, and Swiss Institute of Bioinformatics (SIB), Davos, Switzerland; [i] S. Karger AG, Basel, Switzerland; [j] Advancience AG, Allschwil, Switzerland; [k] Clinical Pharmacy and Epidemiology, Department of Pharmaceutical Sciences, University of Basel, Basel, Switzerland; [l] Hospital Pharmacy, University Hospital Basel, Basel, Switzerland; [m] Idorsia Pharmaceuticals Ltd., Translational Science, Allschwil, Switzerland; [n] aiNET GmbH, Basel, Switzerland

Keywords

Traditional biomarkers · Digital biomarkers · Semantics · Biomarker classification · Digital health

Abstract

The identification and application of biomarkers in the clinical and medical fields has an enormous impact on society. The increase of digital devices and the rise in popularity of health-related mobile apps has produced a new trove of biomarkers in large, diverse, and complex data. However, the unclear definition of digital biomarkers, population groups, and their intersection with traditional biomarkers hinders their discovery and validation. We have identified current issues in the field of digital biomarkers and put forth suggestions to address them during the DayOne Workshop with participants from academia and industry. We have found similarities and differences between traditional and digital biomarkers in order to synchronize semantics, define unique features, review current regulatory procedures, and describe novel applications that enable precision medicine.

Twitter handles: @BaselAreaSwiss @aiHealthLab

Enkelejda Miho
FHNW University of Applied Sciences Northwestern Switzerland
Hofackerstrasse 30
CH–4132 Muttenz (Switzerland)
E-Mail enkelejda.miho @ fhnw.ch

Peter Groenen
Idorsia Pharmaceuticals Ltd
Hegenheimermattweg 91
CH–4123 Allschwil (Switzerland)
E-Mail peter.groenen @ idorsia.com

Introduction

Biomarkers are an integral part of biomedical research and clinical practice. Many common research assays and clinical tests, such as measuring cholesterol levels, blood pressure, body temperature, or pulmonary function tests serve as biomarkers for assessing health states. In clinical trials, well-defined traditional biomarkers are important and accepted metrics to objectively assess clinical status, therapeutic effects such as early predictors for clinical endpoints, and adverse events. Due to the globally increasing importance of biomarkers in modern healthcare, in 1998 the US National Institute of Health (NIH) Biomarkers Definitions Working Group began to define clinically associated terms and opinions on the qualification of novel methodologies for medicine development. Since, they have also been adopted by the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) [1]. A biomarker is defined as “a characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention including therapeutic interventions. This can include molecular, histological, radiographic, or physiologic characteristics” [2].

The FDA-NIH Joint Leadership Council put together the Biomarkers, Endpoints and other Tools Resource (BEST) in order to harmonize translational science by standardizing medical terms and improve scientific understanding regarding biomarkers in clinical development and practice [2]. In BEST, the FDA-NIH Biomarker Working Group defined two categories of biomarkers: (i) disease-associated biomarkers (susceptibility/risk biomarker, diagnostic biomarker, prognostic biomarker, monitoring biomarker), and (ii) drug-related biomarkers (predictive biomarker, pharmacodynamics/response biomarker, safety biomarker). For example, body mass index (BMI) measurement is a risk biomarker used to assess the risk for metabolic diseases such as diabetes mellitus, hypertension, and dyslipidemia, and can lead to preventative health measures [3]. Biomarkers can also be safety indicators for determining health status, such as liver function, by measuring serum creatinine, blood urea nitrogen, and cystatin C [4, 5]. The working group also describes other clearly standardized biomarker-relevant definitions, e.g., “fit for purpose” and “context of use (COU)” [2].

Due to the long and resource-intensive process, incorporation of fully qualified biomarkers in the medical field is very challenging. It begins with biomarker discovery, animal experiments, clinical or epidemiological studies, analytical validation, and interventional studies with the gold standard endpoints [6–8]. Additionally, there are many stakeholders involved in the qualification of biomarkers: academia, technology companies, biopharmaceutical industry, regulatory agencies, doctors, and patients add to the complexity of biomarker qualification and acceptance.

As digital devices have begun to be integrated into the medical landscape, digital biomarkers have become an exciting new tool for advancing precision medicine and supporting clinical trials. Digital biomarkers are objective, quantifiable, physiological, and behavioral measures that are collected by means of digital devices that are portable, wearable, implantable, or digestible [9]. These data are often used to explain, influence, and/or predict health-related outcomes. Individualized measures can now be recorded to create personalized baselines for health. The inclusion of digital biomarkers has specifically impacted the field of neurology where there is a great unmet need for objective and non-invasive biomarkers. Currently, a number of digital biomarkers are being tested for feasibility and reliability in Parkinson’s and Alzheimer’s disease and clinical outcome assessments [10, 11].

As the emerging field of digital biomarkers joins traditional biomarkers in the health(care) innovation process [12], it is crucial to address rising new possibilities and challenges, but most importantly define a standardized nomenclature and process to allow for a more rapid clinical uptake. To tackle this pressing need, the BaselArea.Swiss organized the DayOne

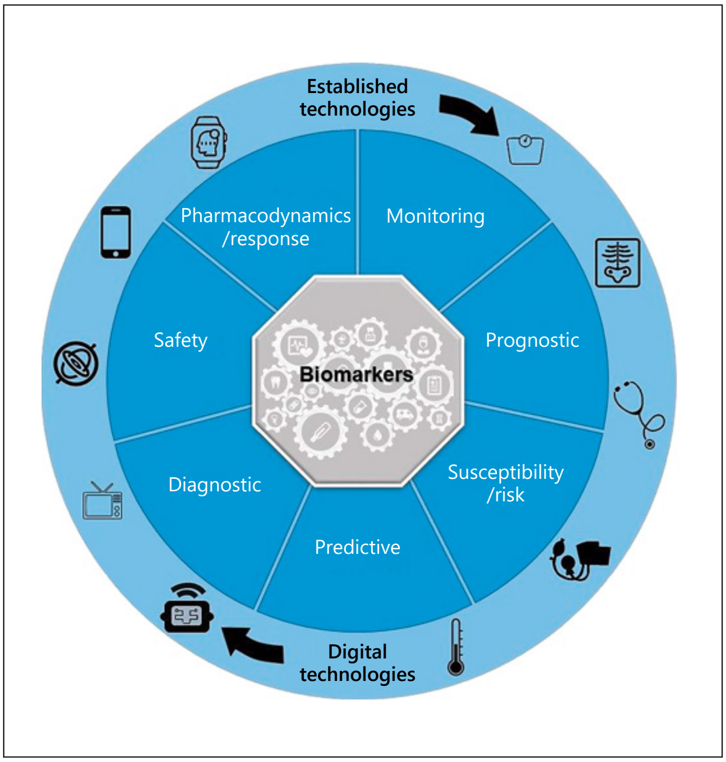


Fig. 1. Digital technologies have enabled the measurement of digital biomarkers. Traditional biomarkers are divided into seven categories depending on the clinical goal for the biomarker (i.e., safety, predictive, diagnostic, etc.). We propose that digital technologies enable the measurement of digital biomarkers, which contribute new and unique features (e.g., longitudinal, continuous measurements) while sharing the same traditional biomarker clinical goals.

Workshop “Traditional and Digital Biomarkers – Two Worlds Apart” on March 27, 2019 in Basel, Switzerland. A diverse and relevant group of stakeholders from academia and industry attended, bringing together multiple viewpoints. In this perspective, we present our findings on identifying similarities and differences between traditional and digital biomarkers and a discussion on how these fields could be further harmonized by addressing semantics, features unique to each area, and novel applications in the medical field.

A Comparison of Traditional and Digital Biomarker Characteristics

Despite the apparent similarities between traditional and digital biomarkers in the way they address important challenges related to the health-disease continuum, there are also important differences in culture, innovation, scientific and technical maturity, and the nature of the data. Simply, digital biomarkers fall within the scope of traditional biomarkers in relation to addressing health-related questions, with use of a digital and portable technology that adds new dimensions, unique features and challenges (Fig. 1; Tables 1, 2).

Traditional biomarkers are often well embedded into clinical practice and research, usually in proximity to the pathological event of interest. They are generally limited in analytical complexity and can range from being qualitative to quantitative. However, such traditional biomarkers can be invasive and expensive to measure. Due to the dynamic, complex nature of disease, traditional biomarkers often present an incomplete view due to the limited number of measurements that can be collected over time (“snapshot” problem). Alternatively, digital biomarkers are usually less or non-invasive, modular, and often cheaper to measure. They can produce qualitative and quantitative measurements, but most importantly, they provide easier and cheaper access to continuous and longitudinal measurements. Nonetheless, digital biomarkers are still new and therefore are not commonly implemented

Table 1. Overview of traditional and digital biomarker features, novel applications, and published examples

Biomarkers		
Definition: "a characteristic that is measured as an indicator of normal biological or pathogenic processes, or responses to an exposure or intervention" [2]		
Feature in digital biomarkers	Novel applications	Examples
Individual/ population level health data	Clinical trial patient stratification Clinical trial patient recruitment Development of personalized/population baselines Population health management	Detection risk of dementia with computerized cognitive testing [13] Wearable sensors to monitor gait performance in Huntington disease (recorded >14,000 assessments vs. 20 in the clinic) [14]
Longitudinal and continuous measurements	More data points will lead to higher granularity and stronger analysis Clearer understanding of health status High resolution of data to stratify subgroups within a population Measures change over time Serves as a control for disease states Measure/observe episodic medical occurrences in real-time, outside of clinic	Analysis of finger tapping and memory tests conducted on smartphones for characterization of longitudinal characteristics of Parkinson's disease [15] Voice, finger tapping, gait, and reaction time on smartphone application for developing an objective measure of PD severity [16] Wearable devices for real-time detection of epileptic seizures to better understand the condition and option to contact emergency services [17]
Passive monitoring	Facilitates monitoring outside of hospital Objective data (individual-independent) Low patient burden; higher adherence	Automated analysis of free speech to predict psychosis onset in youths [18] Passively acquired accelerometer data as primary endpoint to measure the effect of isosorbide mononitrate treatment in patients with heart failure [19]

Table 2. Current challenges of digital biomarkers in medical settings

Challenges	Explanation
Semantics	Lack of standardization in the choice of words describing novel concepts in digital biomarkers (e.g., longitudinal data)
Data standardization	Large number of heterogeneous data sources, formats, scales
Data privacy: protection of user data, anonymization	Regulatory guidelines on privacy, right to opt out, informed consent, data ownership is not fully developed, continuously changing, and different between countries
Data storage	Large volumes of complex data are being generated but there are few guidelines on how they should be stored
Identification of relevant data and interpretation	Identification of pertinent data is important for analysis, and how to correctly analyze the data and use accurate baselines is still unclear; this process is often not transparent due to proprietary issues
Regulatory approval: data analytics, algorithm and tool validation	Regulatory standards lag behind rapid innovation in tool, analytics, and algorithm development

in clinical development and practice. Ease-of-use and low cost come to the detriment of often measuring distally from pathological events, and producing large, complex data that are challenging for data analytics. Some additional drawbacks include data integrity assessment and reliability, hardware malfunctions, identification of relevant data, multiple interpretations, and baseline determination.

Standardizing Novel Semantics Is a Challenge

In any emerging field, a high degree of ambiguity in semantics is typical. In digital biomarkers this results from the merging of several fields (healthcare, engineering, and analytics), from the commercial marketing of mobile products directly to the public, and from a lack of shared practices and standardization. Further, proprietary devices and corresponding digital biomarkers are introduced in the market with differing terminologies used for medical, scientific, and marketing purposes. For example, gait – the way an individual walks – is termed and described differently depending on the context (e.g., medical, engineering, analytics) and would be described differently for marketing purposes depending on the target demographic group. Adding to this disparity, there is also a cultural and historical ambiguity in terminology. There are terms in the regulatory space that have legacy connotations that may no longer apply, or are co-opted or newly created without sufficient standardization and consensus among participating stakeholders. These semantic issues generate confusion by creating conflicting or divergent terms. While a certain degree of semantic ambiguity is present, we propose using traditional biomarker semantics to describe digital biomarkers. While terms such as "surrogate endpoint," and "validation" can apply to both types of biomarkers, additional concepts such as continuous and longitudinal data must be defined to encompass these new parameters and how these terms relate to each other.

Community-based semantic disambiguation in emerging areas is no trivial pursuit and requires an agile, iterative process of refining and shaping terms that enables the community to focus on areas that need our attention first, considering the costs of non-action. For these new terms to be incorporated and modified, a living and dynamic standard for terms needs to be assembled and maintained. This allows new terms to be integrated while removing outdated, irrelevant terms. This living dictionary will need to be flexible, keep up with the rapid speed of technological innovation, and reflect community decisions by expert decision-makers, producers, and users of data, and be based in the regulated field of patient healthcare.

Digital Biomarker Users Consist of a Diverse Population

While traditional biomarkers are applied to a population group seeking to improve or assess their health status, the digital biomarker application landscape is broader. It encompasses three different population groups: (i) the casual, curiosity, or fitness-based individual for non-clinical use (often in the healthier section of the health-disease continuum), (ii) the commercial-based group, and (iii) the regulatory and clinical trial group (typically in the more diseased cluster of the health-disease continuum; Fig. 2). To fully understand and address the challenges in the field of digital biomarkers, the diversity of the population groups that produce these data needs to be identified and attributes described. Figure 1 depicts three different population groups and highlights several unique characteristics within these groups. The first group falls into the casual non-clinical use of digital biomarkers, which includes those who are curious, fitness-trackers, the "quantified self," and individual-driven phenotyping or status tracking. This population group consists of the largest number of individuals, but contains the least amount of reliable, validated, and consistent data, and is at a higher risk of data breach [20, 21]. These attributes prevent the use of this type of data in a medical context. Technological industries associated with this group typically change products rapidly, preventing them from complying with complex regulatory settings which would normally slow down the ability to adapt products quickly to a rapidly changing market.

The second population consists of users of digital biomarkers for commercial purposes such as fitness coaching for professional athletes and for facilitating decision-making in

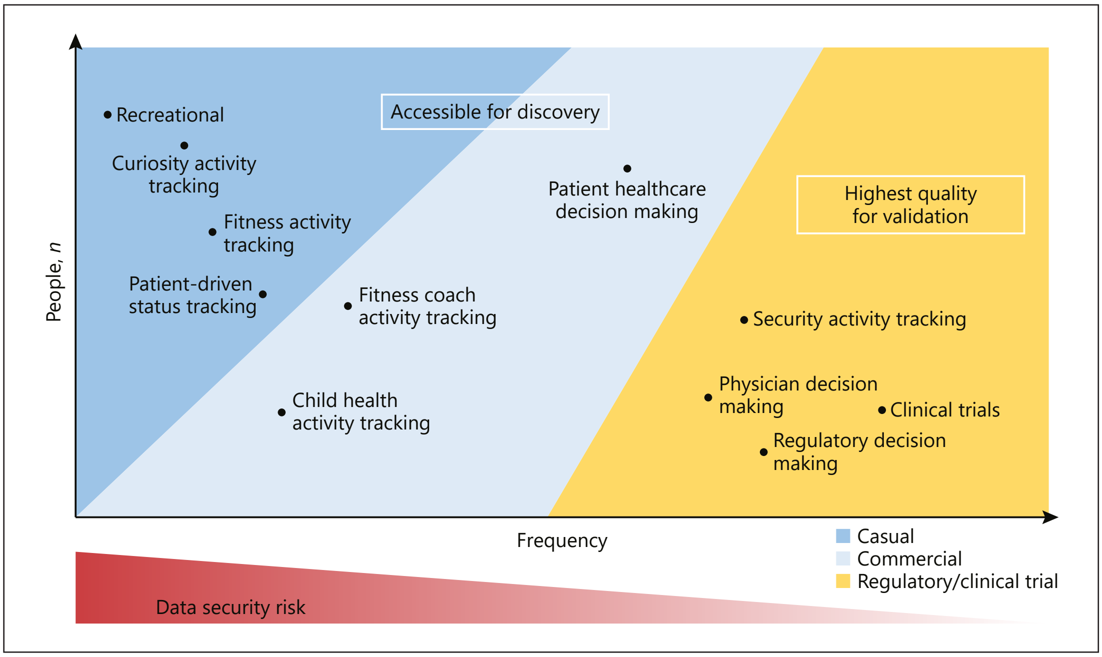


Fig. 2. User populations. Different populations of data producers of potential digital biomarkers and device users are grouped based on application, size of population, consistency of measure, and data security risk, and are categorized based on the best potential use of the data for discovery of novel digital biomarkers or validation. The *y*-axis describes the number of individuals measured and the *x*-axis describes the consistency of the measure. Consistency is defined as accuracy, precision, reproducibility, calibration, and traceability.

healthcare. These users are often located somewhere in the center of the health-disease continuum. This group is highly variable in size but is generally smaller than the first population group. For example, ORECCO, a privately-owned company, provides professional athletes with performance results based on the digital biomarkers they collect [22].

The third population group consists of the device users and data producers involved with clinical or regulatory decision-makers, pharmaceutical companies, medtech, or physicians. This population group is the smallest and the security against improper alteration of the data needs to be high due to the medical and regulatory implications. The consistency and quality of the data (i.e., accuracy, precision, reproducibility, calibration, and traceability) is the highest, containing strict definitions, protocols, and predetermined analyses, making this group ideal for digital biomarker validation. In this group, physicians specialized in particular diseases typically recruit participants according to specific inclusion and exclusion criteria, and traditionally no information optimized for interpretation by patients goes back to the patients.

Despite their differences, the three population groups can provide valuable information with regard to population baselines and health status (casual and commercial groups), which can be formally implemented for biomarker validation (regulatory/clinical study group). In the digital biomarker landscape, the number of individuals measured, data security, and consistency of data vary greatly. Due to the diversity of features demonstrated by each population group depicted in Figure 2, there is an expected level of confusion and uncertainty regarding how to define terms and apply regulatory applications. We need to recognize that there is a breadth of digital biomarker applications and each should be defined separately, standardized appropriately, and regulated accordingly depending on the population group.

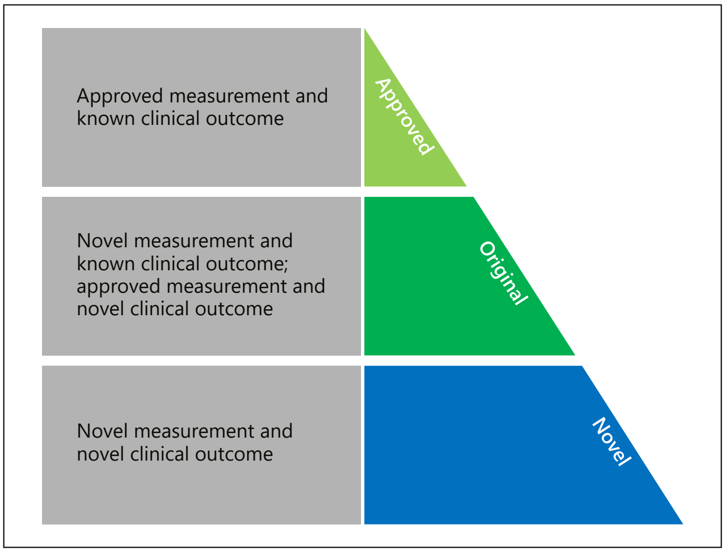


Fig. 3. Biomarker classification. Both traditional and digital biomarkers can be classified based on the status of a particular measurement to a particular clinical status or outcome. A digital biomarker either replaces a non-digital biomarker (Approved), opens a new field (Novel) or is a hybrid that on the one hand replaces and on the other hand opens a new field (Original).

The Classification of Digital Biomarkers

Although there are many perspectives on how digital technologies can be classified (i.e., clinical outcomes or biomarkers), we set to further clarify and describe digital biomarkers in the scope of clinical development and practice. We propose a method of classifying digital biomarkers by taking into account the digital measurement tool and clinical outcome assessment (COA; Fig. 3, adapted from [23]). COA is the monitoring of how a patient feels, functions, or survives. An important corollary to the COA definition is that a biomarker is not a measurement of how a patient feels, functions, or survives. Digital biomarkers and their associated clinical assessment outcomes are partitioned into three types. “Approved” describes biomarkers that improve or accelerate generally accepted practices. These are digital biomarkers that consist of approved measurements such as heart rate, pulse, and known clinical outcomes, such as cardiac risk [24]. “Original” characterizes biomarkers from two scenarios: (i) a novel measurement with a known clinical outcome such as gait measurement for Parkinson’s disease assessment [10] and (ii) an approved measurement, such as heart rate to describe a novel clinical outcome such as depression [25]. “Novel” describes a novel measurement such as physical activity or facial expressions, with a novel clinical outcome such as depression [26]. “Approved” digital biomarkers will likely be the first ones to be used as they are already medically validated and implemented in the clinic. Digital biomarkers in “Original,” and especially in “Novel,” will require rigorous testing and validation to become adopted in clinical development. This new classification system will help facilitate and clarify the type of biomarker and clinical outcome being described and facilitate regulations.

Regulatory Concepts in Digital Biomarkers

The regulatory procedures concerning biomarkers, including digital, is complex because they can be applied to a broad spectrum of uses such as drugs/biologicals and medical devices, and regulated differently in various countries. In Switzerland, the council directive concerning medical devices 93/42/EEC (Swiss Agency for Therapeutic Products, Swissmedic) defines "medical device" broadly as any instrument, apparatus, appliance, software, material, or other article, whether used alone or in combination, used specifically for diagnostic and/or therapeutic purposes. In the USA, the Food and Drug Administration (FDA) oversees the Center for Drug Evaluation and Research (CDER), the Center for Biologics Evaluation and Research (CBER), and the Center for Devices and Radiological Health (CDRH), which manage the approval of new drugs, biologics, and medical devices, respectively. Currently in Europe, the European Medicines Agency (EMA) manages the approval of medical devices and drugs, but recently, a new legislation has created the Medical Device Regulation (MDR), which will oversee all new medical devices and enforce more stringent rules compared to the CDRH for medical device approval in 2020 [27]. Generally, in the USA, medical devices with software that operates the device and sensors have been categorized as "software in medical device" (SiMD). With the introduction of digital biomarkers, new terminology was introduced: "software as a medical device" (SaMD). SaMD is defined as software that performs a medical function without being part of the hardware (e.g., machine learning tools in mobile apps) [28]. An example of an SaMD is the Apple Watch software for atrial fibrillation detection where the Apple Watch serves a component supporting digital biomarker measurements (atrial fibrillation).

For drug and biological measurements and approvals, digital tools are being adopted in the first three phases of clinical trials, and the data derived need to be collected early and consistently to develop the necessary evidence for internal decision-making (phase I and II clinical trials) and regulatory approval (phase III) [28]. In the regulatory area, tools and measurements for medical product development are termed drug development tools or medical device development tools. The FDA also has a Prescription Drug-Use-Related Software that is developed for use with prescription drugs (e.g., drug ingestion tracking, dose calculation, remainders, and drug instructions). Similar to the drug-related products described above, medical devices go through a process for clearance or approval [28]. In this process, the technical aspects of the design of the product are more carefully considered. The intended use of the device determines if and how the item is regulated. If the medical device is only intended to be used for wellness purposes then no regulation is necessary, but if the same device claims to diagnose or monitor a health condition, it needs to be regulated. Whether software is considered a device is determined by a regulatory body and will be highly dependent on the intended functions. The FDA continues to update guidelines in order to clearly determine which devices should be regulated, but the distinctions of what is considered a medical device is still unclear, and often exceptions are made.

Novel Applications for the Fields of Precision Medicine

Digital biomarkers are also extremely well suited for applications and medical areas that rely on subjective measures such as neurology and psychiatry. Digital biomarkers can provide objective measurements to base and support diagnosis, prognosis, and measure therapeutic outcomes. In fields like Alzheimer's and Parkinson's disease, digital outcome measurements bring a paradigm shift of how treatment outcomes are measured and assessed, particularly in very early (prodromal) disease states [29]. Currently, many pharmaceutical companies are running pilot studies to test the feasibility of using digital biomarkers. For example, Roche has built a Parkinson's disease app to measure active and passive PD tests and assess disease severity [30], Biogen is in partnership with PatientsLikeMe trying to understand physical activity measurements in patients with multiple sclerosis [31], and Neurotrack has remotely measured cognitive ability to assess cognition in patients with Alzheimer's disease [32]. In psychiatry, using "digital behavioral biomarkers" such as phone usage patterns in health disorders has had encouraging results [33]. Additionally, digital biomarkers are predictive for the diagnosis and prognosis of symptom severity in patients with major depressive or bipolar disorders [34]. Unfortunately, there are few published results and currently it is difficult to assess the overall feasibility, technology and biology behind digital biomarkers.

Conclusion

Digital continuous monitoring has brought a wealth of new data to healthcare. Retrospectively in science, there have been many instances where new technology has propelled a paradigm shift that enabled new questions and novel insights. Digital biomarkers may not shift the medical paradigm directly, but these novel ways of measuring health status provide observations and perspectives into disease that were unavailable before. They supplement and enhance conclusions from traditional biomarkers (e.g., dynamic changes measured with digital biomarkers during the course of a disease). Detailed longitudinal measurements, combined with accurate and precise assessments from molecular characterization of health and disease, have the potential to redefine diagnosis and the medical classification of diseases. Similar to novel stratifications in cancer groups due to molecular profiling, digital biomarkers can further support precision medicine and lead to innovative treatments. Lastly, digital biomarkers provide information that has the potential to greatly influence not only what we know about disease and prevention, but our very understanding of health.

Acknowledgements

We would like to acknowledge the DayOne community.

Disclosure Statement

J.M. declares holding shares in Merck Pharmaceuticals (MSD outside the USA). P.M.A.G. declares holding shares in Idorsia Pharmaceuticals. E.M. declares holding shares in aiNET GmbH. L.B. is a publishing manager at Karger AG. The remaining authors declare no competing interests.

Funding Sources

Authors are grateful to FHWN University of Applied Sciences and Arts Northwestern Switzerland and BaselArea.Swiss for funding this research.

Author Contributions

All authors contributed to the discussions and wrote the manuscript. L.M.B, J.M., and E.M. produced the figures.

References

1 European Medicines Agency, Committee for Medicinal Products Human Use (CHMP)European Medicines Agency. (2018) (available at https://www.ema.europa.eu/en/committees/committee-medicinal-products-human-use-chmp)

2 FDA-NIH Biomarker Working Group. BEST (biomarkers, endpoints, and other tools) resource. Silver Spring: FDA; 2016. Available from: http://www.ncbi.nlm.nih.gov/books/NBK326791/

3 Bays HE, Chapman RH, Grandy S; SHIELD Investigators' Group. The relationship of body mass index to diabetes mellitus, hypertension and dyslipidaemia: comparison of data from two national surveys. Int J Clin Pract. 2007 May;61(5):737–47.

4 Kar S, Paglialunga S, Islam R. Cystatin C is a more reliable biomarker for determining eGFR to support drug development studies. J Clin Pharmacol. 2018 Oct;58(10):1239–47.

5 Bonventre JV, Vaidya VS, Schmouder R, Feig P, Dieterle F. Next-generation biomarkers for detecting kidney toxicity. Nat Biotechnol. 2010 May;28(5):436–40.

6 Kraus VB. Biomarkers as drug development tools: discovery, validation, qualification and use. Nat Rev Rheumatol. 2018 Jun;14(6):354–62.

7 Allinson JL. Clinical biomarker validation. Bioanalysis. 2018 Jun;10(12):957–68.

8 Hayes DF. Biomarker validation and testing. Mol Oncol. 2015 May;9(5):960–6.

9 Dorsey ER, editor. Digital biomarkers. Basel: Karger. Available from: https://www.karger.com/Journal/Home/271954

10 Lipsmeier F, Taylor KI, Kilchenmann T, Wolf D, Scotland A, Schjodt-Eriksen J, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. Mov Disord. 2018 Aug;33(8):1287–97.

11 Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. NPJ Digit Med. 2019;2:9.

12 Wright JM, Jones GB. Harnessing the Digital Exhaust: incorporating wellness into the pharma model. Digit Biomark. 2018 Jan-Apr;2(1):31–46.

13 Ritchie K, Carrière I, Su L, O'Brien JT, Lovestone S, Wells K, et al. The midlife cognitive profiles of adults at high risk of late-onset Alzheimer's disease: the PREVENT study. Alzheimers Dement. 2017 Oct;13(10):1089–97.

14 Andrzejewski KL, Dowling AV, Stamler D, Felong TJ, Harris DA, Wong C, et al. Wearable Sensors in Huntington Disease: A Pilot Study. J Huntingtons Dis. 2016 Jun;5(2):199–206.

15 Prince J, Arora S, de Vos M. Big data in Parkinson's disease: using smartphones to remotely detect longitudinal disease phenotypes. Physiol Meas. 2018 Apr;39(4):044005.

16 Zhan A, Mohan S, Tarolli C, Schneider RB, Adams JL, Sharma S, et al. Using Smartphones and Machine Learning to Quantify Parkinson Disease Severity: The Mobile Parkinson Disease Score. JAMA Neurol. 2018 Jul;75(7):876–80.

17 Bruno E, Simblett S, Lang A, Biondi A, Odoi C, Schulze-Bonhage A, et al. Wearable technology in epilepsy: the views of patients, caregivers, and healthcare professionals. Epilepsy Behav. 2018 Aug;85:141–9.

18 Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, Ribeiro S, Javitt DC, Copelli M, Corcoran CM. Automated analysis of free speech predicts psychosis onset in high-risk youths. NPJ Schizophr. 2015;1:15030.

19 Redfield MM, Anstrom KJ, Levine JA, Koepp GA, Borlaug BA, Chen HH, et al.; NHLBI Heart Failure Clinical Research Network. Isosorbide Mononitrate in Heart Failure with Preserved Ejection Fraction. N Engl J Med. 2015 Dec;373(24):2314–24.

20 Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA. 2015 Feb;313(6):625–6.

21 Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil. 2015 Oct;7(1):24.

22 Orreco. Data driven solutions for optimizing athlete performance. Galway: Orreco; 2019. Available from: https://orreco.com/

23 Wang T, Azad T, Rajan R. The emerging influence of digital biomarkers on healthcare. San Francisco: Rock Health; 2019. Available from: https://rockhealth.com/reports/the-emerging-influence-of-digital-biomarkers-on-healthcare/

24 Perret-Guillaume C, Joly L, Benetos A. Heart rate as a risk factor for cardiovascular disease. Prog Cardiovasc Dis. 2009 Jul-Aug;52(1):6–10.

25 Carney RM, Freedland KE. Depression and heart rate variability in patients with coronary heart disease. Cleve Clin J Med. 2009 Apr;76(suppl 2):S13–7.

26 Girard JM, Cohn JF, Mahoor MH, Mavadati S, Rosenwald DP. Social Risk and Depression: Evidence from Manual and Automatic Facial Expression Analysis, Proc Int Conf Autom Face Gesture Recognit, 1–8 (2013).

27 Swissmedic. Information on the EU regulations (MDR, IVDR). Bern: Swissmedic; 2019. Available from: https://www.swissmedic.ch/swissmedic/en/home/medizinprodukte/neue-eu-verordnungen-mdr-ivdr-/aktuelles-zur-umsetzung.html

28 Coravos A, Goldsack JC, Karlin DR, Nebeker C, Perakslis E, Zimmerman N, et al. Digital medicine: a primer on measurement. Digit Biomark. 2019;3:31–71.

29 Gold M, Amatniek J, Carrillo MC, Cedarbaum JM, Hendrix JA, Miller BB, et al. Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. Alzheimers Dement. 2018 May;4:234–42.

30 Roche. Roche technology measures Parkinson's disease fluctuations. Basel: Roche; 2019. Available from: https://www.roche.com/media/store/roche_stories/roche-stories-2015-08-10.htm
31 Moore C. PatientsLikeMe/Biogen study evaluates wearable devices to monitors activity of people with MS. Philadelphia: Multiple Sclerosis News Today; 2015. Available from: https://multiplesclerosisnewstoday.com/multiple-sclerosis-news/2015/04/21/patientslikemebiogen-study-evaluates-wearable-devices-monitors-activity-people-ms/
32 Neurotrack [Internet]. Redwood City: Neurotrack; 2019. Available from: https://www.neurotrack.com/science/
33 Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Margrethe Christensen E, Bardram JE, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. Int J Methods Psychiatr Res. 2016 Dec;25(4):309–23.
34 Jacobson NC, Weingarden H, Wilhelm S. Digital biomarkers of mood disorders and symptom change. NPJ Digit Med. 2019;2:3.

Wearable Devices in Clinical Trials: Hype and Hypothesis

Elena S. Izmailova[1], John A. Wagner[1] and Eric D. Perakslis[1]

The development of innovative wearable technologies has raised great interest in new means of data collection in healthcare and biopharmaceutical research and development. Multiple applications for wearables have been identified in a number of therapeutic areas; however, researchers face many challenges in the clinic, including scientific methodology as well as regulatory, legal, and operational hurdles. To facilitate further evaluation and adoption of these technologies, we highlight methodological and logistical considerations for implementation in clinical trials, including key elements of analytical and clinical validation in the specific context of use (COU). Additionally, we provide an assessment of the maturity of the field and successful examples of recent clinical experiments.

WHY CONSIDER USING DIGITAL DEVICES IN CLINICAL TRIALS?

Use of (and hype surrounding) wearable technologies has skyrocketed in recent years. We define here wearable technologies as sensors and/or software applications (apps) on smartphones and tablets that can collect health-related data remotely, i.e., outside of the healthcare provider's office. The data can be collected passively or may require a user's input. An accelerometer embedded in a wristband or a cell phone is an example of a sensor passively collecting data about a person's physical activity and movement. Software (e.g., ePRO (electronic Patient Reported Outcome)) can output a patient's report capturing health-related information, collected by means of a cell phone app or a web-based interface. Additionally, some technologies, such as smart-cap bottles designed to monitor medication adherence, can use a combination of a sensor and app-based data collection. The event recording is triggered by a user action (opening the bottle), but the data are transmitted from a sensor to a server passively via Bluetooth. The transmission is mediated by a cell phone app.

Ten years on since the introduction of the iPhone, we have witnessed an almost complete change in how people communicate with each other, access media/content, and interact with that content. Most noticeably, in healthcare and beyond, this shift has led to a complete change in the expectations surrounding reporting of events. Digital disease detection has shifted outbreak-detection timeframes from months to hours with social media.[1] The US Food and Drug Administration (FDA) now encourages safety adverse event reporting via mobile apps. Hospitals are using Fitbits on inpatients to monitor recovery and mobility. Patients interact regularly online with healthcare facilities. Twitter and other social media can report and post opinions on products and services far faster and more broadly than almost any business.[2,3]

At the same time, rising costs of healthcare are of immense concern and the possibility of healthcare virtualization via digital devices has been heralded by relentless hype. For remote monitoring of cardiovascular parameters, activity (including gait, balance, and many other forms of motion measurement), body temperature, galvanic skin response, blood oxygen saturation, and multisensor/multisystem monitoring,[4] advanced wearable device research and development is continuously improving. Common form factors include wearable watches/bracelets, patches, textiles, and garments (**Table 1**). All of these sensor devices are being built with the ability to monitor continuously and communicate data in real time or intermittently. While maturity, promise, and quality all vary greatly at the moment, clearly these sensors and devices have the potential to become an integral part of the future of healthcare and biopharmaceutical development.

PROMISES AND CHALLENGES OF USING WEARABLES IN CLINICAL TRIALS

Promises in healthcare

Wearable devices can collect data on a 24/7 basis in natural settings as people go through their daily routines at home and work. The data collection can be enhanced by digital diaries depicting key features of personal health and lifestyle. The best-known wearable devices are commercial fitness trackers that collect mobility and some vital sign data.[5] Similar wearables cannot be

---

[1]Takeda Pharmaceuticals International Co., Cambridge, Massachusetts, USA. Correspondence: Elena S. Izmailova (elena.izmailova@takeda.com)

Received 10 October 2017; accepted 29 November 2017; advance online publication 2 April 2018. doi:10.1002/cpt.966

Table 1 Examples of wearable sensors

Device type	Data collected	Examples
Wrist worn	Actigraphy, HR (Heart Rate), BP (Blood Pressure), EDA (Electrodermal activity)	Actiwatch Spectrum by Phillips, ActiGraph Link by ActiGraph, E4 by Empatica, ViSi Mobile by Sotera Wireless
Skin patch	ECG (Electrocardiography), actigraphy, skin temperature	BioStampRC by MC10, HealthPatch by Vital Connect, BodyGuardian by Preventice
Cuffs	BP, HR	Intellisense Digital BP Monitor by Omron Healthcare
Finger worn	HR, SpO2	iSpO2 Pulse Oximeter by Massimo
Clothing embedded sensors	HR, HRV (Heart Rate Variability), ECG, Breathing Rate, actigraphy	Smart shirts by Hexoskin
Headbands	EEG (Electroencephalogram), EMG (Electromyography)	EMOTIV EPOC by Emotiv, 4D FORCE by 4D FORCE

marketed as medical devices unless the device performance has been established prior to release to the market. This is a big step forward compared to the traditional means of health-related data collection. For example, basic physiological data, e.g., vital signs and telemetry, are traditionally collected only during doctor's office visits or as a part of medical product clinical trial procedures. These data represent a very limited snapshot of a person's phenotype and physiology. Inferences about a person's health are made based on the extrapolation of such a snapshot to extended periods of time, potentially weeks and months. This extrapolation is also based on patients' memory recall of incidents preceding the office visit. Decisions about the patient's health, disease status, and treatments are made comparing data collected in doctor's offices to population averages, which may or may not be relevant to a particular individual. Additionally, there are well-known issues related to in-clinic measurement of vital signs, including white-coat hypertension.[6] There is a growing recognition that population-based values need to be adjusted for factors such as age, gender, medication status, demographics, and other factors.[7,8] These adjustments can be made if there are data available for specific subpopulations of interest. This may also be done using the individual's own baseline data collected over extended periods of time, which would enable a precision medicine approach. Data frequently collected over extended periods of time can provide deeper understanding of disease variability, which is likely to be an important contributor to treatment response variability. Having larger and denser datasets will help to characterize intra- and interpatient variability. Additionally, there is growing evidence that replacing paper diaries with electronic versions can greatly improve the quality of subjectively reported outcome data,[9–11] such as pain and functional status, by ensuring compliance, timely collection of the data, avoidance of secondary data entry errors, and reduced administrative burden.[11] Replacing paper diaries and patient memory recall with electronic means of data collection is likely to continue and expand with technological advances in the future. Moreover, wearable device data combined with other data such as genomics or other high-throughput technologies have the potential to create a comprehensive multilayer picture of a person's health and can deepen our understanding of how to combine genotyping with deep phenotyping.

Promises in drug development

The applications mentioned above are also attractive for drug development in both early- and late-stage clinical trials. Collecting dense data from trial participants using wearables in natural settings—often not collectible otherwise—may fundamentally change how clinical trials are designed and conducted. In early clinical drug development, collection of dense physiological data may identify early safety issues and inform dose adjustments and dosing frequencies, or lead to discontinuation of development of certain drug candidates. The study subjects would not have to be confined to the pharmacology units all the time to have the data collected. In the late stages of clinical development, creating novel endpoints by means of wearable technologies has applications in multiple disease areas (**Table 2**). These novel endpoints may provide more sensitive measures of disease activity compared to traditional scales, enabling faster and more objective readouts in clinical trials. Additionally, sensors can provide objective measures of traditionally subjectively reported outcomes, such as pain and fatigue, complementing or even completely replacing self-reports. Another attractive feature includes portability to home settings and simplification of measures traditionally done in hospitals. Sleep data collection by means of actigraphy can serve as an example.[12] Important parameters of sleep, such as sleep duration and number and duration of awakenings, can be collected by wrist-worn actigraphy devices. This could replace sleep studies that are not practical for long-duration monitoring and provide data collected in natural home settings, which are more likely to represent a person's regular sleep patterns. Although actigraphy data do not provide details on a deeper level, e.g., sleep phases, the procedure is very noninvasive and easy to implement. Actigraphy-based sleep data also highlights the need for clinical validation of new wearable-based endpoints.

Other promising wearable technology can be seen in phone/tablet apps. The best-known examples include medication adherence monitoring, medication reminders, and patient engagement. Medication adherence is a big area of concern in multiple therapeutic areas.[13] The reasons behind nonadherence are multifaceted and include socioeconomic factors, access to health care, communication means with healthcare professionals, patients' education, and understanding of the impact of nonadherence to the treatment outcome.[14] Moreover, cell phone apps can provide data to monitor medication adherence and help with timely

Table 2 Novel endpoints: application, benefits, and examples

Application	Benefit	Examples and references
Safety monitoring/patient phenotyping	• Early safety signal, dose and frequency adjustments, discontinuation of certain drug candidates • Better understanding of mechanistic and pharmacological drug profile if combined with PK and wet lab test data	Vital sign, e.g. HR, RR, skin temperature, BP, and actigraphy[37,39]
Novel endpoints	• Mobility as a measure of quality of life • Sleep studies in the home settings for extended periods of time • More sensitive measures than traditional clinical scales in movement disorders	• Actigraphy in Oncology[53] • Actigraphy as a measure sleep in a home settings[54–56] • Gait and tremor in Parkinson's disease[57,58]
Medication adherence monitoring and intervention	• Improved adherence • Informed decisions about dose adjustments • Increased efficiency in postmarket data collection	• Adherence surveys • Drug intake reminder apps • Objective data on drug intake - smart cap bottles
Patient enrollment and retention in clinical trials	• Fewer obstacles to enroll in clinical trials • Reduced burdens for patients to participate • Increased patient outreach	• Remote enrollment and consent apps • Reminder apps about study procedures and clinical trial progress

intervention by medical personnel and caregivers.[15] Medication reminder apps, enhanced by alert personalization and available to both patients and caregivers, were found to improve medication adherence.[16] Additionally, a number of digital technologies were developed to collect objective adherence data with smart-cap bottle and blister pack technologies. However, the effectiveness of these technologies in improving patient adherence has yet to be confirmed in well-powered, controlled studies.[17]

Cell phone apps and web-based interfaces are increasingly used for remote patient enrollment, patient consent, and retention in clinical trials, making the process more convenient and enabling better outreach to remote patients. Clinical trial patient retention may be enhanced by delivering app-mediated reminders, providing information about upcoming visits and operational updates about clinical trial conduct, encouraging compliance, facilitating communication with medical personnel, and making the logistics of participation easier.

The totality and combination of applications can provide a basis for telemedicine and enable partially or completely remote clinical trials, bringing drug development to difficult-to-reach populations. Time and cost could be reduced by decreasing the number of clinic visits and potentially by avoiding use of other expensive medical devices such as telemetry. Time, convenience, and cost savings are big potential benefits of wearable devices, although currently development and adoption costs are militating against such savings. Nonetheless, data delivered by wearable technologies have the potential to improve detection of treatment effects and demonstrate how these effects relate to underlying disease characteristics, improving our understanding of the treatment–response relationship and enhancing personalized medicine.

CHALLENGES

The promising potential of wearable devices has attracted enormous attention, including the start of experiments,[18] and a number of deals between biopharmaceutical, contract research organization (CRO), and device companies have been announced.[19,20] Nevertheless, the major impact expected from digital technologies on biopharmaceutical R&D has not yet materialized.[21] The reasons behind the lack of major transformation include scientific, regulatory, ethical, legal, data management, infrastructure, analysis, and security challenges.

Scientific

Many devices, particularly consumer-grade, are marketed with promises to improve health and wellness with no scientific evidence behind this claim.[5] Properly designed, well-powered studies with a clear statement of a medical problem are required, rather than technology choice-seeking applications.[22] Moreover, drug development and device engineering are historically separate scientific fields. On the one hand, biopharmaceutical R&D scientists are generally not familiar with devices, which creates a barrier for adoption of wearable technologies in drug development clinical trials. On the other hand, device engineers are not conversant with the drug development process and regulatory requirements for drug approvals. The solution would be to bring device engineers into drug development to educate biopharmaceutical R&D and enable adoption of device technologies.

Regulatory

In the US, the drug and device marketing approval paths are separate and the oversight is done by different divisions of the FDA. The majority of wearable devices are classified as Class II devices cleared as 510(k), which requires establishing technical performance in comparison to a predicate (i.e., legally marketed) device that uses a similar engineering solution. The requirement doesn't include establishing an association with a clinical outcome such as a disease condition. This requirement exists only for 510(k) *de novo* devices when there is no predicate device available. Therefore, a device under consideration needs to be tested in a specific population relevant to the device label claims in order to establish an association with a disease condition. If such a 510(k)-cleared device is intended to support an efficacy claim on a drug label, a link between the device readout and an efficacy parameter of

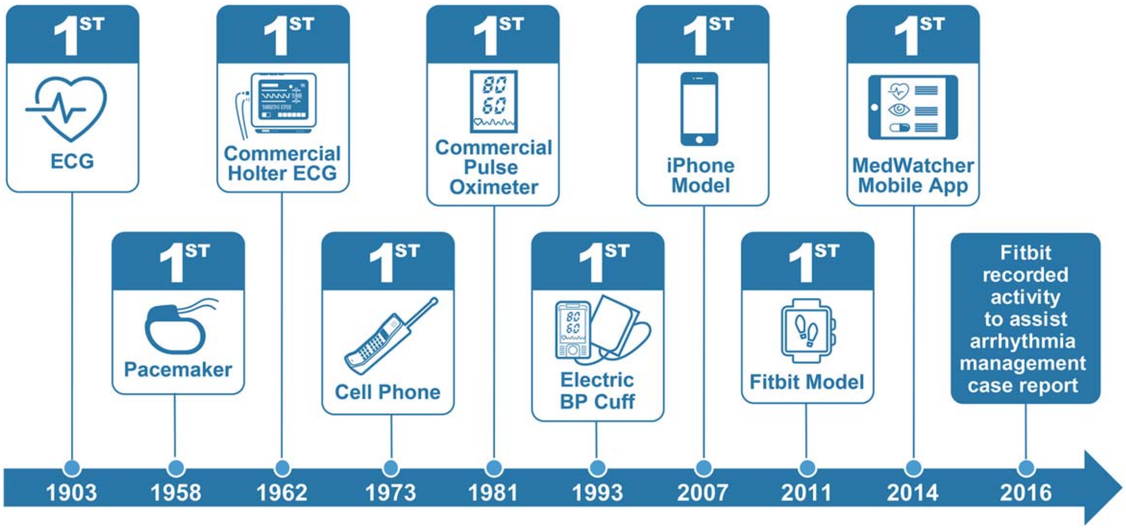


Figure 1 The timeline for market release of technologies enabling wearable device use in healthcare.

interest needs to be established in the context of drug development. It also has to be supplemented by the device analytical performance data indicating that the device is appropriate for an intended use. Additionally, the field is plagued by a lack of shared understanding of methodologies and terminology. A similar issue was successfully overcome in the field of laboratory biomarkers with the widely accepted concept of "fit-for-purpose validation" and well-developed and shared terminology.[23–25] The same approach can be adopted by the wearable device field and several precompetitive initiatives have made significant progress towards achieving this goal.[26,27]

Data infrastructure, processing, analysis, and interpretation

The infrastructure challenges are multifaceted. Drug development clinical teams are not familiar with the massive amounts of 24/7 data to be processed and integrated with the rest of study data. The sensor data structure is very different compared to traditional data collected at predefined timepoints by clinical sites and consists of multiple layers: raw unfiltered data, raw filtered data to eliminate invalid data in accordance with the scoring algorithms, data consisting of the secondary derivatives, and data derived from the secondary derivatives for interpretation. The outstanding questions include: who is the data originator, what constitutes source data, which datasets are required to maintain an audit trail, and what should be reported as a final result. These are the topics of debate by the industry and the regulators, but the recommendations that would help to harmonize the field have not been established. Moreover, the processing and analysis of massive data, as well as result visualization and interpretation, presents a formidable challenge. Machine-learning methods enabling automated data processing and an improved signal recognition were demonstrated to be useful in solving this issue.[11] Additionally, there are no well-developed standards that would help to organize, annotate, and standardize the data and provide data mapping tools to electronic data capture (EDC) databases. The lack of mobile technology data standards is exacerbated by the fact that wearable devices sometimes report variables pertinent to the same phenomenon (e.g., mobility) but use different terminology, and data processing algorithms are not disclosed. The solution should include industry-wide standards for data and terminology, processing principles for similar sets of data, and transparency requirements around data processing algorithms.

Ethical and legal

This category of challenges includes data ownership and sharing, consent requirements, privacy, security, and substantial geographical differences in approaches to addressing these challenges. US and European legislation seems headed in different directions concerning scope, consent, data sharing, and processing.[28] In the US, consumer-grade and medical devices are regulated differently. The data obtained via medical devices are covered by HIPAA and require patient consent for data collection and sharing. On the other hand, the data obtained by consumer-grade devices, although it may contain legitimate health information such as disease condition, lifestyle, biometric, mobility, and behavioral patterns, can be shared in a deidentified, aggregate manner without explicit stipulation concerning who will have access to the data. In the EU, new General Data Protection Regulation (GDPR) regulations do not draw distinctions pertaining to a device type and cover all data generated by wearable devices or apps in the medical context.[29] Additionally, the EU requires clearly defined purposes for data use, consent for data reuse and sharing, and allows patients to withdraw their consent at any time.

Data security

In the practical consideration of privacy, security, and compliance, it can be helpful to separate compliance from privacy and security, as compliance tends to be retrospective in nature, but ensuring privacy and security must be proactive and forward-looking.[30] Much has been written about general and advanced privacy and security with respect to medical data and devices.[31,32]

Table 3 New families of privacy and security controls

Control family	Key example controls
Access Control	Account Management, Access Enforcement, Information Flow Enforcement
Awareness & Training	General Awareness Training, Role-based Training
Audit & Accountability	Audit Event Management, Audit Review Analysis & Reporting
Assessment, Authorization & Monitoring	Annual Assessments, Assessment Guidelines, Independent Assessment
Configuration Management	Baseline Configuration, Configuration Change Control
Contingency Planning	Contingency Plan, Contingency Training, Contingency Plan Testing
Identification and Authentication	User Management, Device Management, Management of Unique Identifiers
Individual Participation	Individual Consent, Redress, Access, Privacy Notices and ACT Statements
Incident Response	Incident Response Policies & Procedures, Training, Testing, Handling, Monitoring
Maintenance	Controlled Maintenance, Maintenance Tools, Personnel, Local & Non-local
Media Protection	Media Access, Media Marking, Storage & Transport, Sanitization and Use
Privacy Authorization	Authority to Collect, Purpose and Sharing
Physical and Environmental Protection	Physical Access Authorization & Control, Monitoring
Planning	Security & Privacy Plans, Updates, Rules of Behavior, Impact Assessments
Program Management	Program Plan, Roles, Resources, Inventory, Architecture and Performance
Personnel Security	Personnel Screening, Risk Designation, Transfer and Termination
Risk Assessment	Security Categorization, Assessment and Vulnerability Scanning
System and Services Acquisition	Resource Allocation, Systems Lifecycle, Acquisition and Documentation
System and Communications Protection	Application Partitioning, Security Function Isolation, Boundary Protection
System and Information Integrity	Flaw Remediation, Malicious Code Protection, Monitoring, Alerts & Advisories

Fortunately, guidance recently released by the US National Institute of Standards and Technology (NIST) details new families of privacy and security controls that can be used as the basis of design and audit, as shown in **Table 3**.[33] Focusing specifically on wearable sensors and devices, the guidance deems it essential that all personally identifiable information (PII) and all personal health information (PHI) must be protected, and that the devices themselves be protected from any form of outside interference, whether accidental or malicious. The predominant generic issues include: the device security of any mobile devices, tablets, and cell phones that are used to collect, store, or transmit information; the potential complications of commingling study sponsor-collected PHI on the personally owned device of a research study participant; secure data transmission and receipt; secure account management; data encryption; data blinding; and data backup and device fidelity. It is essential to understand that these concepts are generic by necessity. Specific solutions will always be required depending on the exact device model, the specific device operating system, the intended method of network connectivity, the intended data capture and processing strategy, and many other variables that will be study-specific. Using several potential methods of network connectivity as examples, **Figure 2** illustrates just some of the most common and potential cyber threat vectors that exist for the three primary types of device connection: Bluetooth, WiFi, and cellular as described by the NIST. The take-home message here is simply that cyber security is increasingly complex, but also well understood and manageable. Success requires a thorough benefit–risk assessment by experts just like any other medical intervention.

SPECIAL CONSIDERATIONS FOR CLINICAL TRIALS

The application of wearable devices to clinical trials and drug development is in a similar state to that of biomarkers in the early 2000s. At that time, considerable confusion abounded regarding the appropriate use and validation of biomarkers. Tremendous efforts were applied to biomarker activities resulting in refined approaches, particularly the definition and framework for analytical validation, clinical validation, and qualification.[24,25,34] Considerations for the use of wearable devices in a clinical trial should include primarily scientific aspects with a patient-centric approach in mind (**Figure 3**). However, operational aspects, such as patient and site personnel training, device acceptability to patients and patient compliance, data reporting, and transfer and management are critical for obtaining valid and interpretable data. In addition, there is a critical role for validation, both analytical and clinical, in the utility of wearable devices (**Figure 3**).

Scientific considerations

The scientific approach should start with a health condition or an aspect of health important to patients that has not been addressed to a satisfactory level by current standards of disease management care. Once it is defined, a scientific hypothesis

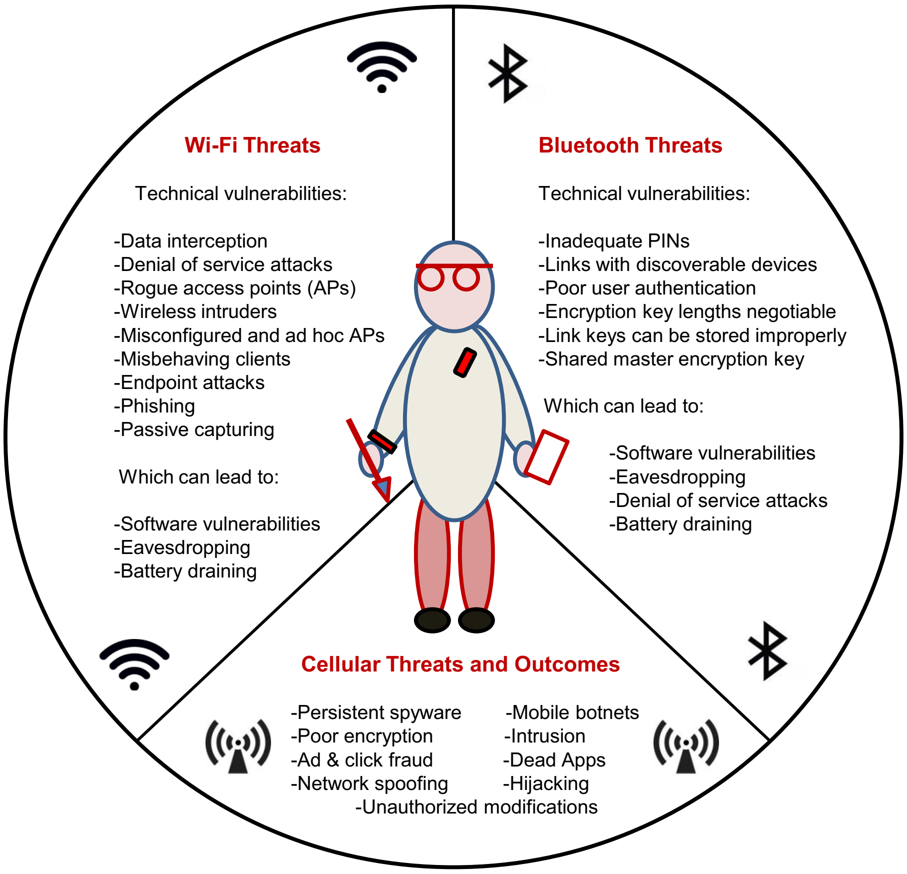


Figure 2 Most common and potential cyber threat vectors.

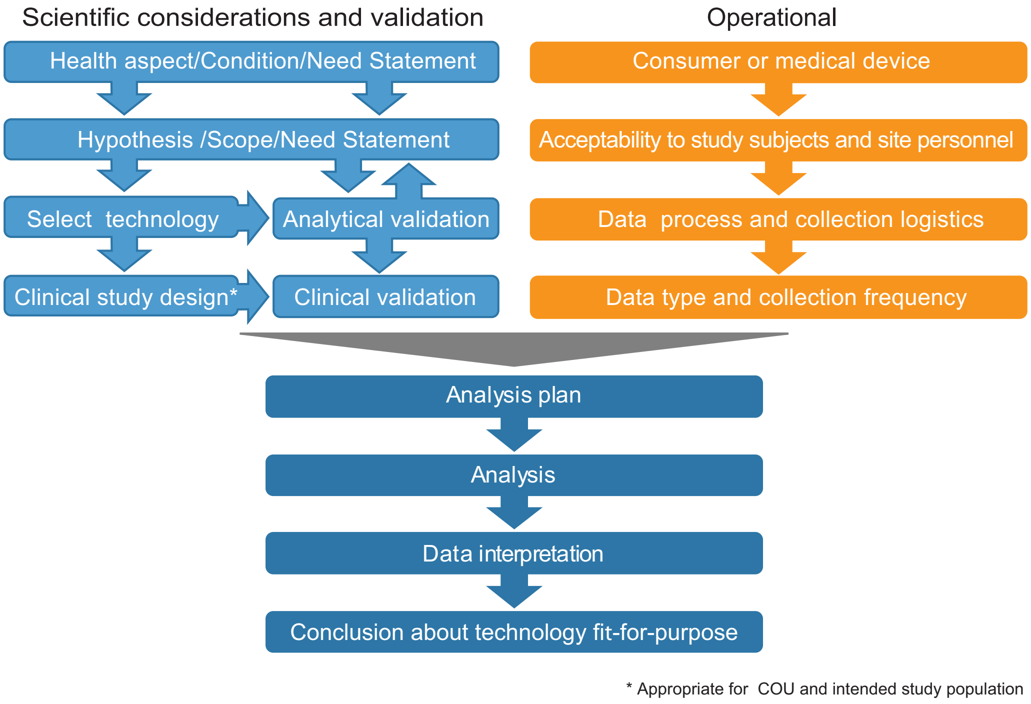


Figure 3 Scientific, validation, and operational considerations for wearable device implementation in clinical trials.

should be formulated to define the scope of an experiment to be conducted. For example, current assessment of morning stiffness and sleep in rheumatoid arthritis (RA) patients is based on self-reports. The standard tools of data collection include patient self-reports during the doctor's office visits based on memory recalls and patient diaries. Having objective data reflective of these health parameters can be very informative for patient care management including management of adverse events, medications, and dose adjustments. Once the scope is defined, the next step would entail finding a suitable technology to capture the data of interest. In the case of RA, study results indicate that wrist-worn actigraphy devices can differentiate RA patients from healthy controls and can provide useful information about mobility in the context of drug treatment.[35,36] The hypothesis should be tested as one of the objectives in a clinical study. The hierarchical order of an objective of interest, e.g., primary, secondary, or exploratory, will depend on the strength of evidence supporting the hypothesis. The testing can be achieved in an observational or an interventional study. An observational study would be appropriate when no data or limited data about the link between a disease/health aspect and device-derived readouts exist. An interventional study is more appropriate if the goal is to establish a process for wearable data collection in the context of drug treatment and to support efficacy claims or guide treatment decisions. Additionally, a device under consideration should be appropriate for a given study population.

The general validation framework includes a need statement, context of use (COU), analytical validation, clinical validation, and qualification, if necessary for a regulatory purpose (**Figure 3**). A need statement is a concise and coherent description of the knowledge gap or drug development need (e.g., improved safety monitoring) and interfaces with the scientific aspect of the wearable. The COU, which also interfaces with the scientific aspects of a wearable, is a concise description of how a wearable is intended to be used in drug development. With a particular COU, analytical validation establishes if the device performance characteristics are acceptable. Analytical validation or technical performance established for purposes of 510(k) clearance would entail establishing device performance parameters under conditions as close as possible to real-life use. This goal can be achieved by comparing device performance to a traditional tool for collecting the data if available,[37] or another device with well-established performance.[7] Some of the analytical validation parameters may be already established during device calibration done by the device manufacturer and may include important information such as conformity to a gold standard and sensor precision under various testing conditions, but may require an independent validation in the COU. Understanding performance characteristics is necessary for deciding if a device can measure what is needed in a particular COU. If a medical device is under consideration, device performance is established for the purposes of device clearance. However, it may not be appropriate in an intended study population or COU. For instance, if a device has been tested in normal healthy volunteers but is intended for future use in a particular disease, both the hardware and the software performance need to be established in the context of disease to render the device use as "fit-for-purpose." The lack of testing in the intended study population may result in inappropriate data processing and even loss of the data.[38] Also, with a particular COU, clinical validation establishes that the wearable device acceptably identifies, measures, or predicts the concept of interest. Clinical validation includes establishing an association with a specific disease condition to make sure that the data are interpretable and provide useful information for patient care management.[37,39] Both analytical and clinical validation can be done in dedicated device evaluation studies or can be incorporated as one of the endpoints in drug development clinical trials. In the first scenario, multiple devices may be evaluated with appropriate controls embedded in the study, e.g., drugs modulating blood pressure for blood pressure monitoring devices. The disadvantage of this study type is the lack of assessment of device impact on other study procedures routinely performed in drug development, such as frequent blood draws for pharmacokinetics (PK) or imaging procedures. In the second scenario, adding devices to drug development clinical studies as exploratory endpoints provides an opportunity to establish tolerability and acceptability of the device by the study participants and sites in the context of other study procedures. These considerations are a starting place, but require input from stakeholders and a further discussion between the biopharmaceutical industry, device manufacturers, and regulators. It is conceivable that qualification will ultimately be necessary for wearables, similar to surrogate endpoints. Based on a formal regulatory process, it is a conclusion within the stated COU that a drug development tool can be relied upon to have a specific interpretation and application in medical product development and regulatory review. We are not aware of any instances of wearable use requiring this level of scrutiny.

Device choice and logistical considerations

Both consumer and medical-grade devices can be considered for drug development clinical trials. Medical-grade devices require less work prior to inclusion in clinical trials, as their performance may be established for the purpose of a clearance or approval process and the information is available on the device label. That said, consideration of the intended COU is necessary prior to application. However, consumer-grade devices may not yet have established performance, and device analytical and clinical validation studies are needed to ascertain that a device of interest is fit-for-purpose. The raw and derivative data availability from the device should be considered carefully, as often only secondary derivatives and summary data are available; this may provide an incomplete audit trail. Device acceptability by study subjects is critical to successful implementation. Device technical characteristics such as size, convenience to wear, battery life, and impact on daily life activities should be considered carefully. These characteristics may require patients' input prior to study initiation to ensure successful adoption of a technology. If technology acceptance by users is not known before the study start, a small pilot study may be warranted to obtain these data, as acceptance will have a major impact on patients' compliance. We found that having hands-on experience by clinical scientists directly involved in clinical study design and conduct is highly beneficial. It accelerates device implementation by clinical teams and allows scientists to rule out early devices that are unlikely to be easily accepted by study participants and may not provide interpretable data. Devices are usually administered by clinical site personnel, trained to pass information to the subjects and be available to help if study subjects are experiencing difficulties. In addition to subject and site personnel training, the data process flow should be mapped before the study start to evaluate the impact of data flow on study participants and other clinical trial procedures. Examples include requiring a cell phone for data synchronization, specific phone models compatible with apps, translations if needed, frequency of data synchronization, and specific computer models for device docking. Compliance of study subjects contributing data should be monitored. Interventions such as reminders to the subjects should be implemented to improve compliance if it falls below a certain threshold.

Decisions need to be made up front about the timing of data processing into secondary derivatives and data review. If data need to be reviewed in near real-time, the data processing, analysis, and visualization need to be established and tested before the study start. Follow-up procedures, if warranted, need to be determined as a part of a clinical study protocol. Retrospective data processing and analysis are more suitable for exploratory endpoints, as they provide more room for experimentation with raw data processing and visualization options, and can be done in an iterative manner. Data use should be clearly defined in the study protocol and it should be stipulated whether such use has any impact on patient care or any other study procedures. In addition, decisions would have to be made on how to handle subjects who may have an allergic or any other adverse reaction to the components of wearable devices. Depending on the intended use of the data, subjects with known adverse reactions to the components of a device may be excluded from the study or allowed to participate in other study procedures; this is appropriate if consent to the wearable device portion of the study is optional and the lack of participation does not have a major impact on overall study data integrity.

Considerations for including devices in the clinical studies are multidimensional (**Figure 3**). R&D and healthcare organizations have a number of hurdles to overcome to make wearable technology implementation a routine procedure. Further development of analytical and clinical validation methodologies and the wide adoption of devices according to the fit-for-purpose principle will remain critical for future success.

WHAT THERAPEUTIC AREAS ARE MOST APPROPRIATE AND WHY?

In theory, wearables can be used broadly across therapeutic areas for deep phenotyping, detection and interpretation of adverse events, assessment of quality of life, and measurement of efficacy. Wearable and digital approaches could provide signal detection for conditions such as depression by measuring increases in sleep or decrease in activity, signs associated with depression. For example, wearables were recently suggested to be helpful in the detection of early signs of Lyme disease.[7] Any therapeutic intervention that may impact quality of life could benefit from measurement of movement or in some cases where a patient diary is required. One example is a collaboration between PatientsLikeMe and Biogen to better characterize multiple sclerosis patients,[40] where activity and mobility are clearly tied to quality of life. Some therapeutic areas may not require use of a wearable, but rather simple mobile phone applications such as Apple's ResearchKit.

Since many wearable devices can readily measure heart rate as well as blood pressure, the cardiovascular therapeutic area is a major focus for use of wearable devices. Cardiac monitoring in both healthy individuals and specific disease populations allows monitoring for cardiac events 24/7 and enables better-informed care. Cardiovascular disease areas in which wearable devices have been or could be used include congestive heart failure, hypertension, and dysrhythmias. For example, The Zio Patch (iRhythm Technologies, San Francisco, CA) is a single-lead electrocardiographic, continuously recording ambulatory adhesive patch, recently approved by the FDA. In a recent study, the device's 14-day monitoring of beat-to-beat cardiac rhythm had a 57% greater diagnostic yield than the standard 24-h Holter monitoring.[41]

Neuroscience uses of wearable devices are manifold, including the monitoring of sleep, cognition, and movement disorders. Wearable devices commonly measure selected sleep parameters and activity. To assess patients for obstructive sleep apnea outside the laboratory setting, use of medical devices has been steadily increasing.[42] IBM Watson Health and the American Sleep Apnea Association have launched the SleepHealth app to conduct a study identifying connections between sleep habits and health outcomes. This app will record movement and heart rate during sleep and track connection between sleep quality and daytime activities, alertness, productivity, general health, and medical conditions. It will amass the largest collection of sleep data to date. Parkinson's disease is another area that has shown promising results and insight via wearables and machine-learning techniques. The sensors in wearable devices can be paired with mobile phone apps to measure symptoms such as tremor, balance, gait, memory, and some vocal characteristics.

There are examples of wearable use in respiratory diseases, immunology, and rheumatology. For example, GlaxoSmithKline (GSK Philadelphia, PA) (in collaboration with Medidata and POSSIBLE Mobile) are starting an RA trial called PARADE8.[43] It is expected to evaluate 300 patients through a mobile application that tracks common RA symptoms such as joint pain and fatigue, and gathers these data through a mix of surveys and sensor-enabled tests (e.g., recording motion through wrist exercises). This trial is gathering data on the everyday lives of people with RA to gain insight and learn more about the condition. WristOx2 by Nonin Medical (Plymouth, MN) is a pulse oximeter that monitors and measures heart rate and blood oxygen levels, and is targeted towards people who have asthma and are at risk of chronic pulmonary obstruction disease. In 2014, Novartis (Hanover, NJ) launched an observational trial with Qualcomm Life (San Diego, CA) collecting biometric data from chronic lung disease patients in their homes using smartphones connected to Qualcomm's cloud-based 2net Platform.[19]

Another therapeutic area addressed by wearable devices is metabolic disorders, including diabetes and obesity. A recent systematic review of mHealth (Mobile Health)-related studies on diabetes and obesity treatment and management found that over half of the reported positive effects of interventions based on primary outcomes.[44] Accurate glucose monitoring is something currently in development, as it is not readily available in smartwatches, but several companies are developing prototypes. For example, Dexcom (San Diego, CA) have developed a continuous glucose monitoring application that uses a dermal implant with a probe capable of monitoring blood glucose every 5 min, eliminating the need for finger sticks. The Freestyle Libre Flash Glucose Monitoring System by Abbott (Abbott Park, IL) is a wearable skin sensor that has received regulatory approval. Recently, a pilot study of a patient-centered, smartphone-based, diabetes care system found that a 12-week application of the system to patients with inadequately controlled type 2 diabetes resulted in a significant HbA1c reduction.[44]

PROGRESS IN CLINICAL TRIALS TO DATE

Recent reviews of wearable monitoring systems have shown that the key implementation challenges are patient and provider engagement, connectivity and device communication, and clinical validation.[45] Per earlier discussion, we have emphasized the importance of rigorous clinical investigation in a stepwise manner where devices are tested in successively less controlled circumstances prior to full investigation in patients' homes. Several wearable devices (ViSi Mobile and HealthPatch) designed for continuous vital signs monitoring were studied in a general hospital ward and compared with vital signs measurements by nurses.[37] The study showed generally promising results, including patient and clinician experiences, but the number and types of artifacts/errors demonstrated the need for significant improvement before equivalence with traditionally used measures can be achieved.

We conducted similar experiments in interventional clinical studies. Our goal was to evaluate 510(k)-cleared wearable devices in the context of drug development clinical trials and ascertain whether devices of interest are fit-for-purpose for vital sign and cardiac rhythm monitoring; this was done in normal healthy volunteers for the purposes of deep phenotyping and expanded safety monitoring. Our experimental design included establishing both analytical and clinical validation by comparing device performance with conventional measures done at the sites, and testing devices in experiments with certain clinical positive controls such as an increased heart rate after certain drug administration. Additionally, we queried important operational parameters such as acceptability by the study subjects and site personnel, and we collected data on subjects' compliance and gained institutional experience with logistics of implementation of wearable technologies. Our data indicate that the technologies are acceptable to the study subjects; however, compliance may decrease when subjects use the devices at home. The feedback from the site personnel indicated high rates of adoption and eagerness of use with a clear need for dedicated technical training and hands-on experience before the launch. Analytical validation experiments demonstrated variable concordance with traditional measures, depending on the variable of interest. Higher concordance was observed with the data collected by another device vs. data collected manually. Consistent with the findings reported by other groups,[37,39] these devices have a propensity to generate a number of artifacts that should be reduced before further broad implementation of technology for safety monitoring. Additionally, we found that a combination of vital sign monitoring with actigraphy readouts, such as mobility counts, facilitates interpretation of vital sign values not collected at the resting state. Overall, our results demonstrated feasibility of collection of vital sign data using wearable devices; however, implementation of such devices for safety monitoring should proceed with caution and should include mandatory verification that a technology of interest is fit-for-purpose.

Looking at ours and others' studies, we see a common theme of great progress and promise but also of technologies that are not quite ready for prime time. Looking at other sensor/device domains, we see similar themes. For example, in a recent in-clinic validation study of a cuffless device for measuring blood pressure, the device demonstrated less than a 5-mmHg variance from conventional measurement in 46% of the study population, but 23% of the originally recruited subjects had to be excluded upfront due to device calibration error.[39]

As previously discussed, there is a sharp difference in measurement accuracy and data/device fidelity between clinical and consumer-grade motion detecting devices but we are hopeful that this gap will eventually disappear. Within the clinical setting, motion detection sensors are being successfully used in increasingly complex observation and analysis scenarios. In one recent motion measurement study of early Parkinson's disease patients, timed "Up and Go" tasks were measured with far greater than 90% sensitivity, but this level of clinical-grade motion measurement required the wearing of special suits that had 17 sensors per body segment.[46]

CONCLUSION

Wearable technologies are promising and have the potential to fundamentally change healthcare and drug development by changing the means of collecting, processing, and visualizing health data. Potential applications are diverse, have utility in multiple therapeutic areas, and are likely to evolve rapidly. The ultimate goal should be a better understanding of disease variability, responses to treatment along with a reduction of healthcare costs, and increasing efficiency in conducting clinical trials. Additionally, adopting new ways of remote data collection can bring new treatments and care management to all patients in need. Challenges presented by adoption of wearable technologies are not insignificant. The scientific community would benefit from frequent information exchange to share the results and learning experiences; this would facilitate the development and adoption of best practices for technology implementation, data collection, analysis, and interpretation. Currently, the field is full of enthusiasm, but more data are needed from rigorously designed studies to displace the hype and adopt scientific methodologies to generate and test scientific hypotheses. Further dialog between the biopharmaceutical industry and device manufacturers to develop methodological approaches and shared understanding of the experiments is required to fulfill the requirements of analytical and clinical validation. This conversation would constitute a major step forward facilitating the adoption of wearable technologies in clinical trials.

Definitions

Analytical validation. Establishing that the performance characteristics of a test, tool, or instrument are acceptable in terms of its sensitivity, specificity, accuracy, precision, and other relevant performance characteristics using a specified technical protocol (which may include specimen collection, handling and storage procedures). This is validation of the test's, tool's, or instrument's technical performance, but is not validation of the item's usefulness.[24]

Clinical validation. Establishing that the test, tool, or instrument acceptably identifies, measures, or predicts the concept of interest.[24]

Consumer devices are devices marketed directly to individuals. An example would be a Fitbit wrist-worn device or an iPhone. The individuals are responsible for managing their devices, including data backup and the decisions around software upgrades. Some of these devices fall under the FDA definition of general wellness products and are considered low-risk devices.[47] There is no requirement to establish device performance before the release to the market.

Context of use (COU). A statement that fully and clearly describes the way the medical product development tool is to be used and the medical product development-related purpose of the use.[21]

Medical devices are defined by the FDA as is "an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component part, or accessory which is . . . intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals."[48]

Medical devices have to be approved/cleared by the FDA before they can be released to the market. The approval/clearance path depends on the intended use of the device and also upon indications for use. Devices are classified into Class I, II, and III devices based on the risk the device poses to the patient and/or the user.[49] Depending on the device classification, a Premarket Approval (PMA) or 510(k) clearance is required before release to the market.[50] The FDA stipulates 510(k) as: "a premarket submission made to FDA to demonstrate that the device to be marketed is at least as safe and effective, that is, substantially equivalent, to a legally marketed device. Submitters must compare their device to one or more similar legally marketed devices and make and support their substantial equivalency claims."[51] Medical devices are also a subject to HIPAA security and privacy rules.

Medical need/necessity. The AMA defines medical necessity as: "Health care services or products that a prudent physician would provide to a patient for the purpose of preventing, diagnosing or treating an illness, injury, disease or its symptoms in a manner that is: (a) in accordance with generally accepted standards of medical practice; (b) clinically appropriate in terms of type, frequency, extent, site, and duration; and (c) not primarily for the economic benefit of the health plans and purchasers or for the convenience of the patient, treating physician, or other health care provider."[52]

Wearable technologies are sensors and/or software applications on smartphones and tablets that can collect health-related data remotely.

ACKNOWLEDGMENTS

The authors thank Mike Mangiapane for help with editing the article, and Justo Torres for help with Figures 1 and 3 graphical designs.

CONFLICT OF INTEREST

Elena Izmailova, Eric Perakslis, and John Wagner are employed by Takeda Pharmaceuticals International Co.

AUTHOR CONTRIBUTIONS

Concept and design: Izmailova, Perakslis, Wagner. Drafting of the article: Izmailova, Perakslis, Wagner. Editing and final revision of the article; Izmailova, Perakslis, Wagner.

---

1. Anema, A. *et al*. Digital surveillance for enhanced detection and response to outbreaks. *Lancet Infect. Dis.* **14**, 1035–1037 (2014).
2. Brownstein, J.S., Freifeld, C.C. & Madoff, L.C. Digital disease detection—Harnessing the Web for public health surveillance. *N. Engl. J. Med.* **360**, 2153–2155, 7 (2009).
3. Rudner, J., McDougall, C., Sailam, V., Smith, M. & Sacchetti, A. Interrogation of patient Smartphone activity tracker to assist arrhythmia management. *Ann. Emerg. Med.* **68**, 292–294 (2016).
4. Majumder, S., Mondal, T. & Deen, M.J. Wearable sensors for remote health monitoring. *Sensors (Basel)* **17**, (2017).
5. Piwek, L., Ellis, D.A., Andrews, S. & Joinson, A. The rise of consumer health wearables: promises and barriers. *PLoS Med.* **13**, e1001953 (2016).
6. Schwartz, J.E. *et al*. Clinic blood pressure underestimates ambulatory blood pressure in an untreated employer-based US population: results from the masked hypertension study. *Circulation* **134**, 1794–1807 (2016).
7. Li, X. *et al*. Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. *PLoS Biol.* **15**, e2001402 (2017).
8. SPRINT challenge <https://challenge.nejm.org/pages/winners>
9. Stone, A.A., Shiffman, S., Schwartz, J.E., Broderick, J.E. & Hufford, M.R. Patient compliance with paper and electronic diaries. *Controll. Clin. Trials* **24**, 182–199 (2003).
10. Lee, S.J., Kavanaugh, A. & Lenert, L. Electronic and computer-generated patient questionnaires in standard care. *Best Pract. Res. Clin. Rheumatol.* **21**, 637–647 (2007).
11. Coons, S.J., Eremenco, S., Lundy, J.J., O'Donohoe, P., O'Gorman, H. & Malizia, W. Capturing patient-reported outcome (PRO) data electronically: the past, present, and promise of ePRO measurement in clinical trials. *Patient* **8**, 301–309 (2015).
12. Martin, J.L. & Hakim, A.D. Wrist actigraphy. *Chest* **139**, 1514–1527 (2011).
13. Cutler, D.M. & Everett, W. Thinking outside the pillbox—medication adherence as a priority for health care reform. *N. Engl. J. Med.* **362**, 1553–1555 (2010).
14. Brown, M.T. & Bussell, J.K. Medication adherence: WHO cares? *Mayo Clin. Proc.* **86**, 304–314 (2011).
15. Basch, E. *et al*. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. *JAMA* **318**, 197–198 (2017).

16. Thakkar, J. *et al*. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. *JAMA Intern. Med.* **176**, 340–349 (2016).
17. Volpp, K.G. *et al*. Effect of electronic reminders, financial incentives, and social support on outcomes after myocardial infarction: The HeartStrong randomized clinical trial. *JAMA Intern. Med.* **177**, 1093–1101 (2017).
18. Steinhubl, S.R. *et al*. Rationale and design of a home-based trial using wearable sensors to detect asymptomatic atrial fibrillation in a targeted population: The mHealth Screening To Prevent Strokes (mSToPS) trial. *Am. Heart J.* **175**, 77–85 (2016).
19. Novartis and PPD partnership news. Available at: <http://www.fiercebiotech.com/medical-devices/novartis-ppd-conduct-clinical-pilot-study-using-medical-device-connectivity-via-a>
20. GSK partnership news. Available at: <http://www.mobihealthnews.com/38325/glaxosmithkline-is-testing-mobile-health-sensors-for-clinical-trials>
21. Perakslis, E.D. Strategies for delivering value from digital technology transformation. *Nat. Rev. Drug Discov.* **16**, 71–72 (2017).
22. Developing Novel Endpoints Generated by Mobile Technology for Use in Clinical Trials. Available at: <https://www.ctti-clinicaltrials.org/projects/novel-endpoints>
23. Lee, J.W. *et al*. Fit-for-purpose method development and validation for successful biomarker measurement. *Pharmaceut. Res.* **23**, 312–428 (2006).
24. Group, F.-N.B.W. In: BEST (Biomarkers, EndpointS, and other Tools) Resource (Food and Drug Administration (US)/National Institutes of Health (US), Maryland, 2016).
25. Amur, S., LaVange, L., Zineh, I., Buckman-Garner, S. & Woodcock, J. Biomarker qualification: toward a multiple stakeholder framework for biomarker development, regulatory acceptance, and utilization. *Clin. Pharmacol. Ther.* **98**, 34–46 (2015).
26. Scientific and Technological Issues Surrounding the Use of Mobile Devices in Clinical Trials. Available at: <https://www.ctti-clinicaltrials.org/projects/mobile-devices>
27. RADAR-CNS Remote Assessment of Disease and Relapse in Central Nervous System Disorders. Available at: <https://www.imi.europa.eu/content/radar-cns>
28. A comprehensive report concerning US and EU privacy matters can be found at: <https://www.democraticmedia.org/CDD-Wearable-Devices-Big-Data-Report>
29. CIPL GDPR Interpretation and Implementation Project. <https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_gdpr_project_risk_white_paper_21_december_2016.pdf>
30. Perakslis, E.D. & Stanley, M. A cybersecurity primer for translational research. *Sci. Transl. Med.* **8**, 322ps2 (2016).
31. Perakslis, E.D. Cybersecurity in health care. *N. Engl. J. Med.* **371**, 395–397 (2014).
32. Martin, G., Martin, P., Hankin, C., Darzi, A. & Kinross, J. Cybersecurity and healthcare: how safe are we? *BMJ (Clin. Res. Ed.)* **358**, j3179 (2017).
33. NIST Special Publication 800-53 Revision 5. <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwin-de0yrnWAhULqFQKHboMDPgQFgg-MAA-url=https%3A%2F%2Fcsrc.nist.gov%2Fpublications%2Fdrafts%2F800-53%2Fsp800-53r5-draft.pdf&usg=AFQjCNGdIfjvpf-cqYzqa7ddeFmuzLoIkw>
34. Leptak, C. *et al*. Evidentiary criteria framework to advance biomarker qualification. *Sci. Transl. Med.* (2017).
35. Clarke, L.L., Wilson, S. & Kirwan, J.R. Using actigraphy to measure sleep patterns in rheumatoid arthritis: a pilot study in patients taking night-time prednisone. *Musculoskeletal Care* **11**, 179–185 (2013).
36. Hashimoto, T. *et al*. Physical activity of elderly patients with rheumatoid arthritis and healthy individuals: an actigraphy study. *BioPsychoSocial Med.* **9**, 19 (2015).
37. Weenk, M. *et al*. Continuous monitoring of vital signs using wearable devices on the general ward: pilot study. *JMIR mHealth uHealth* **5**, e91 (2017).
38. Hernandez-Silveira, M. *et al*. Assessment of the feasibility of an ultra-low power, wireless digital patch for the continuous ambulatory monitoring of vital signs. *BMJ Open* **5**, e006606 (2015).
39. Schoot, T.S., Weenk, M., van de Belt, T.H., Engelen, L.J., van Goor, H. & Bredie, S.J. A new cuffless device for measuring blood pressure: a real-life validation study. *J. Med. Internet Res.* **18**, e85 (2016).
40. PatientsLikeMe Study Monitors Walking Activity in People with MS. <http://news.patientslikeme.com/press-release/patientslikeme-study-monitors-walking-activity-people-ms>
41. Barrett, P.M. *et al*. Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring. *Am. J. Med.* **127**, 95.e11–17 (2014).
42. Kirsch, D.B. PRO: sliding into home: portable sleep testing is effective for diagnosis of obstructive sleep apnea. *J. Clin. Sleep Med.* **9**, 5–7 (2013).
43. GSK starts RA trial on Apple ResearchKit, advancing industry use of iPhone-run studies. <http://www.fiercebiotech.com/it/gsk-starts-ra-trial-apple-researchkit-advancing-industry-use-iphone-run-studies>
44. Wang, Y., Xue, H., Huang, Y., Huang, L. & Zhang, D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. *Adv. Nutr. (Bethesda, MD)* **8**, 449–462 (2017).
45. Baig, M.M., GholamHosseini, H., Moqeem, A.A., Mirza, F. & Linden, M. A systematic review of wearable patient monitoring systems — current challenges and opportunities for clinical adoption. *J. Med. Syst.* **41**, 115 (2017).
46. Nguyen, H.P. *et al*. Auto detection and segmentation of physical activities during a Timed-Up-and-Go (TUG) task in healthy older adults using multiple inertial sensors. *J. Neuroeng. Rehab.* **12**, 36 (2015).
47. General Wellness: Policy for Low Risk Devices Guidance for Industry and Food and Drug Administration Staff. <https://www.fda.gov/downloads/MedicalDevices/NewsEvents/WorkshopsConferences/UCM519701.pdf>
48. FDA medical device. <https://www.fda.gov/aboutfda/transparency/basics/ucm211822.htm>
49. FDA device classification. <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/default.htm>
50. PMA and 510K. <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/default.htm#510k>
51. 510(K) premarket notification. <https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketNotification510k/default.htm>
52. Medical necessity definition by AMA. <http://www.nationalacademies.org/hmd/~/media/8D03963CAEB24450947C1AEC0CAECD85.ashx>
53. Beg, M.S., Gupta, A., Stewart, T. & Rethorst, C.D. Promise of wearable physical activity monitors in oncology practice. *J. Oncol. Pract.* **13**, 82–89 (2017).
54. Sarkanen, T.O., Alakuijala, A.P.E., Dauvilliers, Y.A. & Partinen, M.M. Incidence of narcolepsy after H1N1 influenza and vaccinations: Systematic review and meta-analysis. *Sleep Med. Rev.* (2017) [Epub ahead of print].
55. Filardi, M., Pizza, F., Martoni, M., Vandi, S., Plazzi, G. & Natale, V. Actigraphic assessment of sleep/wake behavior in central disorders of hypersomnolence. *Sleep Med.* **16**, 126–130 (2015).
56. Alakuijala, A., Sarkanen, T. & Partinen, M. Polysomnographic and actigraphic characteristics of patients with H1N1-vaccine-related and sporadic narcolepsy. *Sleep Med.* **16**, 39–44 (2015).
57. Silva de Lima, A.L. *et al*. Large-scale wearable sensor deployment in Parkinson's patients: The Parkinson@Home Study Protocol. *JMIR Res. Protoc.* **5**, e172 (2016).
58. Peterson, D.S. & Horak, F.B. The effect of levodopa on improvements in protective stepping in people with Parkinson's disease. *Neurorehab. Neural Repair* **30**, 931–940 (2016).

PERSPECTIVE

Data Science Approaches for Effective Use of Mobile Device–Based Collection of Real-World Data

Larsson Omberg[1,*], Elias Chaibub Neto[1,*] and Lara M. Mangravite[1,*]

The use of mobile health for monitoring disease outside of the clinic has opened new opportunities for drug testing and monitoring. In particular, these tools are providing new experimental designs for collection of real-world data. These technologies and queries, although promising, require the application of analytical methods that can accommodate the uncontrolled, unmonitored, individualized, and, often, near continuous data streams. Here, we discuss opportunities and ramifications on analytical considerations.

Mobile health, that is, the evaluation of health outside of the clinic using wearables and smartphones, and, more broadly, the collection of real-world evidence,[1] provide opportunities to advance multiple goals for monitoring drug response, including the monitoring of efficacy through digital biomarkers that can be used as primary end points for drug efficacy, monitoring of patient-reported outcomes and/or quality of life measures, and of toxicities and/or response to long-term exposures. Although digital end points are of interest to regulatory agencies such as the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) and are starting to be integrated as primary end points into clinical trials,[2] adoption is slow. In part, this is due to difficulties in quantifying the accuracy of measures when they are collected in an unmonitored manner and in an uncontrolled setting. Indeed, the ability to develop robust measures that are reliably accurate requires both an expanded validation plan designed to pressure-test the measure across a range of conditions and a good understanding of the impact that variations in daily living can have on data collection. Because interpretation of mobile health data involves the processing and analysis of high dimensional, longitudinal sensor data collected in continuous or near continuous data streams, it requires the use of statistical approaches that account for repeat measures as well as extensive use of signal processing and/or machine-learning techniques. These approaches provide opportunity for sensitive, individualized monitoring of drug responses. Here, we provide a short introduction to the importance of appropriate usage of analytic and machine-learning techniques for the interpretation of mobile health data (see also refs. 3–5). This includes a description of the types of experiments and data that can be collected using mobile health and some examples from the literature that highlight important analytical considerations. Although these observations are relevant to any device that is collecting sensor data in a continuous or near continuous manner, we exemplify these issues using our own experience with the development and analysis of smartphone-based measures.

FROM SENSORS TO MACHINE LEARNING AND SOME CONSEQUENCES

Modern smartphones have opened a wealth of possibilities to extend electronic health monitoring for two reasons: (i) The always connected nature and computational power of smartphones allows for rapid data collection and (ii) the large number of embedded sensors allows for multimodal data collection.[6] A typical phone has sensors that can measure acceleration, rotation rates, magnetic fields, sound levels, record audio and video, and record time and touch through the screen, among other capabilities. Sensor-based data collection performed in the context of protocols, or tasks designed to capture disease relevant behavior, can be used to generate hundreds of phenotypic measurements, including those that mimic evaluations typically performed in the clinic (e.g., sit to stand test for mobility or blood pressure measurement). They can also be used to passively collect measurements during daily activities (e.g., mobility analysis performed while an individual is walking). In

---

[1] Sage Bionetworks, Seattle, Washington, USA. *Correspondence: Larsson Omberg (larsson.omberg@sagebionetworks.org), Elias Chaibub Neto (elias.chaibub.neto@sagebionetworks.org), and Lara M. Mangravite (lara.mangravite@sagebionetworks.org)

Received November 18, 2019; accepted January 13, 2020. doi:10.1002/cpt.1781

either case, high dimensional data streams are generated that require extensive processing and analysis to be converted into phenotypic measures. For gait analysis, inertial measuring units embedded in wearables and phones collect time series data consisting of 100 Hz recordings on 6 axes (three from the accelerometer and three from the gyroscope). Data such as these can be analyzed in three ways. First, features with established clinical relevance (e.g., gait speed) can be extracted through signal processing. For this approach, algorithms are manually evaluated and tuned to maximally approximate the desired phenotypic measure. Although this first approach provides measures with clear clinical interpretation, it can limit use of the full spectrum of information provided in the collected data. To address this, one could opt to use traditional machine learning in conjunction with signal processing to select a subset of promising features from a larger set of exploratory features generated by signal processing. This data-driven approach might be better able to distinguish disease state across heterogeneous populations, as it works by optimizing on the outcome of interest. In the case of gait, this second approach is suited to identifying a broader set of gait disturbances in addition to gait speed. Finally, machine-learning methods based on deep-learning models have also been used to generate features in an automatic and data-driven way, bypassing the need for signal processing.

Although machine-learning approaches provide the opportunity to develop more comprehensive digital measures, the use of machine learning must be done appropriately in order to avoid subtle errors. Because they are data driven, machine-learning approaches will leverage any source of variation in a dataset, including variability due to biology, technical artifacts, and even random noise (especially in small datasets). Identification of biologically relevant measures requires disciplined analysis. This is typically addressed by using two datasets—training data is used to train models and select potential features, whereas a separate validation dataset, assumed to contain similar biological but different technical variation, is used to evaluate the predictive performance of the trained model and confirm the biological relevance of the new features (e.g., by comparing then to existing clinically validated outcomes or severity measures). Because two datasets are not always available in mobile health studies, a single dataset is often split to support both training and validation functions. This can be problematic for small datasets. There are many papers reporting positive validation results of digital measures. Many of these results are developed using machine learning in small sample size studies, which can promote exaggerated results that will not replicate in other datasets. This is best addressed by reporting the uncertainty in measure performance. As an example, a study reporting diagnostic accuracy using the area under the receiver operating curve should be expected to report error bars as a means to help readers understand the uncertainty in the reported performance. In addition, the choice of performance metric is dependent on the nature of the data. Reporting on the incorrect metric (e.g., area under the receiver operating curve in extremely imbalanced datasets) can lead to inflated interpretation of accuracy.[7]

APPLICATION AND CONSEQUENCES OF LONGITUDINAL SAMPLING

A major benefit of mobile health is the opportunity to tailor health monitoring to each individual. This is of particular benefit for conditions and treatments that present in a highly heterogeneous manner across individuals or change dynamically over time. Because mobile health provides longitudinal data collection with frequent sampling, it can be used to capture individualized changes over time by using personalized models or n-of-1 analysis.[3] Analysis of frequently sampled longitudinal data requires an analytical approach that is distinct from those used for sparsely sampled data. Although repeated measures collected from an individual are autocorrelated, a common mistake observed in the literature is to assume that these repeated measures are independent. If not taken into account, autocorrelation can lead to an inaccurate estimate of the number of false-positive discoveries in an analysis. Notably, this can result in either an underestimate or an overestimate depending on whether the autocorrelation is positive or negative.[3,4] Furthermore, the incorrect use of the repeated measurements in population level analysis, such as classification studies can lead to *identity confounding* artifacts, where the classifier is mostly distinguishing differences across individuals instead of differences across conditions or disease states. A recent literature review of mobile health classification studies demonstrated that 47% had artificially inflated the performance of their measures through failure to account for the identity of individual data points.[8] Our own quantification of this effect across three studies showed that identity confounding can be many times larger than the effect of the condition that was being studied.[9] As with the analytical issues described above, proper interpretation of analyses using mobile health studies for classification requires reporting of how repeat measures were handled.

POSTMARKET MONITORING, OPEN ENROLLMENT, AND THE EFFECTS OF CONFOUNDERS

Fully remote mobile health studies can support low cost enrollment of large swaths of the population as compared to in-clinic studies. Many studies relying solely on mobile health measures have enrolled in the tens of thousands from across distributed geographic regions, providing the opportunity for broad sampling across diverse populations in the real-world setting. This approach can be a good option for postmarket monitoring studies, including to evaluate real-world drug efficacy and toxicity as well as market fit. It can also be used to prescreen for enrollees into clinical trials. In these contexts, data are often collected using open enrollment techniques. Because these can lead to biased sampling of the population, they must be carefully evaluated and interpreted. For example, we recently recruited 17,000 individuals into a Parkinson's disease study using an open enrollment approach. The control population tended to be significantly younger than the Parkinson's disease population (average age 38 vs. 61). Because age was correlated with disease status, machine-learning methods could trivially distinguish between cases and controls by selecting features related to age rather than those related to disease state. With careful consideration these issues can be both assessed and accounted for.[10]

In this case, we did so by rebalancing the populations according to clinical covariates and by measuring performance of the classifier both before and after correction for known covariates.[10]

CONCLUSIONS

The use of mobile health to collect frequent measures in a real-world setting provides a promising tool to aid in drug development and monitoring. Appropriate use and interpretation of these approaches, which also provide great opportunity to monitor lived experience, require careful attention to analytical techniques. Much of the success of mHealth will be dependent on comprehensive validation of developed measure and objective benchmarking of analytical techniques used in their interpretation. With appropriate application, these approaches stand to greatly advance our ability to objectively assess the impact of treatments on individuals' lives.

FUNDING

No funding was received for this work.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

---

1. Sherman, R.E. *et al*. Real-world evidence - what is it and what can it tell us? *N. Engl. J. Med.* **375**, 2293–2297 (2016).
2. Goldsack, J. *et al*. Digital endpoints library can aid clinical trials for new medicines - STAT. STAT <https://www.statnews.com/2019/11/06/digital-endpoints-library-clinical-trials-drug-development/> (2019).
3. Chaibub Neto, E. *et al*. On the analysis of personalized medication response and classification of case vs control patients in mobile health studies: the mPower case study <https://ui.adsabs.harvard.edu/abs/2017arXiv170609574C/abstract> (2017).
4. Barnett, I., Torous, J., Staples, P., Keshavan, M. & Onnela, J.-P. Beyond smartphones and sensors: choosing appropriate statistical methods for the analysis of longitudinal data. *J. Am. Med. Inform. Assoc.* **25**, 1669–1674 (2018).
5. Hicks, J.L. *et al*. Best practices for analyzing large-scale health data from wearables and smartphone apps. *NPJ Digit. Med.* **2**, 45 (2019).
6. Perry, B. *et al*. Use of mobile devices to measure outcomes in clinical research, 2010–2016: a systematic literature review. *Digital Biomarkers* **2**, 11–30 (2018).
7. Cokelaer, T. *et al*. DREAMTools: a Python package for scoring collaborative challenges. *F1000Res.* **4**, 1030 (2015).
8. Saeb, S., Lonini, L., Jayaraman, A., Mohr, D.C. & Kording, K.P. The need to approximate the use-case in clinical machine learning. *GigaScience* **6**, 1–9 (2017).
9. Chaibub Neto, E. *et al*. Detecting the impact of subject characteristics on machine learning-based diagnostic applications. *NPJ Digit. Med.* **2**, 99 (2019).
10. Chaibub Neto, E., Tummalacherla, M., Mangravite, L. & Omberg, L.Causality-based tests to detect the influence of confounders on mobile health diagnostic applications: a comparison with restricted permutations <https://arxiv.org/abs/1911.05139> (2019).

Citation: Clin Transl Sci (2019) 12, 247–256; doi:10.1111/cts.12602

ARTICLE

Evaluation of Wearable Digital Devices in a Phase I Clinical Trial

Elena S. Izmailova[1,*], Ian L. McLean[1], Gaurav Bhatia[2], Greg Hather[1], Matthew Cantor[2], David Merberg[1], Eric D. Perakslis[1], Christopher Benko[2] and John A. Wagner[1]

We assessed the performance of two US Food and Drug Administration (FDA) 510(k)-cleared wearable digital devices and the operational feasibility of deploying them to augment data collection in a 10-day residential phase I clinical trial. The Phillips Actiwatch Spectrum Pro (Actiwatch) was used to assess mobility and sleep, and the Vitalconnect HealthPatch MD (HealthPatch) was used for monitoring heart rate (HR), respiratory rate (RR), and surface skin temperature (ST). We measured data collection rates, compared device readouts with anticipated readings and conventional in-clinic measures, investigated data limitations, and assessed user acceptability. Six of nine study participants consented; completeness of data collection was adequate (> 90% for four of six subjects). A good correlation was observed between the HealthPatch device derived and in-clinic measures for HR (Pearson $r = 0.71$; $P = 2.2\text{e-}16$) but this was poor for RR ($r = 0.08$; $P = 0.44$) and ST ($r = 0.14$; $P = 0.14$). Manual review of electrocardiogram strips recorded during reported episodes of tachycardia > 180 beats/min showed that these were artefacts. The HealthPatch was judged to be not fit-for-purpose because of artefacts and the need for time-consuming manual review. The Actiwatch device was suitable for monitoring mobility, collecting derived sleep data, and facilitating the interpretation of vital sign data. These results suggest the need for fit-for-purpose evaluation of wearable devices prior to their deployment in drug development studies.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Wearable sensors have the potential to collect health-related data remotely, thus enabling acquisition of dense physiological study subject profiles, allowing data collection on an outpatient basis, and thereby reducing the number of clinical study hospital or clinical pharmacology unit (CPU) visits.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ We evaluated the performance of two FDA 510(k)-cleared devices, HealthPatch MD by Vitalconnect and Actiwatch Spectrum Pro by Phillips, for continuous physiological data collection, compared device readouts with conventional analogous measures and published data, and assessed operational feasibility in a residential phase I clinical trial.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ The Actiwatch device was suitable for monitoring mobility, collecting derived sleep data, and providing metadata for interpreting vital sign data. The HealthPatch device was not determined to be "fit-for-purpose" because of the artefacts and the need of extensive, time-consuming manual data review.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ Our study results indicate the need for evaluation of wearable digital device according to fit-for-purpose principle in the context of clinical investigations.

Despite the widespread adoption of consumer digital technologies and their increasing use in healthcare settings, they have yet to find widespread application in industry-sponsored drug development. Some progress has been made, including pilot studies for remotely run clinical trials,[1–3] novel technological solutions to improve medication adherence,[4] and multiple modalities of using digital sensors to create new data streams to improve the collection of health-related data.[5] Despite these efforts, published reports of study results remain limited. Moreover, the results of some studies indicate that digital innovation in health care is more complicated than anticipated[6] and that some technologies do not perform as the researchers had planned.[7] In addition, published results indicate that there is a need for extensive manual review of data and for the investigation of potential device-derived data artefacts, activities that can be time consuming.[8]

---

[1]Takeda Pharmaceuticals International, Inc., Cambridge, Massachusetts, USA; [2]Koneksa Health, Inc., New York, New York, USA. *Correspondence: Elena S. Izmailova (elena.s.izmailova@gmail.com)

Received: 18 September 2018; accepted: 26 October 2018. doi:10.1111/cts.12602

In early-stage drug development clinical trials, vital sign data, such as heart rate (HR) and respiratory rate (RR), are typically collected manually by clinical personnel or by using electronic devices that record these data at discrete single time points. This is generally done at predefined times before, during, and after administration of the study drug, while a subject is a resident at a clinical pharmacology unit (CPU) or returns for follow-up visits. Additional assessments of vital sign data are performed in response to suspected safety or tolerability issues or if the study drug or a challenge agent is expected to have pharmacological effects on vital signs. High-density vital sign data recorded continuously using wearable digital sensors have the potential (i) to provide more information on study subjects' physiological profiles and greater sensitivity for detecting changes in these parameters, (ii) to allow periods of data collection to be done in the outpatient/home setting rather than as inpatients in the residential CPU setting, and (iii) as an aid in interpretation of adverse events, with an overall view to reducing the time of residential observation during phase I studies and the number of follow-up clinic visits needed.

Wearable digital devices may also have utility for evaluating the impact of a novel medicine on disease activity or outcomes. In many therapeutic indications, the impact of a drug on activities of daily living, including physical activity levels and sleep patterns, is captured routinely as an indication of the potentially clinically relevant benefit of therapy or potential negative side effects. These assessments usually rely on the subject's ability to subsequently recall these events for self-completed questionnaires. This type of data can be subjective, vague, and prone to confounding and bias,[9] aspects that may be improved by continuous real-time collection of activity-related data by digital devices to objectively monitor activities of daily living.

The selection and deployment of appropriate wearable digital devices in the context of drug development presents challenges[5] that are similar to those encountered with the introduction of novel laboratory biomarker tests in the early 2000s. To address these challenges, the "fit-for-purpose" concept was developed by the American Association of Pharmaceutical Scientists (AAPS) Biomarker Workshop[10] and advanced further by the US Food and Drug Administration (FDA)-National Institutes of Health (NIH) Biomarkers Endpoints and other Tools (BEST) working group.[11] Using this framework, a potential biomarker should be evaluated for a predefined purpose in the context that it will be used. We applied this approach to evaluate two wearable digital devices that have 510(k) clearance from the FDA: the Phillips Actiwatch Spectrum Pro (Actiwatch) and the Vitalconnect HealthPatch MD (HealthPatch). We incorporated the testing of these devices as an exploratory component in a 10-day residential phase I study recruiting normal healthy volunteers.

The goal of this substudy was to evaluate whether the HealthPatch and Actiwatch devices were fit-for-purpose to enhance vital sign data collection and to capture physical activity in the context of an industry-sponsored early-phase drug development study. Aspects examined included (i) a comparison with the traditional conventional measures performed at the clinical site, (ii) assessment of a diurnal variation of physiological parameters that were expected to conform to expected temporal patterns, and (iii) understanding data limitations and technical issues. We also assessed the operational aspects of device use, including acceptability for the study subjects and the site personnel.

METHODS

The clinical study was conducted at a US-based single-site residential CPU for a 10-day period. All subjects were healthy volunteers recruited from the CPU's panel; they had no clinically significant acute or chronic medical disorders, were taking no concomitant medications, and had no exposure to other investigational agents in the 30 days preceding the study. The devices were deployed during the CPU confinement period only. Informed consent was obtained separately for the device component of the study, which was optional for any subject consenting to participate in the core part of the study. The study conduct was reviewed and approved by the institutional review board.

For the design of the study and authoring of the study protocol, the vital signs, activity, and sleep data produced by wearable devices were treated as "exploratory," used for device evaluation purpose only, and not linked to primary or secondary study endpoints, which included pharmacokinetic and safety assessments. The data were not available to CPU or sponsor staff during the conduct of the study and were not intended to guide clinical care or other decision making.

The Actiwatch[12] was worn on the wrist using a standard wristwatch-style strap and captured data on motion using an accelerometer, which were used to derive information on activity level and sleep. Activity level is summarized using activity counts, a dimensionless measure of motion that removes the effects of gravity, transportation, and other acceleration not indicative of activity. The HealthPatch[13] was applied to the anterior surface of the left upper precordium using an adhesive strip and captured biometric data: HR, RR, skin temperature (ST), and step count. Both devices were intended to be worn throughout the entire 10-day period of confinement in the CPU. At the end of the study, the site personnel and the study participants were asked to complete a satisfaction questionnaire.

Device data collection

The data collected by the Actiwatch were retrieved by periodically connecting it to a laptop computer running study-specific software, which downloaded the epoch level data from the device to the computer before transfer to the Philips database (**Figure S1**). The HealthPatch device recorded a single-lead anterior chest wall echocardiogram (ECG) voltage every 8 ms, and from the resulting R-R interval an estimate of HR was calculated approximately every 4 seconds, averaging 15 HR estimates within a minute. The data collected by the HealthPatch were streamed from the HealthPatch to a companion iPhone application (Healthwatch, version 2.5.4) on a dedicated iPhone 5 via Bluetooth technology (**Figure S1**).

Device data processing

HealthPatch data were first subjected to a quality control step during which invalid readings were filtered using the manufacturer's proprietary software.

To facilitate estimation of data completeness, gap thresholds were defined. This threshold (T) was set to 5 and 30 seconds for HealthPatch and Actiwatch, respectively. Then, for each device–subject data stream, data were sorted in timestamp order, and the intervals between valid recordings were calculated. If an interval was greater than the gap threshold, it was considered a gap (i.e., missing data). Total noncovered time was calculated by summing the length of all gaps. Percent completeness was defined as 100% × (1-(device noncovered time)/total study time).

We calculated compliance separately for each individual and each device using the millisecond coverage technique. This technique is designed to account for the slight variability in the rate at which measurements are taken by measuring the percentage of on-study time that is within T (defined above) seconds of a valid measurement. Compliance was estimated as the proportion (%) of on-study milliseconds within T seconds of a valid measurement.

Summary statistics

Computation of summary statistics across subjects and time points allowed us to explore the reasonableness of the data. We computed the arithmetic mean, SD, and minimum and maximum for all data sources and individuals. In addition, we computed a measurement timeline, averaged across individuals for the full study. This allowed us to explore any diurnal patterns in the data. All statistical analysis was done using the R software package version 3.3.2 with software libraries "plyr" (https://cran.r-project.org/web/packages/plyr/index.html) for data processing and "lme4" (https://cran.r-project.org/web/packages/lme4/index.html) for fitting of linear mixed models.

Assessment of diurnal variation

The degree to which measurements varied as a function of time of day was analyzed by calculating the minute-by-minute averages of HR, RR, and ST measurements and plotting these as a function of time of day. For each measurement, we calculated the minute of the day when the measurement was made. All measurements for the same minute were combined; for example, to calculate the "average HR" for 8:01 AM, all HR measurements taken between 8:01 AM and 8:02 AM were examined for all study days and for all subjects. Measurements were also grouped into "daytime" and "nighttime" periods. For this analysis, daytime was defined as between 8:00 AM and 9:00 PM and nighttime as between 12:00 AM and 6:00 AM. The periods between 6:00 AM and 8:00 AM and between 9:00 PM and 12:00 AM were anticipated to be "grey areas" with considerable variability within and between the subjects as to awake or asleep status during these periods.

Comparison between conventional in-clinic and wearable device measures

We compared the HR, RR, and ST measurements reported by the HealthPatch to the time-matched clinic measurements of HR, RR, and core body temperature (BT), respectively. In-clinic HR was collected using the Dinamap device. The study site's electronic source data system automatically captured the procedure timestamp at the time of collection. BT was collected using an electronic oral thermometer again linked to the Dinamap unit. In-clinic RR was collected manually by the site staff by observing the subjects' chest wall movements and counting respiration cycles over a defined period and entered immediately into the site's system, together with the time of data entry. We mapped the corresponding data from the wearable devices to the in-clinic data and then assessed the degree of concordance between the mapped data points at matched time points. The degree of concordance between in-clinic and wearable device data was determined using three separate strategies: correlation, regression, and Bland-Altman analyses. First, we calculated the Pearson correlation coefficient between the in-clinic and mapped wearable measurements. For regression, we performed ordinary least squares regression with the in-clinic measurement as the independent variable and the wearable measurement as the dependent variable.

For the Bland-Altman analyses,[14] we produced Bland-Altman Mean (BAM) and Bland-Altman Difference (BAD) values. For each in-clinic measurement, BAM was the average of the in-clinic and mapped wearable measurements, and BAD was the difference between the in-clinic and mapped wearable measurement. Bland-Altman plots were generated consisting of a scatterplot of BAM (x-axis) against BAD (y-axis). The points were color-coded by individual to help visualize any individual-specific bias. Computations of mean bias and 95% limits of agreement were also performed. The mean bias was simply the mean of BAD values. The 95% limits of agreement were calculated as 1.96 times the SD of BAD values.

Comparison between HealthPatch and Actiwatch actigraphy data

We investigated the extent of correlation between the HealthPatch step count and Actiwatch activity units. We divided each subject's time in the clinic into 1-hour intervals, beginning and ending on the hour. We then summed total activity units reported for each hour and calculated steps by subtracting step count at the start of the hour from step count at the end. Based on the epoch time (~1 second for HealthPatch and 30 seconds for Actiwatch), we calculated the number of measurements comprising complete data for 1 hour. We excluded any interval that was < 90% complete in either measurement and determined the Pearson correlation coefficient between activity units/hour and steps/hour.

RESULTS

Subject demographics

Six of the nine subjects enrolled in the core clinical study consented to participate in the exploratory wearable digital device evaluation component. Reasons given for nonconsent were the following: a history of prior severe cutaneous hypersensitivity to adhesives (one subject) and the perception that an honorarium payment should be offered by the sponsor for additional study procedures. The demographic profile of the participants was as follows: five men and one woman; age range 18–55 years inclusive; three white subjects, two African-American subjects, one multiracial subject, and one smoker.

Completeness of data collection

For the HealthPatch, data completeness rates among the six subjects ranged from 83.6–99.2% (mean 93.1 ± 7.4%; **Table S1** and **Figure S2**). For the Actiwatch, completeness rates ranged from 62.6–98.6% (mean 88.9 ± 15%); the low rate (62.6%) occurred because the subject removed the device for the last 3 days of the study for unspecified reasons. Periods of loss of valid data for the HealthPatch device were attributed to poor skin contact and subjects removing the devices, again for unspecified reasons. Additional loss of valid Actiwatch data occurred due to device calibration issues and subjects removing devices without reporting this to the site personnel (**Figure S2**).

Comparison of in-clinic and wearable device measures

Comparison of the paired HR data showed a strong correlation between in-clinic and wearable measurements (Pearson's $r = 0.71$, $P = 2.2e\text{-}16$; **Figure 1**), confirmed by regression analysis ($\beta = 0.81$). Bland-Altman analysis (**Figure 1**) showed that, on average, in-clinic measurements were 0.91 bpm lower than their wearable counterparts. We estimated the 95% limits of agreement at 7.2 bpm, corresponding to 11% of the mean HR. Overall, the HR wearable-device data correlated well with the traditional in-clinic counterpart.

Recordings for RR derived from the HealthPatch seemed to be substantially different from the corresponding in-clinic measures. There was no significant relationship between the in-clinic and wearable device measurements (**Figure 1**) by either correlation (Pearson's $r = 0.08$, $P = 0.44$) or regression analysis ($\beta = 0.14$). Bland-Altman analysis corroborated these findings, with 95% limits of agreement of 6.0 breaths/min corresponding to 35% of the mean RR, indicating that the RR reported by the HealthPatch were statistically independent of the RR measured in-clinic.

There was no significant relationship between the ST as reported by the HealthPatch and in-clinic oral BT by either correlation (Pearson's $r = 0.14$, $P = 0.16$) or regression analyses ($\beta = 0.31$). Bland-Altman analysis indicated that ST and oral temperature had different distributions (**Figure 1**).

Actigraphic mobility and sleep data

The actigraphy data indicated much lower movement activity during the nighttime period, as expected. **Table 1**

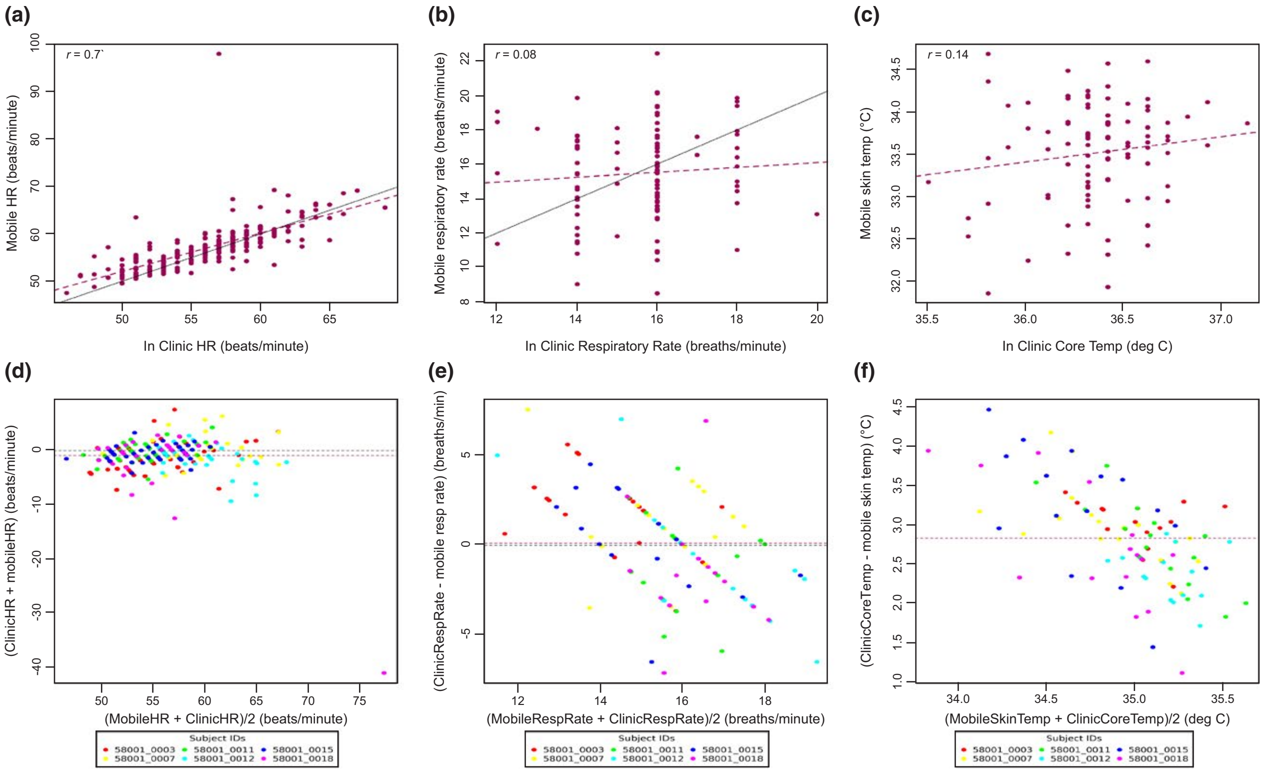


Figure 1 The comparison of in-clinic and wearable measurements for heart rate (HR), respiratory rate (RR), and skin temperature (ST) by correlation, regression (**a**–**c**), and Bland-Altman (**d**–**f**) analyses. The solid lines in **a**, **b**, and **c** are the line $y = x$, which would be expected if concordance were perfect; the dotted line is the actual regression line. In **c** there is no solid line because ST is significantly lower than core temperature. The Bland-Altman analysis results are presented in **d**, **e**, and **f**. The black dotted line is the line $y = 0$ indicating no mean difference between the two measures. The red dotted line is the mean difference line based on the actual data. **f** ST there is no black dotted line because ST is significantly lower than core temperature. **a** and **d** show the analyses for HR, **b** and **e** shows RR, **c** and **f** shows ST. Points are color-coded with a unique color for each subject.

Table 1 Summary statistics for activity measurements and total sleep time derived from the Actiwatch data

Subject ID	Daytime activity, counts/minute min–max	Nighttime activity, counts/minute min–max	TST, minutes min–max	TST, minutes mean ± SD
58001_0003	0–1132	0–801	196–501	367.9 ± 89
58001_0007	0–1453	0–514	190.5–536	376.5 ± 126.7
58001_0011	0–1499	0–584	306–471.5	367.5 ± 51.4
58001_0012	0–1063	0–908	118.5–383.5	268.6 ± 90.3
58001_0015	0–1697	0–998	313–498.5	409.8 ± 59.5
58001_0018	0–1697	0–1063	242–439	362.6 ± 68.7
OVERALL	0–1697	0–1063	118.5–536	358.8 ± 91.1

TST, total sleep time.
Range by endpoints for measurements of daytime activity, nighttime activity, and TST.

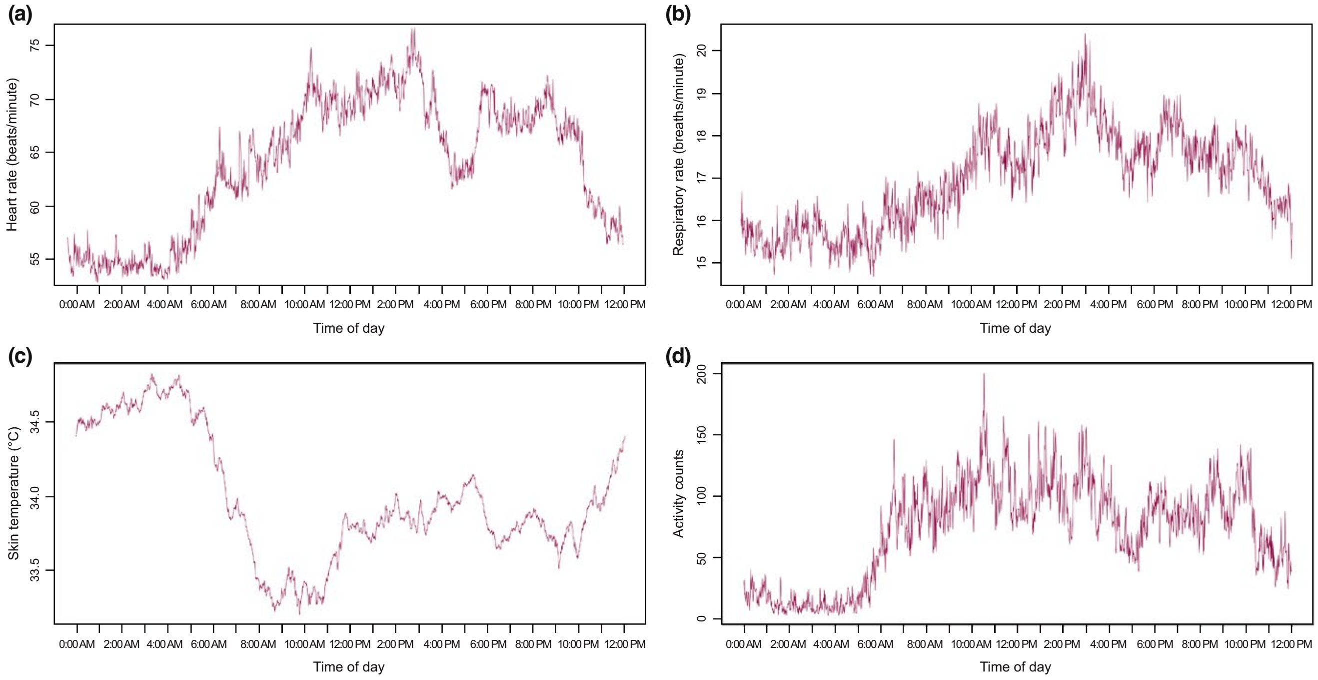


Figure 2 The aggregate diurnal patterns for heart rate (**a**), respiratory rate (**b**), skin temperature (**c**), and activity counts (**d**).

shows a summary of the data range (minimum and maximum) recorded by the Actiwatch. Analysis of total sleep time (**Table 1**) suggested that the subjects' mean sleep time was of 358 ± 91.1 minutes, or ~6 hours per night, with substantial variation between subjects.

Additionally, we compared actigraphy data derived from both HealthPatch (chest worn) and Actiwatch (wrist worn) devices by examining correlation of corresponding device outputs: step counts and activity counts. The comparison of these activity measures indicated that the readouts from these devices were broadly in agreement (**Figure S3**). The degree of correlation varied among study subjects (**Table S2**). However, we consistently observed some low values reported from Actiwatch devices corresponding to 0 values reported by the HealthPatch device (**Figure S3**) in all subjects.

Diurnal patterns

HR, RR, ST, and movement activity all demonstrated significant diurnal variation. HR, RR, and activity displayed very similar temporal patterns: lowest at night, highest in the afternoon, and an early evening nadir (**Figure 2**), as expected. The ST showed a different pattern; we observed the highest ST at night followed by a sharp drop in the morning.

Vital sign data

Most HR values reported by the HealthPatch were within or close to the normal range for healthy adults at rest, ~50–100 beats/min (bpm). Similarly, most reported RR recordings were within the expected physiological range (**Figure 3**). Most measurements for ST were within the range of previously reported for healthy adult normal ST (33–35°C; **Figure 3**) but were significantly different from

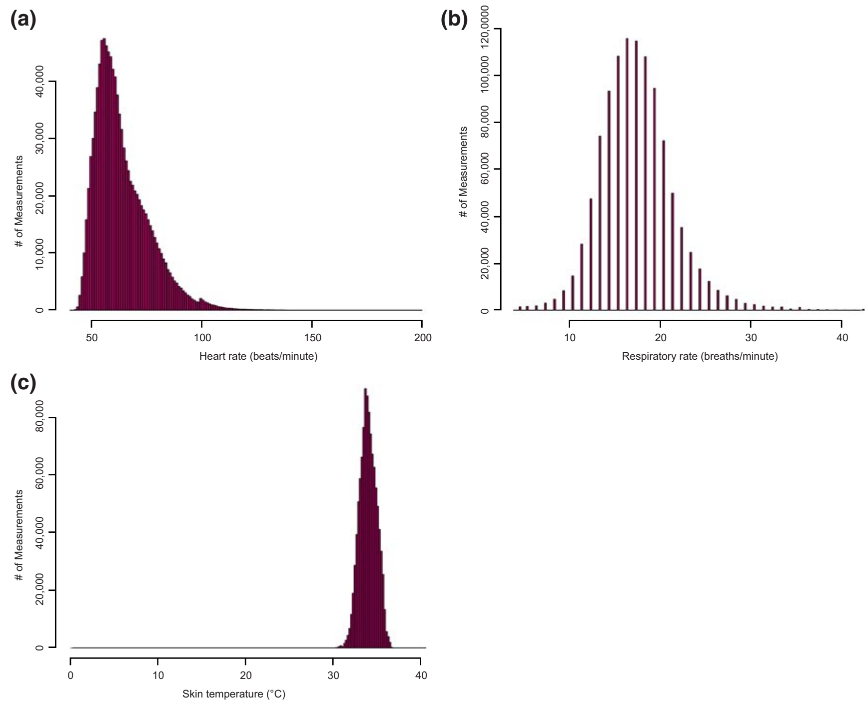


Figure 3 The histogram depicting the range of value distribution for heart rate (**a**), respiratory rate (**b**), and skin temperature (**c**).

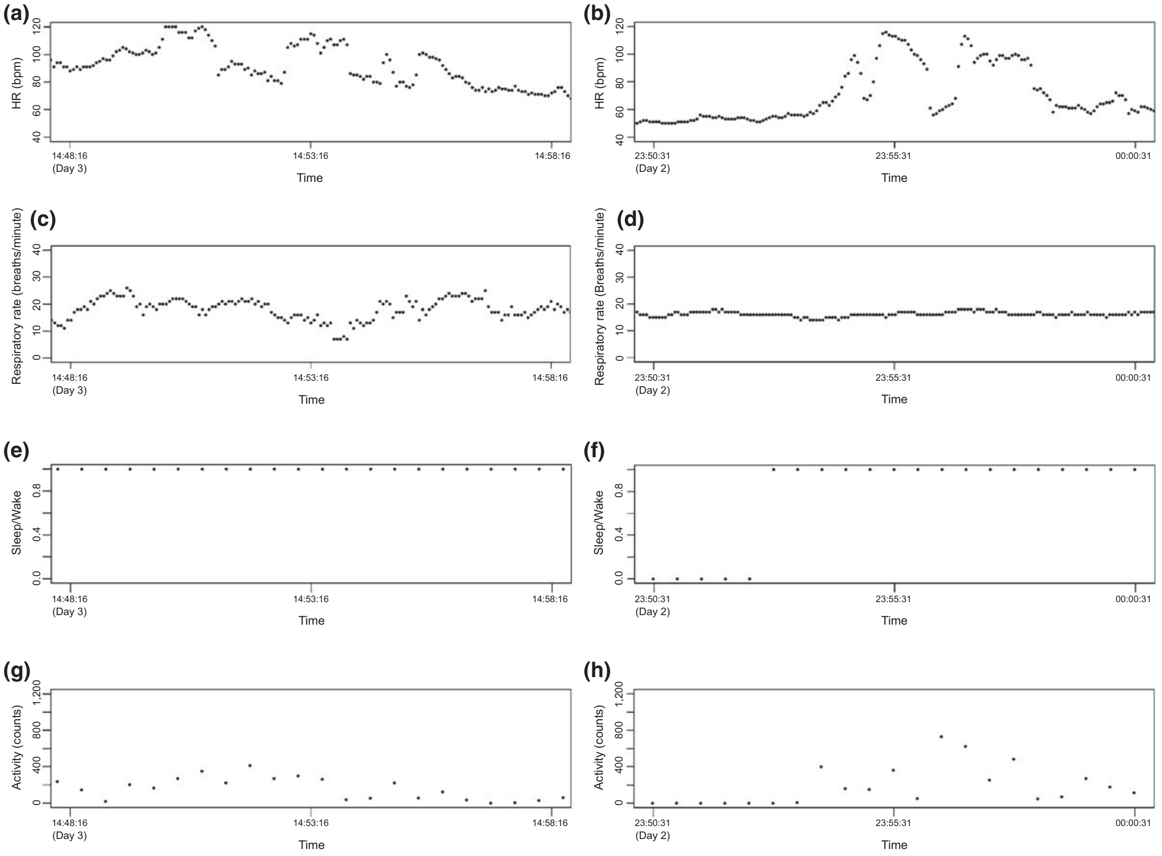


Figure 4 Visualization of heart rate (HR) (**a**, **b**) and respiratory rate (**c**, **d**) temporal patterns in the context of wake–sleep data (**e**, **f**), and subject mobility as assessed by activity counts (**g**, **h**) for subjects 58001_0003 (**a**, **c**, **e**, and **g**) and subject 58001_0011 (**b**, **d**, **f**, and **h**). bpm, beats/min.

the body core temperature as measured with the oral thermometer at the same time point.

Investigation of reported abnormal HR and RR values

To aid interpretation of HR and RR values that were outside of the expected resting range we compared the temporal change patterns of HR and RR as reported by the HealthPatch with activity level and wake-sleep status derived from the Actiwatch. Elevated readings were separated into two categories: (i) modestly elevated values that might be readily ascribed to changes in body position, physical activity, and/or study procedures, and (ii) significantly elevated values – those that would require further attention as they could indicate potential drug-related safety signals should they occur in a phase I setting and be temporally related to exposure to the study drug or to a challenge agent. Readouts for the same period were examined graphically (**Figure 4**), revealing that most episodes of elevated HR and RR occurred at times of increased physical activity. We found that comparing time-matched data by direct visual comparison of HR and RR data on the one hand and activity counts and wake-sleep status on the other was informative for interpreting moderately abnormal readings. However, this visual review process was time-consuming.

There were several episodes when the HR was elevated above 150 bpm or even 180 bpm, which exceeds reported "normal" HR ranges for a comparable healthy volunteer study population.[15] This finding was unexpected given that the study recruited healthy volunteer subjects, that the study subjects were confined to the CPU for the duration of the wearable device evaluation, that no cardiac-related adverse events were expected based on the properties of the investigational compound or were detected during the study by conventional means, and that, in compliance with the study protocol, the subjects were restricted from strenuous physical exercise. Manual reviews of the ECG waveforms were done for periods during which HR values were below 50 bpm or above 180 bpm. This indicated that all device-reported values below the normal range were consistent with the reported in-clinic HR values (**Figure 5a**), consistent with sinus bradycardia occurring in healthy individuals.

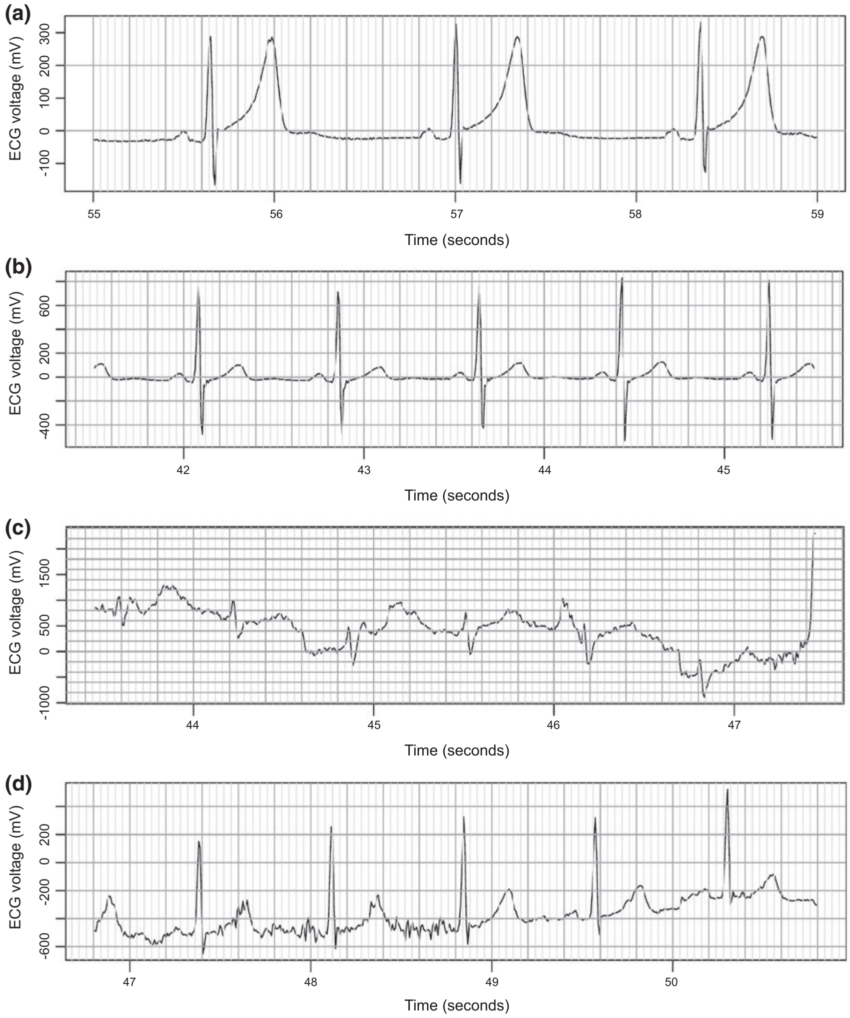


Figure 5 Visual examination of the echocardiogram (ECG) waveforms for episodes with reported heart rates of 44 bpm (**a**), 79 bpm (**b**), 203 bpm (**c**), and 193 bpm (**d**), bpm, beats/min.

However, examination of single-lead ECG strips corresponding to reported HR values above 180 bpm revealed artefacts of data recording and data processing (**Figure 5, c and d**) rather than true episodes of tachycardia. This additional investigation of these reported periods of out-of-range HR was resource-intensive, requiring manual visual review by a physician of each ECG tracing corresponding to these periods.

The number of time intervals containing values outside of normal range that would require follow-up was determined by calculating the number epochs with HR values above 150 bpm and above 180 bpm (**Table 2**). The number of such epochs was highly variable between subjects, with the highest number for subject 58001_0018. We also calculated the number and total duration of gaps in data collection in order to estimate the overall amount of vital sign data not being collected, thus providing an estimate of the likelihood of missing a safety signal if the device had been used as the primary method for collecting vital sign data. **Table S3** shows the number of time intervals and a total duration of such intervals, with the highest number occurring in subject 58001_0018. We found significant gaps in the completeness of vital sign data collection, with between 0.6 and 35.4 hours of data missing over the course of the 10-day study period.

The reported minimum ST values of 0°C were likely due to poor patch adherence or other artefact, despite the apparently valid skin impedance measures.

Study subject and site personnel feedback

At the end of the study, both the study participants and the study coordinators were asked to complete a brief satisfaction questionnaire to assess their comfort level with the device technology used and their willingness to participate in similar study procedures in the future. The majority of the subjects indicated a high level of acceptability for the devices and a willingness to participate in future studies that assessed wearable digital device technologies.

Overall, study coordinators at the CPU reported high levels of satisfaction on the training and technical support they receive. However, in free-form feedback the study coordinators did highlight a desire to gain additional "hands-on" experience with devices and their associated software in order to increase their comfort level.

DISCUSSION

This study assessed the feasibility of using 510(k)-cleared wearable digital devices for collecting physiological and activity data in the context of a residential drug development clinical study together with the performance and fit-for-purpose validation of these devices. The HealthPatch and Actiwatch devices used did not interfere with the other study procedures, such as dosing, safety, and pharmacokinetic sample collection and were well received by the study subjects and the site personnel. Activity counts and sleep duration data derived from the Actiwatch device had face validity, followed expected diurnal patterns, and were consistent with previously published results.[16] The average total sleep time of 6 hours is perhaps shorter than a typical night's sleep for a healthy adult; this may be a consequence of the unfamiliar environment of the CPU. The observed diurnal patterns in HR, RR, ST, and activity were consistent with the previous reports.[17,18] Our experience of periods of missing vital sign data is consistent with previously published results.[8]

Table 2 Quantification of number and percent of epochs with HR values above 150 and 180 bpm

Subject	Number (%) of epochs with HR > 150 bpm	Number (%) of epochs with HR > 180 bpm
58001_0003	29 (0.02)	13 (0.01)
58001_0007	18 (0.01)	5 (0.00)
58001_0011	9 (0.00)	6 (0.00)
58001_0012	67 (0.04)	6 (0.00)
58001_0015	99 (0.06)	33 (0.02)
58001_0018	258 (0.16)	77 (0.05)

bpm, beats/min; HR, heart rate.

There were significant limitations with the HR data produced by the HealthPatch device because of the volume of artefacts produced that require a follow-up and a manual review. Although the HR recordings showed good correlation with traditional in-clinic measures, there were many reported episodes of tachycardia due to voltage artefacts that required manual visual review by a physician to resolve.

Differences between RRs as determined by the device and by manual in-clinic measures may in fact reflect the differences between manual and device-mediated methods, which have been described by other groups.[19] Differences between surface ST and body core temperature are expected.[20] ST is typically lower and more variable than BT[20,21] and is affected by the site of measurement, clothing, environmental temperature, and even emotional state. However, we expected some degree of relationship between these two variables. In general, we observed much smaller variation in core temperature than ST, as expected, which further underscores the distinction between the two parameters.

The comparison of activity data generated by HealthPatch and Actiwatch devices revealed that the measurements were broadly in agreement although not perfectly correlated. Less than perfect correlation was expected because the devices are located on different parts of the body (trunk vs. wrist); in addition, different activity readouts (step counts vs. activity counts) impacted the types of physical activities detected (i.e., walking vs. moving the upper body only). Given that the Actiwatch readout provided data for a wider variety of physical activities, the activity counts derived from Actiwatch were used for vital sign data interpretation. These differences illustrate the need for data standardization for similar device readouts (e.g., variables associated with actigraphy).

The extent to which each device was evaluated was driven by the intended use of the data. We applied more rigor for the HealthPatch data analysis because of its potential to detect a safety signal. The data derived from the Actiwatch device played a secondary role and were largely used to interpret the vital sign values outside of the normal range.

The device evaluation portion of this study has several limitations. In the broad context of drug development, the phase I study in which these devices were evaluated was of low complexity in terms of device implementation logistics, data analysis, and interpretation. The study subjects were inpatients and were monitored by the CPU personnel for the entire duration of the study, which facilitated data interpretation and helped to qualify certain findings as artefacts. In addition, the small size of the study ($N = 6$) allowed the many reports of out-of-range vital signs (e.g., potential tachycardia episodes) to be reviewed manually. This may prove to be too resource-intensive in a larger study. Our findings demonstrated significant challenges with continued device use, data collection, processing, and, most importantly, data interpretation. We anticipate that the impact of these issues would be even higher with a study involving more subjects, if additional procedures were included (e.g., imaging or invasive sampling) or if the device component of the study was done remotely with study subjects wearing the devices at home. We also observed missing data, an effect that is likely to be amplified in study subjects with medical conditions who are seen in usual practice. Adherence was an issue in this study of relatively low complexity, although it was comparable with other similar reports[8]; it is likely to be a limiting factor in subjects with disease conditions as well. In addition, our ability to review information was somewhat restricted, as the device manufacturer uses proprietary software algorithms for initial data processing.

Further limitations on the use of wearable devices in drug-development studies are that with the model of the device evaluated in this study and the available analysis software packages: (i) the derivative data were not available in “real time” during the study, which would delay the detection of acute safety signals and would not enable the investigator or sponsor to make real-time clinical decisions on the management of safety issues, and (ii) the single-lead ECG information that is reported is limited to HR, and, in particular, does not provide interpretation of potential rhythm abnormalities or of important electrocardiographic intervals, such as the QTc time. However, for arrhythmia detection, a single-lead ECG device has the advantages of ease of use and convenience (compared with conventional Holter monitoring[22]) and in this context of use, might be fit-for-purpose with appropriate software development. The use of a chest wall patch device to detect atrial fibrillation in a pragmatic population-based study of over 2,000 subjects was recently reported.[23] In addition, our study demonstrated the critical need for access to the source data in order to evaluate study results, to confirm reports of abnormal activity, and to understand the data limitations. Source data and algorithm transparency remain an issue with both consumer grade and some medical grade devices in the context of clinical investigations.

Our findings indicate the need for careful evaluation of wearable digital devices according to the fit-for-purpose principle before the device-derived data can be used to support primary or secondary study endpoints, irrespective of the regulatory clearance. Regulatory evaluation of device performance under the auspices of the FDA 510(k) clearance program conveys a level of reassurance regarding the technical performance of a device. However, receipt of 510(k) clearance should not be taken to imply that the device is fit-for-purpose for an industry-sponsored drug development study. We did not clinically validate the HealthPatch as fit-for-purpose for augmented physiological data collection because of artefacts, including false-positive HR signals and missing data. This issue of false-positive signals is not inherent to any specific device. Several groups have previously reported a false-positive rate from ECG monitors in the intensive care unit setting as high as 75–93%.[24] Our finding of poor correlation between device-reported and manual in-clinic measurements of RR is also consistent with results reported previously.[19] Nonetheless, the issue of specificity of safety monitoring limits the potential utility digital devices for drug development. We believe that a device similar to the HealthPatch device could be of utility for monitoring study subjects if the issues of false-positive results and missing data are addressed to an acceptable extent. The data derived from such a device can be used in a manner similar to an “early warning score” system[25] to generate signals to be investigated further and facilitate building an investigational drug safety profile early during clinical development.

There are many promising uses of wearable devices in clinical trials as well as several challenges. Potential applications drive toward an enhanced understanding of disease variability, treatment response, safety assessment, innovation in clinical trial design and conduct, as well as increasing efficiency and decreasing costs in clinical trials. Although the promises are clear, the challenges are not insignificant and include scientific, regulatory, ethical, legal, data management, infrastructure, analysis, and security challenges.[5] The current study demonstrates practical examples of scientific/regulatory, data management, infrastructure, and analysis issues, as described above. We did not encounter significant obstacles with ethical, legal, or security issues, but the importance of these may have been diminished by the pilot nature of this substudy.

In summary, comparison between specific wearable digital devices and in-clinic measures established a strong correlation for HR but poor correlation between in-clinic and wearable measurements of RR and of ST using the HealthPatch. We concluded that the HealthPatch was not fit-for-purpose for HR monitoring because of the artefacts it produced and the amount of time required for data processing and review. The number of artefacts would need to be greatly reduced before a wider of implementation of this device in clinical trials. The Actiwatch device was used as a supporting application to interpret the vital sign data and was suitable for the intended purpose of monitoring movement, aiding interpretation of abnormal vital sign data, and collecting certain sleep parameters. For wearable devices to gain wider applicability in drug development, we need to develop and establish acceptance for common issues, including medical need, device choice, context of use, fit-for-purpose validation, and predefined operational requirements, as well as data collection, processing, and interpretation. Careful consideration must be given to clinical validation and context of use to assure that device measurements are fit-for-purpose. Clinical Trial Transformation Initiative made substantial progress in addressing these issues by developing recommendations for implementation of mobile technologies in human experimentation.[26] However, there is a great need to supplement these recommendations with the results derived from clinical studies. The current study illustrates the critical role for evaluation, both analytical and clinical, in the applicability of wearable devices.

Supporting Information. Supplementary information accompanies this paper on the *Clinical and Translational Science* website (www.cts-journal.com).

Figure S1. Data flow diagrams for the Actiwatch and the HealthPatch devices in the phase I study.
Figure S2. Completeness of data collection for the HealthPatch (**a**) and Actiwatch (**b**) devices.
Figure S3. The comparison of activity measurements by the HealthPatch (step counts, x-axis) and Actiwatch (activity counts, y-axis) devices for study subjects 58001_003 (**a**), 58001_007 (**b**) 58001_0011 (**c**), 58001_0012 (**d**), 58001_0015 (**e**), and 58001_0018 (**f**).
Table S1. Completeness of data collection for the HealthPatch and Actiwatch devices for each study subjects.
Table S2. Correlation of activity measurements by the HealthPatch and Actiwatch devices.
Table S3. Quantification of time intervals with missing data.

Acknowledgments. The authors thank Kristina Allikmets and Jason Homsy for help with the data review and interpretation.

Funding. No funding was received for this work.

Conflict of Interest. E.S.I., I.L.M., G.H., D.M., E.D.P., and J.A.W. are employees of Takeda Pharmaceuticals and own stock in the company. G.B., M.C., and C.B. are employees of Koneksa Health and own stock in the company. As Editor-in-Chief for *Clinical and Translational Science*, J.A.W. was not involved in the review or decision process for this paper.

Author Contributions. E.S.I., I.L.M., G.B., D.M., and J.A.W. wrote the manuscript. E.S.I., I.L.M., G.B., M.C., E.D.P., C.B., and J.A.W. designed the research. E.S.I., I.L.M., G.B., G.H., M.C., and C.B. performed the research. E.S.I., I.L.M., G.B., G.H., and D.M. analyzed the data.

1. Mealer, M. *et al.* Remote source document verification in two national clinical trials networks: a pilot study. *PLoS One* **8**, e81890 (2013).
2. Wu, T.C. *et al.* Telemedicine-guided remote enrollment of patients into an acute stroke trial. *Ann. Clin. Transl. Neurol.* **2**, 38–42 (2015).
3. Smalley, E. Clinical trials go virtual, big pharma dives in. *Nat. Biotechnol.* **36**, 561–562 (2018).
4. Rosenbaum, L. Swallowing a spy - the potential uses of digital adherence monitoring. *N. Engl. J. Med.* **378**, 101–103 (2018).
5. Izmailova, E.S., Wagner, J.A. & Perakslis, E.D. Wearable devices in clinical trials: hype and hypothesis. *Clin. Pharmacol. Ther.* **104**, 42–52 (2018).
6. Clifton, L., Clifton, D.A., Pimentel, M.A., Watkinson, P.J. & Tarassenko, L. Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors. *IEEE J. Biomed. Health Inform.* **18**, 722–730 (2014).
7. Roess, A. The promise, growth, and reality of mobile health - another data-free zone. *N. Engl. J. Med.* **377**, 2010–2011 (2017).
8. Weenk, M. *et al.* Continuous monitoring of vital signs using wearable devices on the general ward: pilot study. *JMIR Mhealth Uhealth.* **5**, e91 (2017).
9. Schmier, J.K. & Halpern, M.T. Patient recall and recall bias of health state and health status. *Expert Rev. Pharmacoecon. Outcomes Res.* **4**, 159–163 (2004).
10. Lee, J.W. *et al.* Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm. Res.* **23**, 312–328 (2006).
11. National Center for Biotechnology Information. BEST (Biomarkers, EndpointS, and other Tools) Resource <https://www.ncbi.nlm.nih.gov/books/NBK326791/>. (January 2016). Accessed November 30, 2018.
12. US Food and Drug Administration. 510(k) summary. <https://www.accessdata.fda.gov/cdrh_docs/pdf/K983533.pdf>. (January 2014). Accessed November 30, 2018.
13. US Food and Drug Administration. 510(k) summary. <https://www.accessdata.fda.gov/cdrh_docs/pdf13/K132447.pdf>. (April 2014). Accessed November 30, 2018.
14. Bland, J.M. & Altman, D.G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–160 (1999).
15. Sarzynski, M.A. *et al.* Measured maximal heart rates compared to commonly used age-based prediction equations in the heritage family study. *Am. J. Hum. Biol.* **25**, 695–701 (2013).
16. Thurman, S.M. *et al.* Individual differences in compliance and agreement for sleep logs and wrist actigraphy: a longitudinal study of naturalistic sleep in healthy adults. *PLoS One* **13**, e0191883 (2018).
17. Bracci, M. *et al.* Peripheral skin temperature and circadian biological clock in shift nurses after a day off. *Int. J. Mol. Sci.* **17**, (2016).
18. Degaute, J.P., van de Borne, P., Linkowski, P. & Van Cauter, E. Quantitative analysis of the 24-hour blood pressure and heart rate patterns in young men. *Hypertension* **18**, 199–210 (1991).
19. Smith, I., MacKay, J., Fahrid, N. & Krucheck, D. Respiratory rate measurement: a comparison of methods. *Br. J. Healthcare Assistants* **5**, 18–22 (2011).
20. Xu, X., Karis, A.J., Buller, M.J. & Santee, W.R. Relationship between core temperature, skin temperature, and heat flux during exercise in heat. *Eur. J. Appl. Physiol.* **113**, 2381–2389 (2013).
21. Choi, J.K., Miki, K., Sagawa, S. & Shiraki, K. Evaluation of mean skin temperature formulas by infrared thermography. *Int. J. Biometeorol.* **41**, 68–75 (1997).
22. Barrett, P.M. *et al.* Comparison of 24-hour Holter monitoring with 14-day novel adhesive patch electrocardiographic monitoring. *Am. J. Med.* **127**, 95e11–95e97 (2014).
23. Steinhubl, S.R. *et al.* Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. *JAMA* **320**, 146–155 (2018).
24. Tsien, C.L. & Fackler, J.C. Poor prognosis for existing monitors in the intensive care unit. *Crit. Care Med.* **25**, 614–619 (1997).
25. Mitchell, I.A. *et al.* A prospective controlled trial of the effect of a multi-faceted intervention on early recognition and intervention in deteriorating hospital patients. *Resuscitation* **81**, 658–666 (2010).
26. Clinical Trial Transformation Initiative (CTTI). Program: Mobile Clinical Trials (MCT). <https://www.ctti-clinicaltrials.org/projects/mobile-technologies>. (July 2018). Accessed November 30, 2018.

npj | Digital Medicine

www.nature.com/npjdigitalmed

Corrected: Author Correction

PERSPECTIVE **OPEN**

Developing and adopting safe and effective digital biomarkers to improve patient outcomes

Andrea Coravos[1,2], Sean Khozin[3] and Kenneth D. Mandl[1,4]

Biomarkers are physiologic, pathologic, or anatomic characteristics that are objectively measured and evaluated as an indicator of normal biologic processes, pathologic processes, or biological responses to therapeutic interventions. Recent advances in the development of mobile digitally connected technologies have led to the emergence of a new class of biomarkers measured across multiple layers of hardware and software. Quantified in ones and zeros, these "digital" biomarkers can support continuous measurements outside the physical confines of the clinical environment. The modular software–hardware combination of these products has created new opportunities for patient care and biomedical research, enabling remote monitoring and decentralized clinical trial designs. However, a systematic approach to assessing the quality and utility of digital biomarkers to ensure an appropriate balance between their safety and effectiveness is needed. This paper outlines key considerations for the development and evaluation of digital biomarkers, examining their role in clinical research and routine patient care.

npj Digital Medicine (2019)2:14 ; https://doi.org/10.1038/s41746-019-0090-4

INTRODUCTION

Biomarkers are characteristics (such as a physiologic, pathologic, or anatomic characteristic or measurement) that are objectively measured and evaluated as an indicator of normal biologic processes, pathologic processes, or biological responses to a therapeutic intervention.[1] Building on this standard definition, we describe an emerging class of biomarker, the "digital biomarker", which has important implications for both clinical trials and clinical care. "Digital" refers to the method of collection as using sensors and computational tools, generally across multiple layers of hardware and software. The measurements are often made outside the physical confines of the clinical environment using home-based connected products[2] including wearable, implantable, and ingestible devices and sensors. Digital biomarkers span a broad range of diagnostic and prognostic measurements (Table 1). We discuss development and evaluation of the digital biomarkers, outlining opportunities and challenges associated with their use in clinical research and routine care. As remote monitoring of digital biomarkers becomes increasingly prevalent, we discuss the challenges to patient privacy and patient autonomy.

Just as clinicians must evaluate a drug's safety and effectiveness by critically appraising clinical trials, they will increasingly need to know how to evaluate, select, and "prescribe" digital health tools and biomarkers. Some biomarkers are immediately familiar to patients or physicians as they are digitized versions of well-established metrics—for example, glucometer readings transmitted by Bluetooth, or the timed six-minute walk test measured with the smartphone's built-in gyroscope and accelerometer. Others, such as the smartphone-derived tapping test for Parkinson's disease severity, are novel and evolving.[3] Digital biomarkers are an essential component in autoregulated closed loop systems. For example, in an "artificial pancreas" model, a continuous glucose sensor linked to an insulin pump can automatically dose insulin in patients with diabetes.[4]

THE ANATOMY AND EVALUATION OF DIGITAL BIOMARKERS

Measurements

An input layer such as a camera, microphone, or sensor captures a digital biomarker signal. For example, photoplethysmographs measure blood volume changes in the microvasculature using an optical sensor placed on the skin surface. A signal processing layer, typically an algorithm, converts the input signal into actionable metrics (e.g., oxygen saturation and/or heart rate), or digital biomarkers. Although measuring blood volume changes using photoplethysmography is widely accepted in medical practice, the interplay among hardware, sensors, and algorithms can make the evaluation of emerging digital biomarkers difficult. There are several challenges in deciding not only whether a digital biomarker is valid, but equally important, whether it is "fit-for-purpose", meaning that the product has an explicit context of use, meets appropriate requirements for accuracy and precision, and is accompanied by the metadata needed for analysis and interpretation.[5]

Verification

Analytical verification uses engineering bench tests to ensure that the product is measuring and storing values accurately by confirming the tool's accuracy, precision, and reliability. Confidence in the performance of digital biomarkers is an important consideration for researchers, clinicians, and patients. For example, the verification step ensures that the translation from raw data, e.g., that a heart rate sensor measuring electrical potential in millivolts, faithfully converts that signal into an accurate heart rate, expressed in beats per unit of time.

---

[1]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, USA; [2]Harvard-MIT Center for Regulatory Science, Boston, MA, USA; [3]Food and Drug Administration, Silver Spring, MD, USA and [4]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA
Correspondence: Kenneth D. Mandl (kenneth_mandl@harvard.edu)

Received: 23 November 2018 Accepted: 15 February 2019

Table 1. Digital biomarker examples

Category[a]	Definition[a]	Example[a]	Corresponding Digital Biomarker Examples
Susceptibility/Risk Biomarker	A biomarker that indicates the potential for developing a disease or medical condition in an individual who does not currently have clinically apparent disease or the medical condition.	Breast Cancer genes 1 and 2 (BRCA1/2) mutations may be used as a susceptibility/risk biomarker to identify individuals with a predisposition to develop breast cancer.	[*] Detect cognitive changes in healthy subjects at risk of developing Alzheimer's disease using a video game platform.[18] [**] Classify adults at high risk of late-onset Alzheimer's disease using computerized cognitive testing.[19] [*] Reduce key risk metrics for anterior cruciate ligament injury during jump landings using inertial sensor-based feedback.[20]
Diagnostic Biomarker	A biomarker used to detect or confirm the presence of a disease or condition of interest or to identify individuals with a subtype of the disease.	Repeated blood pressure readings obtained outside the clinical setting in adults 18 years and older may be used as a diagnostic biomarker to identify those with essential hypertension.	[*] Diagnose ADHD in children using eye vergence metrics.[21] [*] Detect arrhythmias using convolutional neural networks and a wearable single-lead heart monitor.[22] [*] Detect depression and Parkinson's disease using vocal biomarkers.[23] [*] Diagnose asthma and respiratory infections using smartphone-recorded cough sounds.[24]
Monitoring Biomarker	A biomarker measured serially for assessing the status of a disease or medical condition or for evidence of exposure to (or effect of) a medical product or an environmental agent.	Prostate-specific antigen (PSA) may be used as a monitoring biomarker when assessing disease status or burden in patients with prostate cancer.	[**] Monitor signs of Parkinson's disease using smartphone-based measurements.[25] [*] Quantify Parkinson's disease severity using smartphones and machine learning.[3] [**] Track time and location of short-acting beta-agonist inhaler use using an attached wireless sensor.[26] [*] Predicting sleep/wake patterns from a 3-axis home-based accelerometer using deep learning.[27] [*] Detection of nocturnal scratching movements in patients with atopic dermatitis using accelerometers and recurrent neural networks.[28]
Prognostic Biomarker	A biomarker used to identify the likelihood of a clinical event, disease recurrence, or progression in patients who have the disease or medical condition of interest.	Increasing prostate-specific antigen (PSA) may be used as a prognostic biomarker when evaluating patients with prostate cancer during follow-up, to assess the likelihood of cancer progression.	Stratify mental health conditions and predict remission using passively collected smartphone data.[29] Detect post-acute care deterioration in patients at home, applying machine learning to multi-sensor digital ambulatory monitoring.[30]
Predictive Biomarker	A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from exposure to a medical product or an environmental agent.	Human leukocyte antigen allele (HLA)–B*5701 genotype may be used as a predictive biomarker to evaluate human immunodeficiency virus (HIV) patients before abacavir treatment, to identify patients at risk for severe skin reactions.	Predict autism risk in the siblings of children with autism, using an EEG biomarker.[31] Detect asymptomatic atrial fibrillation (AF) as a stroke risk factor, remotely through a connected device.[32]
Pharmaco-dynamic/ Response Biomarker	A biomarker used to show that a biological response has occurred in an individual who has been exposed to a medical product or an environmental agent.	Blood pressure may be used as a pharmacodynamic/response biomarker when evaluating patients with hypertension, to assess response to an antihypertensive agent or sodium restriction.	Measure cognitive performance with the Cambridge Neuropsychological Test Automated Battery (CANTAB) to test the effects of erythropoietin.[33] Measure blood pressure using a digital sphygmomanometer to assess response to antihypertensive therapy.[34]

[a] Selected from the FDA-NIH "Biomarkers, EndpointS, and other Tools" (BEST) classification for traditional biomarkers
[*] Digital biomarker under development
[**] Digital biomarker in use (in a clinical trial or an FDA cleared/approved digital health product, or a digital health app in use not requiring approval)

Validation

As with diagnostics, the performance of digital biomarker algorithms may vary across different patient populations, producing different rates of false-positive or false-negative outputs in different groups. Validation addresses whether the measurement is applicable in the target population and context of use,[6] which would render digital biomarker "fit for purpose". For example, a tool measuring sleep and waking periods perform against

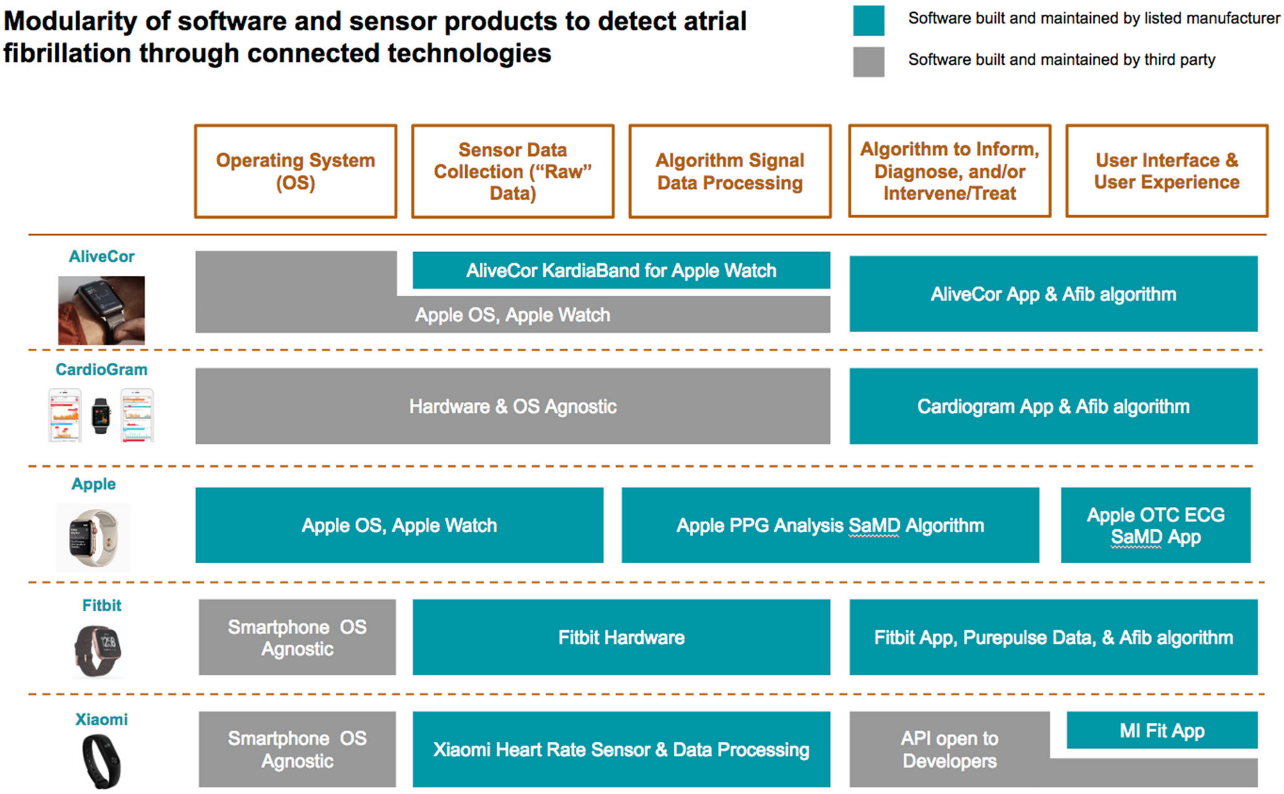


Fig. 1 Digital biomarker products. Five products, all detecting a similar digital endpoint, are constructed with differing, modular approaches. In the first column are five products to detect atrial fibrillation: AliveCor, CardioGram, Apple Watch plus ECG App, Fitbit, and Xiaomi. Across the top, are major software modules comprising the product, from the operating system on the left to the user interface on the right. Some modules are created by the product manufacturer and others by a third party. If the listed organization manufactures the component, the module is represented in green. If instead it is created by a different party, the color is gray. These differently composed products require different strategies for verification, validation, and likely also regulatory clearance. Figures are reused with permission from the copyright owners, and the Apple watch image is Courtesy of Apple Inc

polysomnography may perform differently in a patient population with insomnia versus sleep apnea versus healthy volunteers.

Modularity

Digital biomarker products can be composed of multiple individual software and hardware components. When the components are interoperable, they can be mixed and matched as modular components to assemble a diverse array of offerings. For example, the US Food and Drug Administration (FDA) recently approved the Dexcom integrated continuous glucose monitoring system as the first type of continuous glucose monitoring system that can be used in a modular fashion with other compatible medical devices and electronic interfaces, including automated insulin dosing systems and diabetes management devices.[7]

Software and hardware manufacturers have started to specialize in modular pieces of a connected product's data flow tool chain (Fig. 1).

Regulation of modular components

The FDA regulatory process can often address particular, modular, components along a digital biomarker's measurement apparatus. The FDA is piloting a program that would "pre-certify" companies and their policies[8] in order to offer a streamlined path to market for their product-level approvals and modifications.

Historically, most of the software-products have been categorized as software in a medical device (SiMD), which operates the device and sensors (e.g., firmware). More recently, digital biomarker components are categorized as software as a medical device (SaMD) solutions. SaMDs can perform a medical function without being part of a hardware medical device (e.g., machine-learning based tools in mobile apps[8]) have novel properties and potential for wider adoption. Definitions distinguishing SaMD from SiMD are evolving. The FDA recently cleared two SaMDs compatible with the Apple Watch for detection of atrial fibrillation. The first is an "over the counter" electrocardiogram app for display of atrial fibrillation[9] and the second can notify the user of an irregular rhythm.[10] The hardware, the Apple Watch, serves as a component supporting digital biomarker measurement. The Apple Watch over the counter EKG app and irregular rhythm notifications are FDA cleared as SaMDs.

While modularity enables mixing and matching across a variety of components, it can also be a source of potential error. For example, performance changes to an operating system may affect the speed of computation[11] and, for example, corrupt measurement of a Parkinson's tapping test, which uses a smartphone to calculate a digital biomarker based on timed reaction.

POTENTIAL BENEFITS AND RISKS OF DIGITAL BIOMARKERS

As new modalities are incorporated into connected devices, mobile apps, and software products for patients at home, a natural area of growth in biomarker collection is remote collection of patient-generated measurements. As digital biomarkers are increasingly used as endpoints in clinical trials, we anticipate that clinicians will have a growing number of validated means of gathering clinical insights on patients remotely. However, incorporation of these tools in clinical research is dependent on accelerating the development of new study designs such as those employed in decentralized clinical trials, where many of the trial participant touchpoints occur at home.[12] Furthermore, verification and validation of digital biomarkers require a uniquely collaborative approach, with engineering, data science, health information technology, and clinical research functions tightly coordinated as integrated multidisciplinary units.

New digital biomarkers are directly targeting clinical management. The Empatica Embrace Watch, for example, is a “smartband” wrist-device that measures sympathetic nervous impulses at the skin and infers parasympathetic activity from heart rate variation. Its algorithm detects seizures and its associated app suite can alert care providers. There are many examples of digital biomarkers in use or actively under development today, as well as computational metrics with potential for development into digital biomarkers (Table 1). We expect that as digital biomarkers become increasingly used in clinical trials, patient and physician adoption will increase in care and self-management. Digital tools also allow deep collection of data on individual trial participants as well as patients in clinical settings, thereby providing an opportunity for “N of 1” clinical investigations, the cornerstone of evidence generation for personalization of care.

As new platforms for connected technologies emerge, “composite” biomarkers simultaneously incorporating multi-sourced physiologic parameters (e.g., blood pressure, heart rate, and oxygen saturation) and patient-reported information can have higher diagnostic and prognostic value. With more data, an algorithm’s accuracy improves. For example, incorporation of the user’s height, weight, age, and gender increases step count accuracy, because a 25-year-old’s gait is not equivalent to that of an 80-year-old. Availability of contextual information will enable more personalized algorithms (e.g., a step count algorithm designed for a population with late-stage Parkinson’s), and also can combine data sources to create novel measures for conditions that have historically struggled to have meaningful endpoints (e.g., brain and nervous system disorders).

Ensuring privacy and autonomy is paramount as digital biomarkers are incorporated into care and self-management, and incentive programs encouraging wellness and treatment plan adherence. While healthcare delivery organizations using digital biomarkers are of course Health Insurance Portability and Accountability Act (HIPAA) covered entities, when citizens engage directly with the technologies or technology companies, HIPAA does not apply.[13] Social media and targeted advertising platforms typically employ end-user-license agreements and terms of service to outline data-sharing rights and privacy policies. However, like informed consent, health data rights should cover a continuum of activities over time. Therefore, data use agreements for digital biomarker development should contain clear statements on conditions for data usage especially for tools that collect near-continuous data, like movement, voice, and other sensitive biometric states.

Connected software products may pose cybersecurity challenges exposing trial participants and patients to privacy breaches or even safety risks. Just as HIPAA and the Common Rule are written to protect a patient’s medical record data and biospecimens, nascent efforts are building protections for digital “specimens”. New frameworks are emerging around the security,[14] ethics,[13] and informed consent challenges,[15] of digital phenotyping technologies.[16] One approach—a promising one for tracking security vulnerabilities and issues of performance, transparency, and accuracy—would require software manufacturers to provide, in premarket submission to the FDA, a “Software Bill of Materials” which is analogous to the ingredient list for a medication.[17]

A challenge to the evaluation of algorithms is that many are proprietary, patented or are trade secrets. For example, the AliveCor, Cardiogram, and Apple atrial detection algorithms and training data sets, for example, are not published. Instead, these companies offer a textual description of what the code does. The Empatica epilepsy monitor, for example, does not readily output raw signal, but instead, only the processed output interpreted by its proprietary algorithm. Hence the impact on a population of a digital biomarker-driven clinical management plan may not always be transparent to patients and clinicians. Testing characteristics, including selected thresholds for action, sensitivity, and specificity should be made transparent to the healthcare professional, regulators, and trial participant and patient users of digital biomarkers.

CONCLUSION

In recent years, digital biomarker development has begun integration into translational and clinical research. An increasing number of industry and academic investigators are at the leading edge of a new wave of innovations.

To accrue maximum benefit to the patient, a safe and effective digital biomarker ecosystem requires transparency of the algorithms, interoperable components with open interfaces to accelerate the development of new multicomponent systems, high integrity measurement systems. The time is now to give forethought to strong incentive structures to promote the safe and effective use of digital biomarkers. Generally, the verification and validation of a digital biomarker should be not construed as a one-time process, but rather, a learning digital health system should continuously collect data and handle modifications and updates overtime. Industry, researchers, regulators, clinicians, and patients have a joint responsibility to design such a learning system that can improve digital biomarker products, empower patients, and improve health and healthcare delivery for everyone

ACKNOWLEDGEMENTS

This work was supported by the PrecisionLink initiative at Boston Children’s Hospital and by a grant from the National Institutes of Health/NIGMS R01GM104303. Andrea Coravos acknowledges support from Harvard-MIT Center for Regulatory Science. For suggesting compelling examples of digital biomarkers for the Table, we gratefully acknowledge Jessie Bakker, MS PhD, Brandon Ballinger, Chris Benko, Brian M. Bot, Jeffrey D. Bower, PhD, Ray Dorsey, MD, Ariel V. Dowling, PhD, Robert Ellis, PhD, Luca Foschini, PhD, Robert Furberg, PhD, Jennifer C. Goldsack, MChem MA MBA, Ankit Gordhandas, Elena Izmailova, PhD, Matthew Johnson, Daniel Karlin, MD MA, Ashley Mateus, PhD, Donald McLaren, PhD, Shyamal Patel, PhD, Barry Peterson, PhD, Ariel Dora Stern, PhD, Iain Simpson, PhD, A. Sofia Warner, MD, William Wood, MD MPH, and Noah Zimmerman, PhD.

AUTHOR CONTRIBUTIONS

All authors (A.C., K.D.M., and S.K.) equally contributed to conceptualization, drafting, editing, and finalizing the manuscript. A.C. created the figure. All authors meet the following criteria: (1) Substantial contributions to the conception or design of the work or the acquisition, analysis or interpretation of the data; (2) Drafting the work or revising it critically for important intellectual content; (3) Final approval of the completed version; and (4) Accountability for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

ADDITIONAL INFORMATION

Competing interests: A.C. has been developing an open-source pre-competitive digital biomarker catalog, at Elektra Labs, a startup company, with funding from the Harvard Business School, the NSF, and the Mount Sinai School of Medicine. The other authors declare no competing interests.

REFERENCES

1. 114th Congress. *H.R.34—21st Century Cures Act* (2015–2016). https://www.nejm.org/doi/full/10.1056/NEJMp1615745.

2. Byrom, B. et al. Selection of and evidentiary considerations for wearable devices and their measurements for use in regulatory decision making: recommendations from the ePRO Consortium. *Value Health* **21**, 631–639 (2018).
3. Zhan, A. et al. Using smartphones and machine learning to quantify Parkinson disease severity: the mobile Parkinson disease score. *JAMA Neurol.* **75**, 876–880 (2018).
4. Kovatchev, B. The artificial pancreas in 2017: the year of transition from research to clinical practice. *Nat. Rev. Endocrinol.* **14**, 74–76 (2018).
5. Atreja, A. et al. *Mobilizing mHealth Innovation for Real-World Evidence Generation* (Duke Margolis Center for Health Policy, https://healthpolicy.duke.edu/sites/default/files/atoms/files/duke-margolis_mhealth_action_plan.pdf, 2018).
6. Izmailova, E. S., Wagner, J. A. & Perakslis, E. D. Wearable devices in clinical trials: hype and hypothesis. *Clin. Pharmacol. Ther.* **104**, 42–52 (2018).
7. Parmar, A. *FDA Clears New Dexcom CGM that Requires No Patient Calibration Earlier than Expected*. https://medcitynews.com/2018/03/fda-clears-new-dexcom-cgm-requires-no-patient-calibration-earlier-expected/ (2018).
8. Shuren, J., Patel, B. & Gottlieb, S. FDA regulation of mobile medical apps. *JAMA* **320**, 337–338 (2018).
9. U.S. Food and Drug Administration. *ECG App: Electrocardiograph Software for Over-the-Counter Use.* https://www.accessdata.fda.gov/cdrh_docs/pdf18/DEN180044.pdf (2018).
10. U.S. Food and Drug Administration. *Irregular Rhythm Notification Feature: Photoplethysmograph Analysis Software for Over-the-Counter Use.* https://www.accessdata.fda.gov/cdrh_docs/pdf18/DEN180042.pdf (2018).
11. Apple. *A Message to Our Customers about iPhone Batteries and Performance.* https://www.apple.com/iphone-battery-and-performance/ (2017).
12. Steinhubl, S. R., McGovern, P., Dylan, J. & Topol, E. J. The digitised clinical trial. *Lancet* **390**, 2135 (2017).
13. Martinez-Martin, N., Insel, T. R., Dagum, P., Greely, H. T. & Cho, M. K. Data mining for health: staking out the ethical territory of digital phenotyping. *npj Digit. Med.* **1**, 68 (2018).
14. Clinical Trials Transformative Initiative. *CTTI Unveils Recommendations for Using Mobile Technologies in Clinical Research.* https://www.ctti-clinicaltrials.org/news/ctti-unveils-recommendations-using-mobile-technologies-clinical-research (2018).
15. Sage Bionetworks. *Elements of Informed Consent.* http://sagebionetworks.org/in-the-news/elements-informed-consent/ (2018).
16. Torous, J., Onnela, J. P. & Keshavan, M. New dimensions and new tools to realize the potential of RDoC: digital phenotyping via smartphones and connected devices. *Transl. Psychiatry* **7**, e1053 (2017).
17. U.S. Food and Drug Administration. *Medical Device Safety Action Plan: Protecting Patients, Promoting Public Health.* https://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM604690.pdf. Accessed 25 Feb 2019.
18. Gold, M. et al. Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. *Alzheimers Dement.* **4**, 234–242 (2018).
19. Ritchie, K. et al. The midlife cognitive profiles of adults at high risk of late-onset Alzheimer's disease: the PREVENT study. *Alzheimers Dement.* **13**, 1089–1097 (2017).
20. Dowling, A. V., Favre, J. & Andriacchi, T. P. Inertial sensor-based feedback can reduce key risk metrics for anterior cruciate ligament injury during jump landings. *Am. J. Sports Med.* **40**, 1075–1083 (2012).
21. Varela Casal, P. et al. Clinical validation of eye vergence as an objective marker for diagnosis of ADHD in children. *J. Atten. Disord.* https://doi.org/10.1177/1087054717749931 (2018).
22. Rajpurkar, P., Hannun, A., Masoumeh, H., Bourn, C. & Ng, A. Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv preprint arXiv:1707.01836, https://arxiv.org/pdf/1707.01836.pdf (2017).
23. Gosh, S. S. & Ciccarelli, G. Speaking one's mind: vocal biomarkers of depression and Parkinson disease. *J. Acoust. Soc. Am.* **139**, 2193 (2016).
24. RespApp. *Diagnosing Respiratory Disease in Children Using Cough Sounds 2 (SMARTCOUGH-C-2).* https://www.clinicaltrials.gov/ct2/show/NCT03392363 (2018).
25. Sage Bionetworks. *Sage Bionetworks in Collaboration with The Michael J. Fox Foundation Announce Winners in the DREAM Parkinson's Disease Digital Biomarker Challenge.* https://www.businesswire.com/news/home/20180117006187/en (2018).
26. Barrett, M. A. et al. Effect of a mobile health, sensor-driven asthma management platform on asthma control. *Ann. Allergy Asthma Immunol.* **119**, 415–421 (2017).
27. Wolz, R., Munro, J., Guerrero, R., Hill, D. L. & Dauvilliers, Y. Predicting sleep/wake patterns from 3-axis accelerometry using deep learning. *Alzheimer Dement.* **13**, P1012 (2017).
28. Moreau, A. et al. Detection of nocturnal scratching movements in patients with atopic dermatitis using accelerometers and recurrent neural networks. *IEEE J. Biomed. Health Inform.* **22**, 1011–1018 (2018).
29. Mindstrong Health. *Mindstrong Health and Takeda Partner to Explore Development of Digital Biomarkers for Mental Health Conditions*. https://www.prnewswire.com/news-releases/mindstrong-health-and-takeda-partner-to-explore-development-of-digital-biomarkers-for-mental-health-conditions-300604553.html (2018).
30. physIQ. *physIQ.* http://www.physiq.com/resources/.
31. Bosl, W. J., Tager-Flusberg, H. & Nelson, C. A. EEG analytics for early detection of autism spectrum disorder: a data-driven approach. *Sci. Rep.* **8**, 6828 (2018).
32. Halcox, J. P. J. et al. Assessment of remote heart rhythm sampling using the AliveCor Heart Monitor to screen for atrial fibrillation: The REHEARSE-AF Study. *Circulation* **136**, 1784–1794 (2017).
33. Kessing, L. V. *Effects of Erythropoietin on Cognition and Neural Activity in Bipolar Disorder (PRETEC-EPO).* https://clinicaltrials.gov/ct2/show/NCT03315897 (2017).
34. Padwal, R. S. Validation of the Omron HEM-9210T by the ANSI/AAMI/ISO 81060-2 with two novel cuffs: wide range and extra-large. *Blood Press Monit.* **22**, 379 (2017).

ARTICLE OPEN

Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants

Abhishek Pratap[1,2]✉, Elias Chaibub Neto[1], Phil Snyder[1], Carl Stepnowsky[3,4], Noémie Elhadad[5], Daniel Grant[6], Matthew H. Mohebbi[7], Sean Mooney[2], Christine Suver[1], John Wilbanks[1], Lara Mangravite[1], Patrick J. Heagerty[8], Pat Areán[9] and Larsson Omberg[1]

Digital technologies such as smartphones are transforming the way scientists conduct biomedical research. Several remotely conducted studies have recruited thousands of participants over a span of a few months allowing researchers to collect real-world data at scale and at a fraction of the cost of traditional research. Unfortunately, remote studies have been hampered by substantial participant attrition, calling into question the representativeness of the collected data including generalizability of outcomes. We report the findings regarding recruitment and retention from eight remote digital health studies conducted between 2014–2019 that provided individual-level study-app usage data from more than 100,000 participants completing nearly 3.5 million remote health evaluations over cumulative participation of 850,000 days. Median participant retention across eight studies varied widely from 2–26 days (median across all studies = 5.5 days). Survival analysis revealed several factors significantly associated with increase in participant retention time, including (i) referral by a clinician to the study (increase of 40 days in median retention time); (ii) compensation for participation (increase of 22 days, 1 study); (iii) having the clinical condition of interest in the study (increase of 7 days compared with controls); and (iv) older age (increase of 4 days). Additionally, four distinct patterns of daily app usage behavior were identified by unsupervised clustering, which were also associated with participant demographics. Most studies were not able to recruit a sample that was representative of the race/ethnicity or geographical diversity of the US. Together these findings can help inform recruitment and retention strategies to enable equitable participation of populations in future digital health research.

npj Digital Medicine (2020)3:21 ; https://doi.org/10.1038/s41746-020-0224-8

INTRODUCTION

Traditional in-person clinical trials serve as the cornerstone of modern healthcare advancement. While a pivotal source of evidence generation for advancing clinical knowledge, in-person trials are also costly and time-consuming, typically running for at least 3–5 years from conception to completion, at a cost of millions of dollars per study. These timelines have often meant that promising treatments take years to get to dissemination and uptake, which can create unnecessary delays in advancing clinical practice. Additionally, clinical research suffers from several other challenges[1,2] including (1) recruiting sufficiently large and diverse cohorts quickly, and (2) tracking day-to-day fluctuations in disease severity that often go undetected in study-related intermittent protocolized in-clinic evaluations.[3,4] Scientists have recently turned to digital technology[5,6] to address these challenges, hoping to collect real-world evidence[7] from large and diverse populations to track long-term health outcomes and variations in disease trajectories at a fraction of the cost of traditional research.[8]

The global penetration[9] and high-frequency usage of smartphones (up to 4 h daily[10,11]) offer researchers a potentially cost-effective means to recruit a large number of participants into health research across the US (and the world).[12,13] In the last 5 years, investigators have conducted several large-scale studies[14–22] that deployed interventions[23,24] and operationally conduct clinical trials[25–27] using mobile technologies. These studies are able to recruit at-scale because participants can be identified and consented[28] to participate in the study without ever having stepped foot in a research lab, with significantly lower costs than conventional clinical trials.[23,24] Mobile technologies also allow investigators an opportunity to collect data in real-time based on people's daily lived experiences of the disease, that is, real-world data.[7] Rather than retrospectively asking people to recall their health over the past week or month, researchers using mobile technologies can assess participants frequently including outside clinic and at important points in time without having to rely on recall that is known to have bias.[29] While these recent studies show the utility of mobile technology, challenges in participant diversity and long-term participant retention still remain a significant problem.[30]

Digital studies continue to suffer from participant retention problems that also plagued internet-based studies[31,32] in the early 2000s.[33–35] However, our understanding of factors impacting retention in remote research remains to be limited. High levels of user attrition combined with variations in long-term app usage may result in the creation of a study cohort that does not represent the characteristics of the initially recruited study population with regards to demographics and disease status. This has called into question the reliability and utility of the collected data from these studies.[36] Of note, the representativeness of remote study cohort (e.g., demographics, geographical diversity, etc) may vary based on the study design and inclusion criteria. Many large-scale digital health studies enroll participants from a general population, where anyone eligible with or without target disease of interest can self-select to join the study. Such strategies

---

[1]Sage Bionetworks, Seattle, WA, USA. [2]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, USA. [3]University of California, San Diego, CA, USA. [4]American Sleep Apnea Association, Washington, DC, USA. [5]Columbia University, New York, NY, USA. [6]Novartis Pharmaceutical Corporation, East Hanover, NJ, USA. [7]GoodRx, Santa Monica, CA, USA. [8]Department of Biostatistics, University of Washington, Seattle, WA, USA. [9]Department of Psychiatry & Behavioral Sciences, University of Washington, Seattle, WA, USA. ✉email: apratap@sagebionetworks.org

may be prone to selection and ascertainment biases.[36] Similarly, cohorts in studies targeting a population with a specific clinical condition of interest may need to be evaluated in the context of the clinical and demographic characteristics of the underlying population with that disease. Evaluation of participant recruitment and retention from large-scale remotely collected data could help detect confounding characteristics that may be present and which have been shown to severely impact the generalizability of the resulting statistical inference.[36,37] Participant retention may also be partially dependent on the engagement strategies used in remote research. While most studies assume participants will remain in a study for altruistic reasons,[38] other studies provide compensation for participant time,[39] or leverage partnerships with local community organizations, clinical registries, and clinicians to encourage participation.[23,24] Although monetary incentives are known to increase participation in research,[40] we know little about the relative impact of demographics and different recruitment and engagement strategies on participant retention, especially in remote health research.

The purpose of this study is to document the drivers of retention and long-term study mobile application usage in remote research. To investigate these questions, we have compiled user-engagement data from eight digital health studies that enrolled more than 100,000 participants from throughout the US between 2014–2019. These studies assessed different disease areas including asthma, endometriosis, heart disease, depression, sleep health, and neurological diseases. While some studies enrolled participants from the general population (i.e., with and without disease of interest) others targeted a specific subpopulation with the clinical condition of interest. For analysis, we have combined individual participant data across these studies. Analysis of the pooled data considers overall summaries of demographic or operational characteristics while accounting for study heterogeneity in retentions (see Methods for further details on individual studies and analytical approach). The remote assessments in these studies consisted of a combination of longitudinal subjective surveys and objective sensor-based tasks including passive data[41] collection. The diversity of the collected data allows for a broad investigation of different participant characteristics and engagement strategies that may be associated with higher retention in the collected real-world data.

RESULTS

Participant characteristics

The combined user-activity data from eight digital health studies resulted in a pool of 109,914 participants who together completed ~3.5 million tasks on more than 850,000 days (Table 1). The demographics of participants across studies (Table 2) varied widely in part due to study-specific eligibility criteria, which may impact the underlying characteristics of the recruited population. Except for three studies (Brighten, Phendo, and Start) that aimed to recruit people with a specific clinical condition of interest, the rest of the five studies enrolled people from the general population with and without the target disease of interest. The majority of participants were between 17–39 years (Median percent of subject age across studies 17–39 = 65.2%, Range = 37.4–91.5%) with those 60 years and older being the least represented (Median age percent greater than 60 = 6%, Range = 0–23.3%). The study samples also had a larger proportion of Females (Median = 56.9%, Range = 18–100%). A majority of recruited participants were Non-Hispanic Whites (Median = 75.3%, Range = 60.1–81.3%) followed by Hispanic/Latinos (Median = 8.21%, Range = 4.79–14.29%) and African-American/Blacks (Median = 3.45%, Range = 2–10.9%). The race/ethnic and geographical diversity of the present sample showed a marked difference from the general population of the US. Minority groups were under-represented in the present sample with Hispanic/Latinos and African-America/Black showing a substantial difference of −8.1% and −9.2% respectively compared to the 2010 US census data s (Table 2, Fig. 1b). Similarly, the median proportion of recruited participants per state also showed notable differences from the state's population proportion of the US (Fig. 1a).

Participant retention

As is the nature of these studies, participants were required to complete all health assessments and other study-related tasks (e.g., treatments) through a mobile application (app) throughout the length of the study. The median time participants engaged in the study in the first 12 weeks was 5.5 days of which in-app tasks were performed on 2 days (Table 2). Higher proportions of active tasks were completed by participants during the evening (4–8 PM) and night (8–12 Midnight) hours (Fig. 2a). Across the studies, the median retention time varied significantly ($P < 1e\text{-}16$) between 2 and 12 days with the Brighten study being an outlier with higher median retention of 26 days (Fig. 2b). A notable increase in median retention time was seen for sub-cohorts that continue to engage with the study apps after day one and beyond (Fig. 2c). For example, the median retention increased by 25 days for the sub-cohort that was engaged for the first 8 days. The participant retention also showed a significant association with participant characteristics. While older participants (60 years and above) were the smallest proportion of the sample, they remained in the study for a significantly longer duration (Median = 7 days, $P < 1e\text{-}16$) compared to the majority younger sample (17–49 years) (Fig. 2d). Participants declared gender showed no significant difference in retention ($P = 0.3$). People with clinical conditions of interest to the study (e.g., heart disease, depression, multiple sclerosis) remained in the studies for a significantly longer time (Median = 13 days, $P < 1e\text{-}16$) compared to participants that were recruited

Table 1. Summary of user-engagement data compiled from eight digital health studies.

Study	Disease focus/study type	Study period	Number of participants	Total participant days	Active tasks completed
Start	Antidepressant Efficacy–Observational	Aug, 2015–Feb, 2018	42,704	280,489	1,219,656
MyHeartCounts	Cardiovascular Health–Observational	Mar, 2015–Oct, 2015	26,902	165,455	305,821
SleepHealth	Sleep Apnea–Observational	Jul, 2015–Jun, 2019	12,914	99,696	401,628
mPower	Parkinson's–Observational	Mar, 2015–Jun, 2019	12,236	104,797	568,685
Phendo	Endometriosis–Observational	Dec, 2016–Jul, 2019	7,802	81,938	735,778
Asthma	Asthma–Observational	Mar, 2015–Dec, 2016	5,875	77,815	175,699
Brighten	Depression–Randomized Control Trial	Jul, 2014–Aug, 2015	876	34,987	45,951
ElevateMS	Multiple Sclerosis–Observational	Aug, 2017–Jul, 2019	605	11,211	31,568
			109,914	856,388	3,484,786

Table 2. Summary of select participant demographics and study-app usage across the eight digital health studies.

	Asthma	Brighten	ElevateMS	mPower	MyHeartCounts	Phendo	SleepHealth	Start	Overall (median)
Age group									
N	2512	875	569	6810	1555	7484	12392	42690	
18–29 (%)	43.31	50.06	10.9	31.5	25.08	55.38	32.79	55.72	38
30–39 (%)	27.83	25.14	26.54	18.37	32.67	36.09	28.72	24.14	27.2
40–49 (%)	14.41	14.74	28.47	13.19	16.27	8.23	20.77	12.38	14.6
50–59 (%)	9.08	6.97	22.14	13.61	12.09	0.25	11	5.26	10
60 + (%)	5.37	3.09	11.95	23.33	13.89	0.04	6.72	2.51	6
Sex									
N	2509	875	329	6916	6976	7532	12558	42704	
Female (%)	39.58	77.83	74.16	28.93	18.94	100	29.14	75.86	56.9
Race									
N	3274	875	334	6884	4703	7530	5311	—	
Non-Hispanic White (%)	68.69	60.11	80.84	75.32	77.95	81.29	74.13	—	75.3
Hispanic/Latinos (%)	13.29	14.29	4.79	8.21	6.97	5.67	12.82	—	8.21
African-American/Black (%)	4.95	10.86	6.89	2.05	3.1	2.71	3.45	—	3.45
Asian (%)	4.98	8.23	2.99	8.4	7.72	2.79	5.87	—	5.9
Hawaiian or other Pacific Islander (%)	0.89	0.57	0	0.28	0.32	0	0.23	—	0.3
AIAN (%)	0.43	0.46	0	0.65	0.53	0.74	0.28	—	0.5
Other (%)	6.78	5.49	4.49	5.1	3.4	6.8	3.22	—	5.1
Duration in Study (Median ± IQR)	12 ± 38	26 ± 82	7 ± 45	4 ± 21	9 ± 19	4 ± 25	2 ± 8	2 ± 16	5.5
Days active tasks performed (Median ± IQR)	4 ± 12	14 ± 58	2 ± 8	2 ± 4	4 ± 7	2 ± 6	2 ± 4	2 ± 4	2

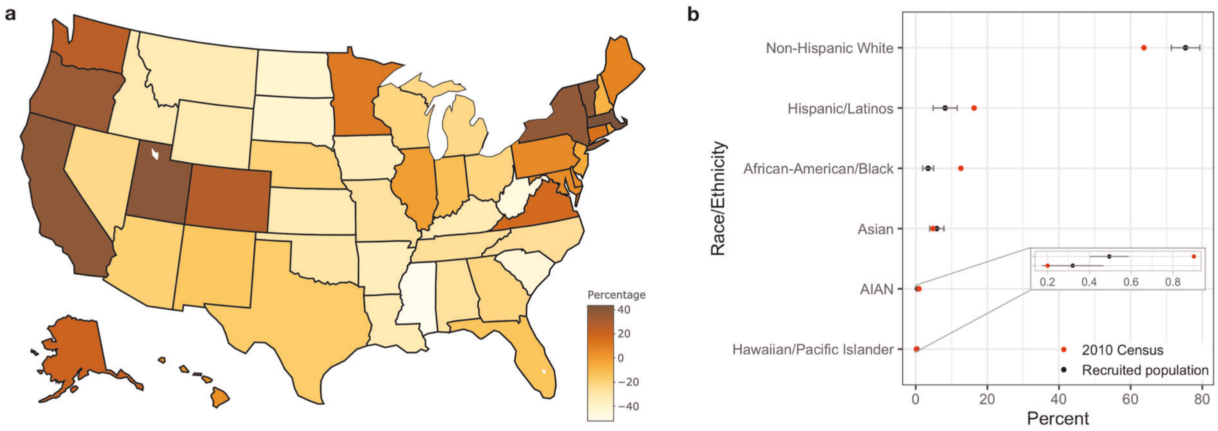


Fig. 1 Geographical and race/ethnic diversity of the recruited participants. a Map of US showing the proportion (median across the studies) of recruited participants relative to state's population proportion of the US and **b** Race/Ethnicity proportion of recruited participants compared to 2010 census data. The median value across the studies is shown by the black point with error bars indicating the interquartile range.

as non-disease controls(Median = 6 days) (Fig. 2e). Median retention time also showed a marked and significant increase of 40 days ($P < 1e\text{-}16$) for participants that were referred by a clinician to join one of the two studies (mPower and ElevateMS) (Median = 44 days) compared to participants who self-selected to join the same study (Median = 4 days) (Fig. 2f). See Supplementary Tables 1–6 for a further breakdown of survival analysis results. Sensitivity analysis by including participants with missing age showed no impact on the association of age with participant retention. However, participants with missing demographics showed variation in retention compared to participants who shared their demographics (Supplementary Fig. 1). This could be related to different time points at which demographic related questions were administered in individual studies.

Participant daily engagement patterns

The subgroup of participants who remain engaged with study apps for a minimum of 7 days, showed four distinct longitudinal engagement patterns (Fig. 3b) with the dedicated users in cluster C1, high utilizers(C2), moderate users(C3) and sporadic users(C4). The participants who did not participate for at least 7 days were placed in a separate group (abandoners, C5*) (See Supplementary Fig. 2 and Methods for cluster size determination and exclusion criteria details). The engagement and demographic characteristics

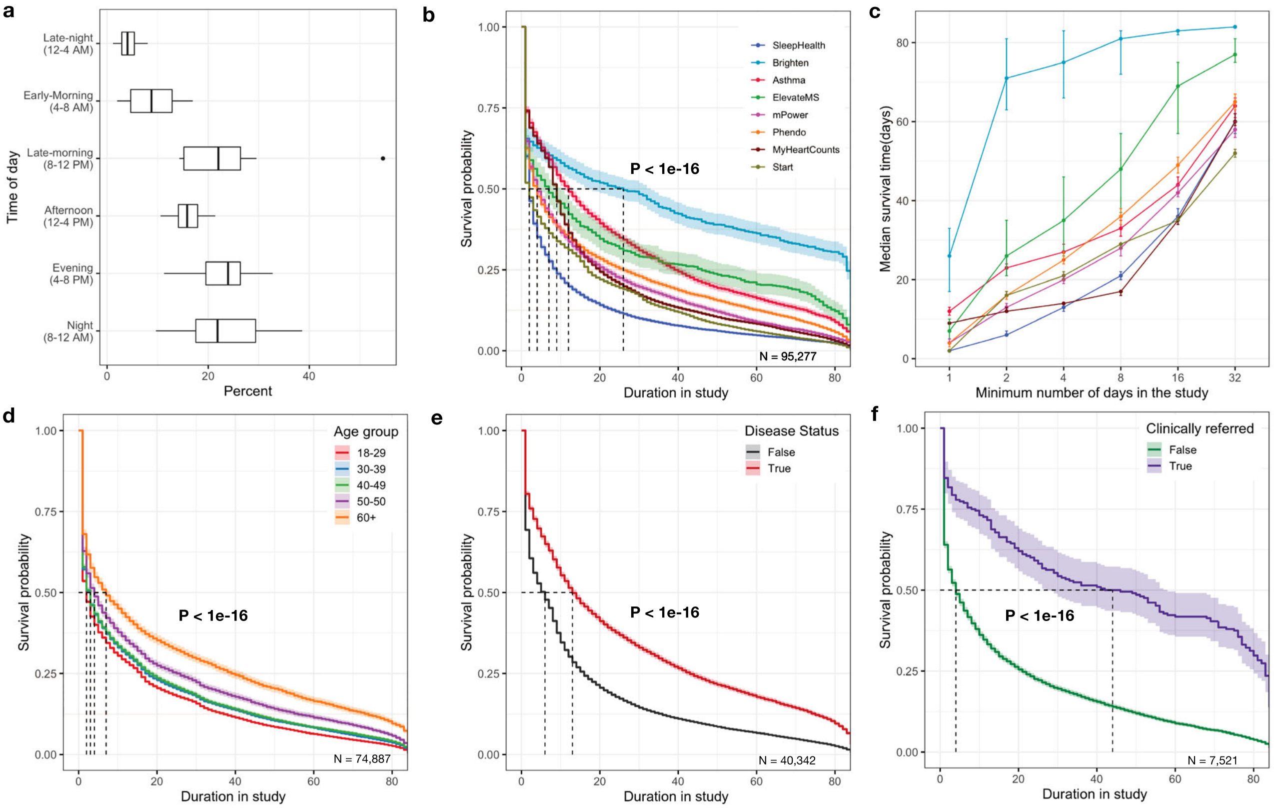


Fig. 2 Factors impacting participant retention in digital health studies. **a** Proportion of active tasks (N = 3.3 million) completed by participants based on their local time of day. The centerline of the boxplot shows the median value across the studies and upper and lower whisker corresponding to outlier point (1.5 times the interquartile range). **b** Kaplan Meir survival curve showing significant differences (P < 1e-16) in user retention across the apps. Brighten App where monetary incentives were given to participants showed the longest retention time (Median = 26 days, 95% CI = 17–33) followed by Asthma(Median = 12 days, 95% CI = 11–13), MyHeartCounts(Median = 9 days, 95% CI = 9–9), ElevateMS(Median = 7 days, 95% CI = 5–10), mPower(Median = 5 days, 95% CI = 4–5), Phendo(Median = 4 days, 95% CI = 3–4), Start(Median = 2 days, 95% CI = 2–2) and SleepHealth(Median = 2 days, 95% CI = 2–2), **c** Lift curve showing the change in median survival time (with 95% CI indicated by error bars) based on the minimum number of days(1–32) a subset of participants continued to use the study apps, Kaplan-Meier survival curve showing significant differences in user retention across **d** Age group, with 60 years and older using the apps for longest duration (Median = 7days, 95% CI = 6–8, P < 1e-16) followed by 50–59 years (Median = 4 days, 95% CI = 4–5) and 17–49 years (Median = 2–3 days, 95% CI = 2–3). **e** Disease status; participants reporting having a disease stayed active longer(N_{50} = 13days, 95% CI = 13–14) compared to people without disease(N_{50} = 6 days, 95% CI = 5–6) and finally **f** Clinical referral; Two studies (mPower and ElevateMS), had a subpopulation, that were referred to the study by clinicians and showed significantly (P < 1e-16) longer app usage period(Median = 44 days, 95% CI = 27–58) compared to self-referred participants with disease (N_{50} = 4 days, 95% CI = 4–4). For all survival curves the shaded region shows the 95% confidence limits based on the survival model fit.

across these five groups (C1-5*) varied significantly. Cluster 1 and 2 showed the highest daily app usage (Median app usage in the first 84 days = 96.4% and 63.1%, respectively) but also had the smallest overall proportion of participants (Median = 9.5%) with the exception of Brighten where 23.7% of study participants were in the dedicated users cluster C1. While daily app usage declined significantly for both moderate and sporadic clusters (C3–21.4% and C4–22.6%), the median number of days between app usage was significantly higher for participants in the sporadic C4 cluster (Median = 5 days) compared to cluster C3 (Median = 2 days). The majority of participants (median 54.6%) across the apps were linked to the abandoner group (C5*) with the median app usage of just 1 day (Fig. 4a, b). Furthermore, distinct demographic characteristics emerged across these five groups. Higher engagement clusters (C1–2) showed significant differences (P = 1.38e-12) in proportion of adults 60 years and above (Median range = 15.1–17.2% across studies) compared to lower engagement clusters C3–5*(Median range = 5.1–11.7% across studies) (Fig. 4c). Minority groups such as Hispanic/Latinos, Asians, and African-American/Black, on the other hand, were represented in higher proportions in the clusters (C3–5*)(P = 4.12e-10) with the least engagement (Fig. 4d) (See Supplementary Table 8 for further details).

DISCUSSION

Our findings are based on one of the largest and most diverse engagement dataset compiled to date. We identified two major challenges with remote data collection: (1) more than half of the participants discontinued participation within the first week of a study and the rates at which people discontinued was drastically different based on age, disease status, clinical referral, and use of monetary incentives, and (2) most studies were not able to recruit a sample that was representative of the Race/ethnicity or geographical diversity of the US. Although these findings raise questions about the reliability and validity of data collected in this manner, they also shed light on potential solutions to overcome biases in populations using a combination of different recruitment and engagement strategies.

One solution could be the use of a flexible randomized withdrawal design.[42] Temporal retention analysis (Fig. 2c) shows that a run-in period could be introduced in the research design, wherein participants who are not active in the study app in the

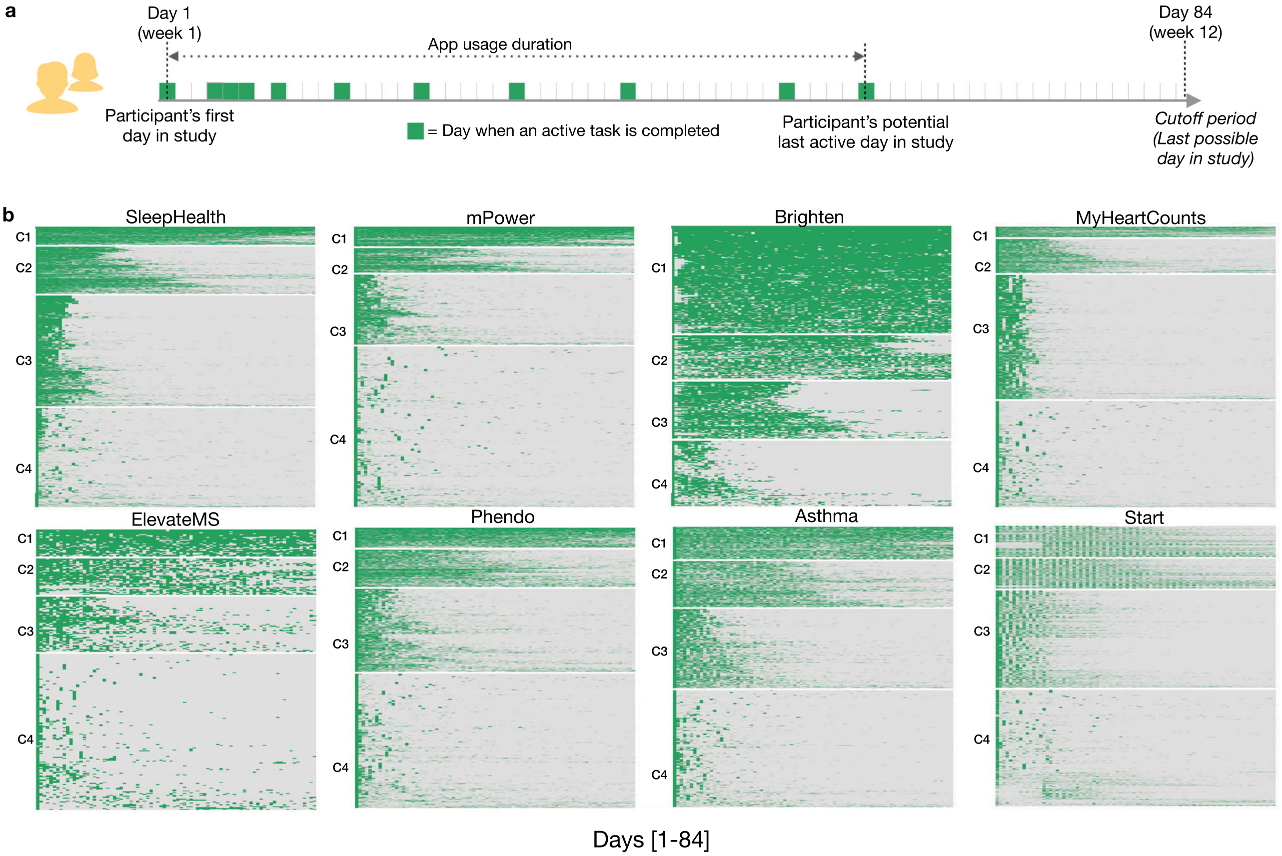


Fig. 3 Daily participant engagement patterns in digital health studies. a Schematic representation of an individual's in-app activity for the first 84 days. The participant app usage time is determined based on the number of days between the first and last day they perform an active task(indicated by the green box) in the app. Days active in the study is the total number of days a participant performs at least one active task (indicated by the number of green boxes). **b** Heatmaps showing participants in-app activity across the apps for the first 12 weeks (84 days), grouped into four broad clusters using unsupervised k-means clustering. The optimum number of clusters was determined by minimizing the within-cluster variation across different cluster sizes between 1–10. Seven out of eight studies indicated four clusters to be an optimum number using the elbow method. The heatmaps are arranged by the highest (C1) to the lowest user-engagement cluster (C4).

first week or two of the study can be excluded after enrollment but before the start of the actual randomized study. The resulting smaller but more engaged cohort will help increase the statistical power of the study but does not fix the potential for non-representative participant bias.[43]

Another solution is to rely on monetary incentives to enhance engagement. Although only one study paid participants, the significant increase in retention and the largest proportion of frequent app users indicate the utility of the fair-share compensation model[1,44,45] in remote research. Such "pay-for-participation" model could be utilized by studies that require long-term and frequent remote participation. Researchers conducting case-control studies should also plan to further enrich and engage the population without the disease. Studies run the risk of not collecting sufficient data from controls to perform case-control analysis with participants without disease seen to be dropping out significantly early. Similarly, more efforts[46–48] are needed to retain the younger population that although demonstrates large enrollment, also features a majority that drops out on day one.

Distinct patterns in daily app usage behavior, also shown previously,[49] further strengthen the evidence of unequal technology utilization in remote research. The majority of the participants found in the abandoners group (C5*) who dropped out of the study on day 1, may also reflect initial patterns in willingness to participate in research, in a way that cannot be captured by recruitment in traditional research. Put another way, although there is significant dropout in remote trials, these early dropouts may be able to yield very useful information about differences in people who are willing to participate in research and those who are not willing to participate. For decades clinical research has been criticized for its potential bias because people who participate in research may be very different from people who do not participate in research.[50–52] Although researchers will not have longitudinal data from those who discontinue participation early, the information collected during onboarding can be used to assess potential biases in the final sample and may inform future targeted retention strategies.

Only 1 in 10 participants were in the high app use clusters (C1-2), and these clusters tended to be largely Non-Hispanic whites and older adults. Minority and younger populations, on the other hand, were represented more in the clusters with the lowest daily app usage (Fig. 4d). The largest impact on participant retention (>10 times) in the present sample was associated with clinician referral for participating in a remote study. This referral can be very light touch in nature, for example in the ElevateMS study, it consisted solely of clinicians handing patients a flyer with information about the study during a regular clinic visit. This finding is understandable, given recent research[53] showing that the majority of Americans trust medical doctors.

For most studies, the recruited sample was also inadequately diverse highlighting a persistent digital divide[54] and continued challenges in the recruitment of racial and ethnic underserved communities.[55] Additionally, the underrepresentation of States in the southern, rural and midwest regions indicates that areas of the US that often bear a disproportionate burden of disease[56] are under-represented in digital research.[56–58] This recruitment bias

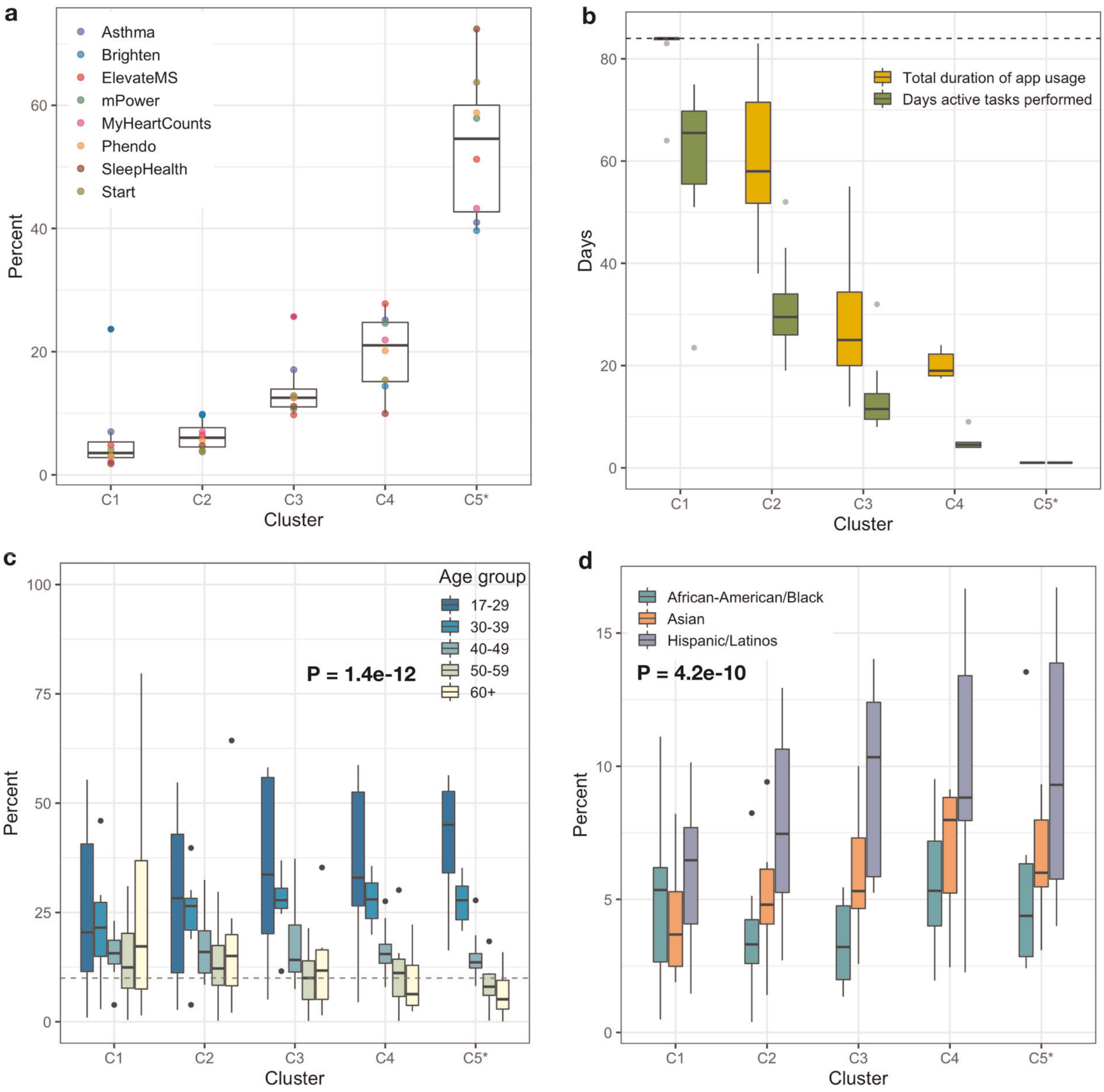


Fig. 4 Comparison of characteristics across participant engagement clusters. a Proportion of participants in each cluster across the study apps, **b** Participants total app usage duration(between 1–84 days) and the number of days participants completed tasks in the study apps, **c** Significant differences [F(4,163) = 18.5, P = 1.4e-12] in the age groups of participants across five clusters and **d** Significant differences [F(2,81) = 28.5, P = 4.1e-10] in proportion of minority population present in the five clusters. C5* cluster contains the participants that used the apps for less than a week and were removed from the clustering; however, they were added back for accurate proportional comparison of participants in each cluster. The centerline of the boxplots shows the median value across the studies for each cluster and upper and lower whisker corresponds to the outlier point that is at least 1.5 times the interquartile range.

could impact future studies that aim to collect data for health conditions that are more prevalent among certain demographic[59] and associated with geographic groups.[60] Recent research[61] has also shown that participants' willingness to join remote research studies and share data are tied to their trust in the scientific team conducting the study including the institutional affiliations of researchers. Using different recruitment strategies[46–48] including targeted online ads in regions known to have a larger proportion of the minority groups, partnerships with local community organizations, clinics and universities may help improve the penetration of remote research and improve diversity in the recruited sample. The ongoing "All of Us" research program that includes remote digital data collection has shown the feasibility of using a multifaceted approach to recruit a diverse sample with a majority of the cohort coming from communities underrepresented in biomedical research.[62] Additionally, simple techniques, such as stratified recruitment that is customized based on the continual monitoring of the enrolling cohort demographics, can help enrich for a target population.

Finally, communication in digital health research may benefit from adopting the diffusion of innovations approach[63,64] that has been applied successfully in healthcare settings to change behavior including the adoption of new technologies.[65–67] Research study enrollments, advertisements including in-app communication and return of information to participants,[68] could be tailored to fit three distinct personality types (trendsetters, majority, and laggards). While trendsetters will adopt innovations early, they are a minority (15%) compared to the majority (greater than two-thirds of the population) who will adopt a new behavior after hearing about its real-benefits, utility and believe it is the status quo. On the other hand, laggards (15%) are highly resistant to change and hard to reach online and as a result, will require more targeted and local outreach efforts.

These results should also be viewed within the context of limitations related to integrating diverse user-engagement data across digital health studies that targeted different disease areas with varying underlying disease characteristics and severity. While we did adjust for potential study level heterogeneity, we were not able to account for within-study differences such as variations in participants' disease severity and any other study-specific temporal changes. For example, the user experience and burden could have changed or improved over time based on changes in the study protocol or other technical fixes in the app. The variations in participant recruitment were not fully documented across studies so not be analyzed and accounted for fully except for clinical referrals in two of the studies. Furthermore, the comparison of participant race/ethnicity and geographical diversity to a general US population was meant to assess the representativeness of the study that is aimed at recruiting from a general population and not necessarily targeted towards a specific clinical condition of interest. Researchers should also prioritize to collect demographic data such as age, gender, race/ethnicity, participant state during onboarding which help characterize user attrition in future studies. While sensitivity analysis showed the main findings from user retention analysis do not change by including participants with missing data, however, missing demographic characteristics remains a significant challenge for digital health (See Supplementary Table 7). Finally, obtaining raw user-level engagement data from digital health studies that is well annotated and computable remains a significant challenge. The present findings are based on eight select US-based digital health studies and thus may not be generalizable broadly. We do, however, hope that this work will help motivate digital health researchers to share user-level engagement data to help guide a larger systematic analysis of participant behavior in digital health studies.

Despite these limitations, the present investigation to the best of our knowledge is the largest cross-study analysis of participant retention in remote digital health studies using individual-level data. While the technology has enabled researchers to reach and recruit participants for conducting large-scale health research in short periods of time, more needs to be done to ensure equitable access and long-term utilization by participants across different populations. The low retention in "fully remote, app-based" health research may also need to be seen in the broad context of the mobile app industry where similar user attrition is reported.[69] Attrition in remote research may also be impacted by study burden[30] as frequent remote assessments can compete with users' everyday priorities and perceived value proposition for completing a study task that may not be linked to an immediate monetary incentive. Using co-design techniques[70] for developing study apps involving researchers and participants could help guide the development of most parsimonious research protocols that fit into the daily lives of people and are still sufficiently comprehensive for researchers.

In the present diverse sample of user-activity data, several cohort characteristics, such as age, disease status, clinical referral, monetary benefits, etc, have emerged as key drivers for higher retention. These characteristics may also guide the development of new data-driven engagement strategies[71,72] such as tailored just-in-time interventions[73] targeting sub-populations that are most likely to dropout early from remote research. Left unchecked the ongoing bias in participant recruitment combined with inequitable long-term participation in large-scale "digital cohorts" can severely impact the generalizability[36,37] and undermine the promise of digital health in collecting representational real-world data.

METHODS

Data acquisition

The user-engagement data were compiled by combining data from four studies that published annotated, accessible, and computable user-level data[16,19,74,75] under qualified researcher program[76] as well as new data from four other digital health studies(SleepHealth,[77] Start,[78] Phendo,[79] and ElevateMS[80]) that were contributed by collaborators. These eight studies aimed at assessing different diseases ranging from Parkinson's(mPower), asthma(Asthma health), heart condition(MyHeartCounts), sleep health (SleepHealth), multiple sclerosis(ElevarteMS), endometriosis(Phendo) to depression(Brighten and Start). Except for three studies (Brighten, Phendo, and Start) that aimed to recruit people with a specific clinical condition, the other five studies enrolled people from the general population. Anyone with and without the target disease that met the other study eligibility criteria could join. The studies recruited participants from throughout the US between 2014–2019 using a combination of different approaches including placing ads on social media, publicizing or launching the study at a large gathering, partnerships with patient advocacy groups, clinics, and through word of mouth. In all studies, participants were enrolled fully remotely either through a study website or directly through the study app and provided electronic consent[81] to participating in the study. Ethical oversight of the eight remote health studies included in the analysis was conducted by the respective institution's Institutional Review Board/Ethics Boards: Brighten (University of California, San Francisco), SleepHealth (University of California, San Diego), Phendo(Columbia University), Start (GoodRx), mPower (Sage Bionetworks), elevateMS(Sage Bionetworks), Asthma(Mt. Sinai), MyHeartCounts(Stanford University). The present retention analysis study used existing de-identified data only and qualifies for exemption status per OHRP guidelines.[82]

The studies were launched at different time points during the 2014–2019 period, including three studies mPower, MyHeartCounts, and Asthma being launched with the public release of ResearchKit framework[83] released by Apple in March 2015. The studies were also active for different time periods including significant differences in the minimum time participants were expected to participate in the studies remotely. While Brighten and ElevateMS had a fixed 12-week participation period, other studies allowed participants to remain active for as long as they desired. Given this variation in the expected participation period across the studies, we selected the minimum common time period of the first 12 weeks (84 days) of each participant's activity in each study for retention analysis. Finally, with the exception of Brighten study which was a randomized interventional clinical trial and enrolled depressed cohort offering them monetary incentives for participation, the rest of the seven studies were observational and did not offer any direct incentives for participation. The studies also collected different real-world data ranging from frequent subjective assessments, objective sensor-based tasks to continual passive data[41] collection.

Data harmonization

User-activity data across all the apps were harmonized to allow for inter-app comparison of user-engagement metrics. All in-app surveys and sensor-based tasks (e.g., Finger tapping on the screen) were classified as "active tasks" data type. The data gathered without explicit user action such as daily step count (Apple's health kit API[84]), daily local weather patterns were classified as "passive" data type and were not used for assessing active user-engagement. The frequency at which the active tasks were administered in the study apps were aligned based on the information available in the corresponding study publication or obtained directly from the data contributing team in case the data were not publicly available. Furthermore, there were significant differences in the baseline demographics that were collected by each app. A minimal subset of four demographic characteristics (age, gender, race, state) was used for participant recruitment and retention analysis. A subset of five studies (mPower, ElevateMS, SleepHealth, Asthma, MyHeartCounts) had enrolled participants with and without disease status and were used to asses retention differences between people with (case) and without (control) disease. Two studies (mPower and ElevateMS) had a subset of participants that were referred to use the same study app by their care providers. For this smaller but unique subgroup, we compared the retention differences between clinically referred participants to self-referred participants.

Statistical analysis

We used three key metrics to assess participant retention and long-term engagement. (1) Duration in the study: the total duration, a study participant remained active in the study i.e., the number of days between the first and last active task completed by the participant, during the first 84 days of each participant's time in the study. (2) Days active in the study: the number of days a participant performed any active task in the app within the first 84 days. (3) User-activity streak: a binary-encoded vector representing the 84 days of potential app participation for each participant (Fig. 3a) where the position of the vector indicates the participant's day in the study and is set to 1 (green box) if at least one active task is performed on that day or else is 0(white). User-activity streak metric was used to assess sub-populations that show similar longitudinal engagement patterns over a 3 month period.

Participant retention analysis (survival analysis[85]) was conducted using the total duration of time in the study as the outcome metric to compare the retention differences across studies, sex, age group, disease status, and clinical referral for study-app usage. The duration of each participant in the study was assessed based on "active task" completion i.e., tasks that require active user input (e.g., a survey or a sensor-based active task). With the underlying user-level engagement data available across selected eight studies, we used an individualized pooled data analysis(IPDA) approach[86] to compare participant retention. IPDA has shown to yield more reliable inference compared to pooling estimates from published studies.[86] Log-rank test[87] was used to compare significant differences in participant retention between different comparator groups of interest. In order to adjust for potential study level heterogeneity, we used a stratified version of the log-rank test. Kaplan-Meier[88] plots were used to summarize the effect of the main variable of interest by pooling the data across studies where applicable. Two approaches were used to evaluate participant retention using survival analysis. (1) No censoring (most conservative)—if the last active task completed by participant fell within the pre-specified study period of the first 84 days, we considered it to be a true event i.e., participant leaving the study (considered "dead" for survival analysis). (2) Right-censoring[88]—to assess the sensitivity of our findings using approach 1, we relaxed the determination of true event (participant leaving the study) in the first 12 weeks to be based on the first 20 weeks of app activity (additional 8 weeks). For example, if a participant completes last task in an app on day 40(within the first 84 days) and then additionally completes more active task/s between week 13–20 he/she was still considered alive (no event) during the first 84 days (12 weeks) of the study and therefore "right-censored" for survival analysis. Given that age and gender had a varying degree of missingness across studies; additional analysis comparing the retention differences between the two sub-groups that provided the demographics and that opted out was done to assess the sensitivity of missing data on main findings.

Unsupervised k-means clustering method was used to investigate the longitudinal participant engagement behavior within each study using the user-activity streak metric (described above). The number of optimum clusters (between 1–10) in each study was determined using the elbow method[89] that aims to minimize the within-cluster variation. Enrichment of demographic characteristics in each cluster was assessed using a one-way analysis of variance. Since the goal of this unsupervised clustering of user-activity streaks was to investigate the patterns in longitudinal participant engagement; we filtered out individuals who remained in the study for less than 7 days from clustering analysis. However, for post hoc comparisons of demographics across the clusters, the initially left-out participants were put in a separate group (C5*). The state-wise proportions of recruited participants in each app were compared to the 2018 US state population estimates using the data obtained from the US census bureau.[90] To eliminate potential bias related to marketing and advertising of the launch of Apple's Research kit platform on March 09, 2015, participants who joined and left the mPower, MyHeartCounts, Asthma studies within the first week of Research Kit launch ($N = 14,573$) were taken out from the user retention analysis. We initially considered using Cox proportional hazards model[91] to test for the significance of variable of interest on user retention within each study accounting for other study-specific covariates. However, because the assumption of proportional hazards (tested using scaled Schoenfeld residuals) was not supported for some studies, these analyses were not further pursued. All statistical analyses were performed using R.[92]

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Aggregated data for all studies are included as part of tables in the main manuscript or supplementary tables. Additionally, individual-level user-engagement data for all studies are available under controlled access through the Synapse (https://doi.org/10.7303/syn20715364).

CODE AVAILABILITY

The complete code used for the analysis is available through a GitHub code repository (https://github.com/Sage-Bionetworks/digitalHealth_RetentionAnalysis_PublicRelease).

Received: 1 October 2019; Accepted: 17 January 2020;

REFERENCES

1. Fogel, D. B. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: a review. *Contemp. Clin. Trials Commun.* **11**, 156–164 (2018).
2. Briel, M. et al. A systematic review of discontinued trials suggested that most reasons for recruitment failure were preventable. *J. Clin. Epidemiol.* **80**, 8–15 (2016).
3. Bradshaw, J., Saling, M., Hopwood, M., Anderson, V. & Brodtmann, A. Fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease is qualitatively distinct. *J. Neurol. Neurosurg. Psychiatry* **75**, 382–387 (2004).
4. Snyder, M. & Zhou, W. Big data and health. *Lancet Digital Health*. https://doi.org/10.1016/s2589-7500(19)30109-8 (2019).
5. Sim, I. Mobile devices and health. *N. Engl. J. Med.* **381**, 956–968 (2019).
6. Steinhubl, S. R., Muse, E. D. & Topol, E. J. The emerging field of mobile health. *Sci. Transl. Med.* **7**, 283rv3 (2015).
7. ElZarrad, M. K., Khair ElZarrad, M. & Corrigan-Curay, J. The US Food and Drug Administration's Real-World Evidence Framework: A Commitment for Engagement and Transparency on Real-World Evidence. *Clin. Pharmacol. Therapeutics*. https://doi.org/10.1002/cpt.1389 (2019).
8. May, M. Clinical trial costs go under the microscope. *Nat. Med.* https://doi.org/10.1038/d41591-019-00008-7 (2019).
9. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. *Pew Research Center's Global Attitudes Project*. https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ (2019).
10. Turner, B. A. Smartphone Addiction & Cell Phone Usage Statistics in 2018. *BankMyCell*. https://www.bankmycell.com/blog/smartphone-addiction/ (2018).
11. Global time spent on social media daily 2018 | Statista. *Statista*. https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/.
12. Steinhubl, S. R. & Topol, E. J. Digital medicine, on its way to being just plain medicine. *npj Digit Med.* **1**, 20175 (2018).
13. Perry, B. et al. Use of mobile devices to measure outcomes in clinical research, 2010–2016: a systematic literature review. *Digit Biomark.* **2**, 11–30 (2018).
14. Trister, A., Dorsey, E. & Friend, S. Smartphones as new tools in the management and understanding of Parkinson's disease. *npj Parkinson's Dis.* **2**, 16006 (2016). https://doi.org/10.1038/npjparkd.2016.6.
15. Dorsey, E. R. et al. The use of smartphones for health research. *Acad. Med.* **92**, 157–160 (2017).
16. Bot, B. M. et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci. Data* **3**, 160011 (2016).
17. Webster, D., Suver, C., Doerr, M. et al. The Mole Mapper Study, mobile phone skin imaging and melanoma risk data collected using ResearchKit. *Sci. Data* **4**, 170005 (2017). https://doi.org/10.1038/sdata.2017.5.
18. Crouthamel, M. et al. Using a ResearchKit smartphone app to collect rheumatoid arthritis symptoms from real-world participants: feasibility study. *JMIR mHealth uHealth* **6**, e177 (2018).
19. McConnell, M. V. et al. Feasibility of obtaining measures of lifestyle from a Smartphone App: the MyHeart Counts Cardiovascular Health Study. *JAMA Cardiol.* **2**, 67–76 (2017).
20. Chan, Y.-F. Y. et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. *Nat. Biotechnol.* **35**, 354–362 (2017).
21. McKillop, M., Mamykina, L. & Elhadad, N. Designing in the Dark: Eliciting Self-tracking Dimensions for Understanding Enigmatic Disease. in *Proc 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* 1–15 (ACM Press, 2018).
22. Waalen, J. et al. Real world usage characteristics of a novel mobile health self-monitoring device: results from the Scanadu Consumer Health Outcomes (SCOUT) Study. *PLoS ONE* **14**, e0215468 (2019).

23. Pratap, A. et al. Using mobile apps to assess and treat depression in Hispanic and Latino Populations: fully remote randomized clinical trial. *J. Med. Internet Res.* **20**, e10130 (2018).
24. Anguera, J. A., Jordan, J. T., Castaneda, D., Gazzaley, A. & Areán, P. A. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. *BMJ Innov.* **2**, 14–21 (2016).
25. Smalley, E. Clinical trials go virtual, big pharma dives in. *Nat. Biotechnol.* **36**, 561–562 (2018).
26. Virtual Clinical Trials Challenges and Opportunities: Proceedings of a Workshop: Health and Medicine Division. (2019). http://www.nationalacademies.org/hmd/Reports/2019/virtual-clinical-trials-challenges-and-opportunities-pw.aspx.
27. Orri, M., Lipset, C. H., Jacobs, B. P., Costello, A. J. & Cummings, S. R. Web-based trial to evaluate the efficacy and safety of tolterodine ER 4 mg in participants with overactive bladder: REMOTE trial. *Contemp. Clin. Trials* **38**, 190–197 (2014).
28. Moore, S. et al. Consent processes for mobile app mediated research: systematic review. *JMIR Mhealth Uhealth* **5**, e126 (2017).
29. Althubaiti, A. Information bias in health research: definition, pitfalls, and adjustment methods. *J. Multidiscip. Healthc.* **9**, 211–217 (2016).
30. Druce, K. L., Dixon, W. G. & McBeth, J. Maximizing engagement in mobile health studies: lessons learned and future directions. *Rheum. Dis. Clin. North Am.* **45**, 159–172 (2019).
31. Christensen, H., Griffiths, K. M., Korten, A. E., Brittliffe, K. & Groves, C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *J. Med. Internet Res.* **6**, e46 (2004).
32. Christensen, H., Griffiths, K. M. & Jorm, A. F. Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ* **328**, 265 (2004).
33. Eysenbach, G. The law of attrition. *J. Med. Internet Res.* **7**, e11 (2005).
34. Christensen, H. & Mackinnon, A. The law of attrition revisited. *J. Med. Internet Res.* **8**, e20 (2006).
35. Eysenbach, G. The law of attrition revisited—Author's -reply. *J. Med. Internet Res.* **8**, e21 (2006).
36. Kaplan, R. M., Chambers, D. A. & Glasgow, R. E. Big data and large sample size: a cautionary note on the potential for bias. *Clin. Transl. Sci.* **7**, 342–346 (2014).
37. Neto, E. C. et al. A Permutation Approach to Assess Confounding in Machine Learning Applications for Digital Health. in *Proc 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining - KDD '19*, https://doi.org/10.1145/3292500.3330903 (2019).
38. Dainesi, S. M. & Goldbaum, M. Reasons behind the participation in biomedical research: a brief review. *Rev. Bras. Epidemiol.* **17**, 842–851 (2014).
39. Balachandran, M. Way To Health. (2019). https://www.waytohealth.org/.
40. Bentley, J. P. & Thacker, P. G. The influence of risk and monetary payment on the research participation decision making process. *J. Med. Ethics* **30**, 293–298 (2004).
41. Laport-López, F., Serrano, E., Bajo, J. & Campbell, A. T. A review of mobile sensing systems, applications, and opportunities. *Knowl. Inf. Syst*. (2019). https://doi.org/10.1007/s10115-019-01346-1.
42. Laursen, D. R. T., Paludan-Müller, A. S. & Hróbjartsson, A. Randomized clinical trials with run-in periods: frequency, characteristics and reporting. *Clin. Epidemiol.* **11**, 169–184 (2019).
43. Pablos-Méndez, A., Barr, R. G. & Shea, S. Run-in periods in randomized trials: implications for the application of results in clinical practice. *JAMA* **279**, 222–225 (1998).
44. Dickert, N. & Grady, C. What's the price of a research subject? Approaches to payment for research participation. *N. Engl. J. Med.* **341**, 198–203 (1999).
45. Gelinas, L. et al. A framework for ethical payment to research participants. *N. Engl. J. Med.* **378**, 766–771 (2018).
46. A guide to actively involving young people in research—INVOLVE. (2004). https://www.invo.org.uk/posttypepublication/a-guide-to-actively-involving-young-people-in-research/.
47. Nicholson, L. M. et al. Recruitment and retention strategies in longitudinal clinical studies with low-income populations. *Contemp. Clin. Trials* **32**, 353–362 (2011).
48. Nicholson, L. M., Schwirian, P. M. & Groner, J. A. Recruitment and retention strategies in clinical studies with low-income and minority populations: progress from 2004–2014. *Contemp. Clin. Trials* **45**, 34–40 (2015).
49. Druce, K. L. et al. Recruitment and ongoing engagement in a UK smartphone study examining the association between weather and pain: cohort study. *JMIR mHealth uHealth* **5**, e168 (2017).
50. Liu, H.-E. & Li, M.-C. Factors influencing the willingness to participate in medical research: a nationwide survey in Taiwan. *PeerJ* **6**, e4874 (2018).
51. Shavers, V. L., Lynch, C. F. & Burmeister, L. F. Factors that influence African-Americans' willingness to participate in medical research studies. *Cancer* **91**, 233–236 (2001).
52. Trauth, J. M., Musa, D., Siminoff, L., Jewell, I. K. & Ricci, E. Public attitudes regarding willingness to participate in medical research studies. *J. Health Soc. Policy* **12**, 23–43 (2000).
53. 5 key findings about public trust in scientists in the U.S. *Pew Research Center.* (2019). https://www.pewresearch.org/fact-tank/2019/08/05/5-key-findings-about-public-trust-in-scientists-in-the-u-s/.
54. Nebeker, C., Murray, K., Holub, C., Haughton, J. & Arredondo, E. M. Acceptance of mobile health in communities underrepresented in biomedical research: barriers and ethical considerations for scientists. *JMIR Mhealth Uhealth* **5**, e87 (2017).
55. Christopher Gibbons, M. Use of health information technology among racial and ethnic underserved communities. *Perspect. Health Inf. Manag.* **8**, 1f (2011).
56. Massoud, M. R., Rashad Massoud, M. & Kimble, L. Faculty of 1000 evaluation for Global burden of diseases, injuries, and risk factors for young people's health during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *F1000 - Post-Publ. Peer Rev. Biomed. Lit*. (2016). https://doi.org/10.3410/f.726353640.793525182.
57. Oh, S. S. et al. Diversity in clinical and biomedical research: a promise yet to be fulfilled. *PLoS Med.* **12**, e1001918 (2015).
58. Kondo, N. et al. Income inequality, mortality, and self rated health: meta-analysis of multilevel studies. *BMJ* **339**, b4471 (2009).
59. Krieger, N., Chen, J. T., Waterman, P. D., Rehkopf, D. H. & Subramanian, S. V. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures—The Public Health Disparities Geocoding Project. *Am. J. Public Health* **93**, 1655–1671 (2003).
60. Roth, G. A. et al. Trends and patterns of geographic variation in cardiovascular mortality among US Counties, 1980–2014. *JAMA* **317**, 1976–1992 (2017).
61. Pratap, A. et al. Contemporary views of research participant willingness to participate and share digital data in biomedical research. *JAMA Netw. Open* **2**, e1915717 (2019).
62. All of Us Research Program Investigators et al. The 'All of Us' Research Program. *N. Engl. J. Med.* **381**, 668–676 (2019).
63. Rogers, E. M., Singhal, A. & Quinlan, M. M. in *An Integrated Approach to Communication Theory and Research*. Taylor & Francis Group. (2019). https://doi.org/10.4324/9780203710753-35.
64. Karni, E. & Safra, Z. in *Risk, Decision and Rationality*. Taylor & Francis Group. (1988). https://doi.org/10.1007/978-94-009-4019-2_11.
65. Greenhalgh, T. et al. Introduction of shared electronic records: multi-site case study using diffusion of innovation theory. *BMJ* **337**, a1786 (2008).
66. Ward, R. The application of technology acceptance and diffusion of innovation models in healthcare informatics. *Health Policy Technol.* **2**, 222–228 (2013).
67. Emani, S. et al. Patient perceptions of a personal health record: a test of the diffusion of innovation model. *J. Med. Internet Res.* **14**, e150 (2012).
68. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy & Committee on the Return of Individual-Specific Research Results Generated in Research Laboratories. *Returning Individual Research Results to Participants: Guidance for a New Research Paradigm*. (National Academies Press, 2018).
69. Mobile Apps: What's A Good Retention Rate? (2018). http://info.localytics.com/blog/mobile-apps-whats-a-good-retention-rate.
70. Yardley, L., Morrison, L., Bradbury, K. & Muller, I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J. Med. Internet Res.* **17**, e30 (2015).
71. O'Connor, S. et al. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. *BMC Medical Informatics and Decision Making* 16 (2016).
72. Pagoto, S. & Bennett, G. G. How behavioral science can advance digital health. *Transl. Behav. Med.* **3**, 271–276 (2013).
73. Nahum-Shani, I. et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Ann. Behav. Med.* **52**, 446–462 (2018).
74. Chan, Y.-F. Y. et al. The asthma mobile health study, smartphone data collected using ResearchKit. *Sci. Data* **5**, 180096 (2018).
75. Hershman, S. G. et al. Physical activity, sleep and cardiovascular health data for 50,000 individuals from the MyHeart Counts Study. *Sci. Data* **6**, 24 (2019).
76. Wilbanks, J. & Friend, S. H. First, design for data sharing. *Nat. Biotechnol.* **34**, 377–379 (2016).
77. SleepHealth Study and Mobile App - SleepHealth. *SleepHealth*. (2019). https://www.sleephealth.org/sleephealthapp/.
78. Website. (2019). https://www.goodrx.com/.
79. Phendo. (2019). http://citizenendo.org/phendo/.
80. elevateMS. (2019). http://www.elevatems.org/.
81. Coiera, E. e-Consent: the design and implementation of consumer consent mechanisms in an electronic environment. *J. Am. Med. Inform. Assoc.* **11**, 129–140 (2003).
82. Electronic Code of Federal Regulations (eCFR). *Electronic Code of Federal Regulations (eCFR)*. (2019). https://www.ecfr.gov/.
83. ResearchKit and CareKit. *Apple*. (2019). http://www.apple.com/researchkit/.

84. Apple Inc. HealthKit - Apple Developer. (2019). https://developer.apple.com/healthkit/.
85. Bewick, V., Cheek, L. & Ball, J. Statistics review 12: survival analysis. *Crit. Care* **8**, 389 (2004).
86. Blettner, M. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int. J. Epidemiol.* **28**, 1–9 (1999).
87. Bland, J. M. & Altman, D. G. The logrank test. *BMJ* **328**, 1073 (2004).
88. Rich, J. T. et al. A practical guide to understanding Kaplan-Meier curves. *Otolaryngol. Head. Neck Surg.* **143**, 331–336 (2010).
89. Parsons, S. Introduction to Machine Learning by Ethem Alpaydin, MIT Press, 0-262-01211-1, 400 pp., $50.00/£32.95. The Knowledge Engineering Review. 432–433 (2005).
90. US Census Bureau. 2018 National and State Population Estimates. https://www.census.gov/newsroom/press-kits/2018/pop-estimates-national-state.html.
91. Kumar, D. & Klefsjö, B. Proportional hazards model: a review. *Reliab. Eng. Syst. Saf.* **44**, 177–188 (1994).
92. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2013). http://www.R-project.org/.

ACKNOWLEDGEMENTS

This work was supported in part by various funding agencies that included the National Institute of Mental Health (MH100466), Robert Wood Johnson Foundation (RWJF - 73205), National Library of Medicine (R01LM013043), National Center for Advancing Translational Sciences (1UL1TR002319-01) and American Sleep Apnea Foundation, Washington, DC. The funding agencies did not play a role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit this article for publication. We also would like to thank Phendo, SleepHealth, and GoodRx groups for sharing user-engagement data from their respective digital health studies for the retention analysis. We also acknowledge and thank researchers from previously completed digital health studies for making the user-engagement data available to the research community under qualified researcher program. Specifically, mPower study data were contributed by users of the Parkinson mPower mobile application as part of the mPower study developed by Sage Bionetworks and described in Synapse [https://doi.org/10.7303/syn4993293]. MyHeartCounts study data were contributed by users of the My Heart Counts Mobile Health Application study developed by Stanford University and described in Synapse [https://doi.org/10.7303/syn11269541]. Asthma study data were contributed by users of the asthma app as part of the Asthma Mobile Health Application study developed by The Icahn School of Medicine at Mount Sinai and described in Synapse [https://doi.org/10.7303/syn8361748]. We would like to thank and acknowledge all study participants for contributing their time and effort to participate in these studies.

AUTHOR CONTRIBUTIONS

A.P. conceptualized the study, wrote the initial analytical plan, carried out the analysis, and wrote the first draft of the paper. A.P. and P.S. integrated and harmonized user-engagement data across the studies. P.A.A. and A.P. managed the data-sharing agreements with data contributors (N.E., M.M., C.S. and D.G.). E.C.N., P.H. and L.O. made significant contributions to the data analysis. P.A.A., L.M. and L.O. helped interpret the analytical findings and provided feedback on the initial manuscript draft. A.P. and P.A.A. made major revisions to the first draft. C.S. and J.W. implemented the data governance and sharing model to enable the sharing of raw user-engagement data with the research community. L.O. and P.A.A. jointly oversaw the study. All authors assisted with the revisions of the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41746-020-0224-8.

Correspondence and requests for materials should be addressed to A.P.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Perakslis *Conflict and Health* (2018) 12:23
https://doi.org/10.1186/s13031-018-0163-z

Conflict and Health

REVIEW Open Access

Using digital health to enable ethical health research in conflict and other humanitarian settings

Eric D. Perakslis[1,2,3]

Abstract

Conducting research in a humanitarian setting requires quantifiable quality measures to ensure ethical study conduct. Digital health technologies are proven to improve research study quality and efficacy via automated data collection, improvement of data reliability, fidelity and resilience and by improved data provenance and traceability. Additionally, digital health methodologies can improve patient identity, patient privacy, study transparency, data sharing, competent informed consent, and the confidentiality and security of humanitarian operations. It can seem counterintuitive to press forward aggressively with digital technologies at a time of heightened population vulnerability and cyber security concerns, but new approaches are essential to meet the rapidly increasing demands of humanitarian research. In this paper we present the case for the digital modernization of humanitarian research in conflict and other humanitarian settings as a vehicle for improved research quality and ethics.

Keywords: Digital health, Humanitarian research, eHealth, mHealth, Data privacy, Data security, Data sharing, Biometric identity, Research ethics

Background

While there is justifiable concern, dialogue and debate on the necessity of research in conflict and humanitarian settings, research is being conducted and likely will continue to be conducted in these settings. Governments, non-governmental organizations (NGOs) and the international community must understand the benefit-risk ratios and efficacy of interventions, and the resulting health, social and economic outcomes following such interventions. Concurrently, it must be clear that the rights, confidentiality and identity of all research subjects are protected and that all possible harms were minimized [1]. Researchers are consistently studying and reviewing these and other ethical obligations, and there has been substantial progress in developing methods and practices to ensure ethical research conduct in humanitarian settings [2, 3]. Despite this progress, questions remain. Is all 'interesting' research necessary? Does the production of high quality evidence hinder the ability to provide the most effective interventions possible to the most vulnerable populations [4]? These questions are difficult and necessary, as consideration of research must always include definitions and measures of benefit-risk ratios and proper ethical oversight before, during, and after humanitarian interventions. This writing proposes that modern digital technologies can improve the ethics and benefits while reducing the risks of research conduct in humanitarian settings.

In recent reviews on health and humanitarian crisis, two of the primary recommendations were the "ethical imperative" of collecting better data and the need for better information systems [5, 6]. Better data is: generated by valid experimental designs; timely; statistically rigorous; properly protected; useful for local authorities; obtained only through proper (truly) informed consent; and capable of providing an evidence base to support the resulting conclusions and recommendations of a given study. Better information systems are: international standards-based; available but secure; quickly available when crises occur; interconnected; cost-effective; and operationally accessible and useful to local

---

Correspondence: eperakslis@gmail.com
[1] Datavant Inc., San Francisco, CA, USA
[2] Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA
Full list of author information is available at the end of the article

authorities. Ideally, better systems can also be an important component in strengthening local health systems as was shown during the recent Ebola epidemic in West Africa [7]. Despite the incessant marketing hype, digital health technologies are designed to improve data quality and security, systems availability and systems flexibility. In addition, digital systems have great potential in patient/subject identity tracking and identity protection, improved communications, data confidentiality, data redundancy, data protection and local workforce development [8].

Overview of the utility of *e*Health and *m*Health in humanitarian research settings

The recent Ebola crisis in West Africa highlighted the limitations of paper systems for patient care, research and logistical support during an infectious disease outbreak. While common and readily available, paper is fragile, easily damaged by weather, greatly limited by the skills/literacy of the user, often difficult to read or reproduce. Most of all, paper itself can be a transmission vector during infectious disease epidemics [9]. Digital technologies, on the other hand, continue to improve and to be proven effective, even in low resource settings when properly engineered and implemented. Success is far from automatic. Essential elements for successful digital technology implementation include proper methodologies, qualified personnel, strong use case and scenario selection, realistic expectations and high-touch stakeholder management [10]. Unless these elements are fully understood and effectively executed, digital technology implementation can result in costly mistakes. The fact that massive amounts of resources continue to pour into digital technologies should provide optimism that these solutions are in close reach and will continue to improve.

One important challenge of conducting research in conflict and humanitarian settings is that it is resource-intensive and could divert attention from patient care [11]. In some cases, this resource burden can be offset by automation. When data is captured electronically, some tasks are automated, and others can be expedited. Conduct of surveys provide the most common example. Paper surveys can be time consuming, illegible, poorly understood by the worker administering the survey, easily damaged by weather or transport and easily stolen or destroyed by hostile actors. The author witnessed this first-hand when deploying a community surveillance mobile app in Kono Sierra Leone during the recent Ebola epidemic. Not only was the app found to be quickly superior to paper for data collection, data quality, data protection and accuracy, the program has proven to be useful for healthcare systems strengthening as there are now more than 100 community health workers using the system [12]. In addition to disease epidemic information, clinical knowledge inside conflict settings has also been successful collected through electronic surveys. One recent study polled healthcare providers within Syria to test awareness of tele-mental health (TMH). The study showed that initial awareness of TMH was low but that the polled physicians were interested, willing to try and thought that such interventions could be effective [13]. Indeed, these tools can actually reach massive numbers of users and bring great utility in a short period of time. This was the case with a Médecins Sans Frontières' Clinical Guidance mobile application which was downloaded in 150 countries and resulted in 250,000 screen views in the first 6 months [14]. These are just a few examples, but the literature and the technology press are blooming with new examples monthly, and best practices are rapidly emerging.

The arguments against the use of digital health tools for the purpose of expanding the reach and minimizing the resource burden of research is that the technology will be too foreign, too complex and too difficult for successful utilization within some settings. While these concerns are valid and important, the landscape is evolving very quickly. Digital transformation appears to be accelerating in low resource areas and conflict zones. One fascinating example is the rapid adoption of cashless currency in challenging settings such as Somaliland. Indeed, even in a country with very high illiteracy rates, it is both simplicity and enhanced functionality that are helping cashless currency flourish [15]. Clearly, the familiarity of cellular phones and tablets is on the rise in low and middle-income countries (LMIC), and this trend will help offset the concern that these technologies appear too foreign.

The potential complexity and difficulty of using digital health technologies must be managed and mitigated carefully by experienced personnel. Digital projects fail primarily due to project management and social issues, regardless of country income level [16]. The most common reasons for failure include avoidance of root cause challenges, unclear or under-articulated goals, lack of proper methodology, lack of understanding of true customer needs, inadequately qualified leadership and staff, poor technology selection, poor communication and poor change management. While information technology (IT) project management is beyond the scope of this writing, Table 1 shows common technology delivery project pitfalls and offers practical guidance [17–22].

Ethical issues of Health Research to be addressed

The basic principles behind ethical human subject research are well articulated and include respect for persons (and their choices), beneficence (the research must do good), non-maleficence (the research does not harm)

Table 1 Best Practices and Common Mistakes in Digital Healthcare Implementation

Challenges in *e*Health & *m*Health Delivery	Best Practices & Specific Methodologies
Poor or limited user involvement & engagement	User-centered design, user co-design, participatory design methodologies
Unclear goals, expectations & scope creep	Develop & use a clear requirements & expectations matrix
Poor sponsor participation & active leadership	Document specific sponsor role requirements & the corresponding relationships to other roles
Poor technology selection	Use an established technology selection framework
Lack of necessary technology skill sets	Understand the necessary roles & recruit, train or contract
Poor project management & lack of formal methodology	Understand & select from 6 most common technology delivery methodologies

and justice (all persons are treated fairly and equally) [23]. These principles are elaborated on within the International Ethical Guidelines for Health-Related Research Involving Humans published by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO) [24]. Despite these guidelines, the application of ethical frameworks to digital health is still new.

It is not always clear how to best apply specific ethical guidelines to new technologies. Technology can be unfamiliar, scary and intimidating. Common concerns include training, accuracy, reliability, privacy, security, inequality and protection of relationships [25]. Despite CIOMS guidance, further subtleties that must be examined are the differences in the collection and use of aggregate population data versus individual patient data. One emergent example is the debate around the use of aggregate phone call detail record data from mobile phone systems within low and middle-income countries (LMIC) [26]. When aggregate data is being made available for research via third parties, how is consent handled? Is the research really in the best interest of the consumers about whom the data was collected? One of the primary requirements of research, of course, is trust between the various actors. However, in fragile states and during conflict, this trust is often missing which can greatly inhibit participation [27]. When considered thoroughly, these concerns are not new, unique or limited to conflict and humanitarian settings. In fact, these concerns are the same as are being actively debated and managed in most healthcare systems, but additional protections must be enabled for the most vulnerable people.

Some argue that the only answer to the collective challenges of dire unmet humanitarian need and significant ethical hurdles is the forward press of innovation. Indeed, Médecins Sans Frontières considers innovation an essential element of humanitarian response and has published a framework for humanitarian innovation that considers harms, benefits, local participation, longer-term consequences and specific delivery methodologies [28]. Similarly, it has recently been suggested that, with proper education and outreach, *m*Health and telehealth offer a relatively low-resource platform for the Sustainable Development Goal (SDG) 3 in conflict-affected populations [29]. There has also been excellent recent work done to assess and describe responsible data approaches for humanitarian settings. Specifics include risk assessment, data value chain, legal foundations, and accountability and best practices [30].

Using the growing body of positive evidence of digital capabilities, an association can be constructed between specific CIOMS guidelines and the best practices of digital technologies as shown in Table 2. Each digital capability enhancement opportunity will be discussed in detail.

Technologies for improved ethical informed consent

Properly documented informed consent is an essential basis of ethical human subject research. All studies are ethically and legally bound to ensure that any and all potential research participants fully understand all aspects of the process they are being asked to undertake. This requires that potential research subjects receive, comprehend and make decisions on information that can be completely beyond experience or understanding. Common challenges include basic literacy, health literacy, the proper local context, cultural competency, proper documentation and the challenge of successfully communicating complex research and clinical protocols [31, 32].

These challenges are not limited to humanitarian or low-resource settings, as the entire world struggles to ensure that the informed consent process truly satisfies its ethical purpose and study documentation purposes [33, 34]. Fortunately, progress is being made, and technology is playing a greater role. For example, digital informed consent tools can include multi-media videos, stories, pop-up definitions and quizzes, all of which have

Table 2 Association of Specific CIOMS Guidelines and Digital Technology Improvement Opportunities

International Ethical Guideline (CIOMS)	Digital Enhancement Opportunities
Informed Consent (Guidelines 9,10,16)	Better comprehension via multi-media, improved privacy, traceability (including ability for withdrawal) & confidentiality
Collection Storage & Use of Data (Guidelines 11&12)	Improved Data Quality, Fidelity, Provenance, Data Reliability
Privacy and Confidentiality (Guidelines 3,4,11,12,20,22)	Digital Identity. Data Resilience, Data Redundancy
Data Transparency & Sharing (Guideline 8,12,22)	Increased Data and Study Transparency

been shown to improve patient comprehension and retention [35]. But are the same principles and tools being used to improve the informed consent process in the industrialized world suitable to humanitarian use? Early evidence is positive, but challenges remain. Multiple studies of informed consent across multiple medical discipline and in various developing nations show that the use of audio and visual multimedia demonstrate quantifiable improvements in understanding and retention [36]. The challenges reported include fear of data and privacy concerns and hesitance by potential research subjects to sign off on the consent forms.

With respect to data and privacy concerns, one of the risks is that more data can be collected than a subject understands. For example, apps could passively capture the GPS coordinates of the exact location of the consent, and this data could be used by other parties if the devices were not adequately protected and controlled [37]. This may be difficult, or impossible, to effectively communicate. With respect to the specific concern regarding signatures, a recent study in northern Ethiopia found that subjects were afraid to sign consent forms due to lack of trust of investigators and the concern that signatures could be related to legal accountability [38]. Clearly, there are important patient sensitivities and concerns regarding privacy, the potential misuse of personal information and fear of unintended consequences. This is where digitized personal identity may help greatly.

Next-generation digital identity and identity/privacy protection

The United Nations Sustainable Goal 16.9 calls for legal identity for all citizens including birth registration by 2030 [39]. This goal is aspirational and complex as no truly ideal global identification strategy exists. The complications of an unprecedented refugee crisis, unstable states and exploding identity theft and misuse in the industrialized world make this a global problem for all peoples, not just those in developing nations. The particular challenges to uniform global identity solutions include the lack of consistent state-issued identification (ID), political instability, corruption and fear of persecution and stigma. Indeed, even in the most developed nations, there is a growing trend of individuals that are choosing to live off-the-grid [40]. For many peoples, feeling safe has a great deal to do with feeling anonymous.

But what aspects of life need to be associated with identity? In the US, past and present identity schemes involve personal information that are based around establishing uniqueness. Date of birth, place of birth, social security number and other personally-identifiable-information (PII) such as home address form the basis of modern identity and, truly, most of this information has likely already been stolen. Estimates vary but the data breaches are now affecting 100 s of millions of citizens per year worldwide. From the healthcare perspective, in the US, personal health information (PHI) is further protected by the Health Insurance Portability and Accountability Act (HIPAA), although here, too, cyberattacks are all too common as this data is considered highly valuable [41, 42]. Many now question this strategy identifying people with personal information and then de-identifying those same people as research subjects by stripping a subset of the personal data. Must unique identification be based on some of the most personal and private attributes of life? Probably not, as newer technologies and smarter identity schemes are rapidly evolving.

Digital identity is likely the best path forward given the complexities and the financial, health and security issues around global identity [43]. Ideal solutions must be truly unique (at least nationally), portable, resilient, inexpensive and standards-based to allow interoperability with national systems. For populations at risk or threat, the additional capability to support those living on-the-grid and off-the-grid will be essential. This can be less tricky than it seems. When unique identity schemes do not depend on associated personally-identifiable data, the risk to individuals is greatly decreased. One recent article suggested that there 'should be an outcry to eliminate the brandishing of birthdates to identify patients in medical encounters' [44].

Fortunately, in the case of biometric ID, the source of uniqueness is not marketable information such as PII or PHI; it is simply biological traits, such as fingerprint or retinal scan, which need not be associated with any personal information to be fully unique [45]. In many ways, this strategy is essentially proactive de-identification according to HIPAA guidelines as long as none of the 18 types of identifiers are ever associated with the ID [46]. These technologies are rapidly evolving. In fact, in 2016, new national electronic ID (*e*ID) programs, most including biometrics, were announced in Algeria, Cameroon, Jordan, Italy, Senegal and Thailand, and pilots were launched in many other nations [47].

Also driving progress are federal government guidelines such as the US NIST SP 800–63 Digital Identity Guidelines. These guidelines provide comprehensive guidance on digital identity, enrollment, identity proofing, authentication and lifecycle management [48]. While not yet mandated, it is expected that these guidelines and equivalents from other nations will set the bar for quality and responsibility of national ID systems and must be carefully considered.

Admittedly, this is a great deal of technology and complexity to comprehend, and many humanitarian missions and settings lack the required technological sophistication to do so. This is where digital identity services can bring excellent value. Digital identity services provide identity solutions 'as a service' and are now being used by many sectors worldwide. The growth of this industry is so explosive that the greatest challenge can be selecting the optimal solution for a given purpose. Here again, there are excellent international guidance documents available to inform and guide [49].

Improving data reliability, fidelity and resilience

For data to be reliable it must be accurate, precise and available. Each of these can be aided by digital tools. For example, digital surveys greatly improve data accuracy and fidelity by enforcing data types such as numerical fields, date fields etc. that ensure proper answers. Multiple choice questions ensure specificity, precision, legibility and suitability of answers. Data availability can be improved by local data caching on devices as well as downloading copies of data via wireless networks (Wi-Fi) or to other devices via peer-to-peer data transfer. Data redundancy is also the best protection against data loss in any setting.

Data resilience is the ability to recover from loss or incident, and this is where digital methods greatly improve upon most paper systems. Paper can get wet or lost, or simply be illegible by the time of intended use. Digital devices offer instantaneous redundancy, even where there is no cellular or Wi-Fi capability, and can store and share copies while offline. Paper also becomes onerous to store and archive. Large missions can quickly compile stacks, boxes and rooms of poorly annotated and filed data making re-use and long-term utility difficult.

Improving data provenance

Another essential element of well-conducted ethical research is proper data provenance. Data provenance is the ability to describe the history and origins of data, a critical element of data reproducibility. The ability of digital systems to create metadata that can be used to establish and ensure data origin, chain of custody and reproducibility is a significant improvement over traditional paper procedures. These improvements can be further enhanced by many pre-existing ontologies that enable the use of data standards and the ability to automate data integrity checking [50, 51]. Lastly, while much of the data provenance and provenance metadata literature is focused upon highly technical and advanced cloud computing environments, it is essential to understand that the concepts are fully amenable to much lower tech environments. Solid experimental data provenance can be established using techniques such as basic labeling and tracking, using proper version control and backups, smart use of data identifiers and even hybrid digital and paper processes [52].

Data quality, data protection and research cyber security

The author has previously provided guidance for research study cyber security and privacy protection so will not dive deeply into these technicalities in this writing [53]. The most important aspects to consider in humanitarian settings are the specificity of the environment, prioritization of data and systems, access and identity management, proper device patching and management, comprehensive daily data backups, good physical security and regular testing of all procedures and technology controls [54].

Fundamental to all security and privacy strategies is an understanding that all data is not of equal risk and importance. In the wrong hands, a clinical case report form that identifies a subject solely based on a unique patient ID, carries much less risk to the patient and/or provider than does the spreadsheet or database that associates personal information with those unique patient IDs.

Data is not of equal risk and this is the basis behind HIPAA, GDPR and other privacy laws. These regulations must be seen as an opportunity to make research more efficient, portable and transparent. Decide what is important and protect what is important. Worry much less about everything else.

In considering data privacy and utility, electronic data can be more useful and secure overall. Consider the case of collecting and managing informed consent forms during any large medical intervention or study. If paper forms and wet signatures are used, what are the odds that a subject could be found and competently re-identified in a crowd fifteen minutes later? Chain of custody of data, including the ability to attach results and documents to particular subjects, is fundamental to ensuring study quality. Now consider the same scenario where an electronic case report form app and a digital biometric identity were used; instant and highly reliable re-contact would be possible and credible as needed.

One last important topic on cyber security is a specific caution around the Android operating system. The Android operating system is far more 'open' than the analogous iOS operating system used by Apple. This has truly enabled rapid and worldwide utility of mobile applications. Most open-source software systems run Android and most reasonably priced phones and tablets run Android. Because Apple controls the entire iOS ecosystem, Apple devices tend to be more secure. They should be, given the $1000 price of the new iPhone! Practicality and economics will cause most of the work to remain on Android, and this is okay, as long as users are vigilant. Technology strategies that rely on Android OS, especially those that handle sensitive information, must be carefully managed. Android devices can indeed be as secure as iOS devices if managed correctly [55].

Pitfalls of digital data management in research settings

In addition to the previous cautions on education and training, project management, proper sponsorship and staff involvement, there are specific cautions that must be understood when implementing digital data collection technologies. First there are the logistical requirements of managing devices, managing users and protecting against theft and misuse. Next are the operational and technical requirements of ensuring that devices can be properly charged, cleaned and kept in good working order. In highly challenging physical environments that may be wet, dry or dusty, proper protective casings and an adequate store of spare devices is required. Lastly, it is undeniable that these technologies and the corresponding preparation and management add financial cost to any research study, but the return on these investments can be extraordinary.

Data and study transparency

It has been argued that the attainable minimal quality standard in epidemiology is reproducibility, and that availability of data sets, software, detailed protocols and statistical approaches enables the types of critical evaluation that ensure study quality and transparency [56]. Maximum transparency is considered an essential element of ethical research as it ensures people are treated properly and that the research itself was conducted with the best interests of the most vulnerable in mind [57]. For industry sponsored clinical trials, transparency via access to data, protocols and results is expected and mandated, although performance varies greatly [58]. While there is clearly much room for improvement in the way that industry shares clinical trial data, the fact is that industry does systematically share data, and there are no truly comparable sharing efforts within academia or the humanitarian sector. Commitment to open sharing of study data would truly raise all boats with respect to the perceptions and concerns regarding the ethics of conducting research in humanitarian settings, and technology can only help.

Data that has been systematically collected, properly managed and evaluated using rigorous statistical methods can be readily examined and evaluated by editors, reviewers and other researchers. Studies that lack well-controlled source data have inadequate chain of custody and lack procedural rigor account for a great proportion of irreproducible research. In contrast, simple checklists have been shown to improve methodological information such as randomization, sample-size calculation and blinding [59]. As previously mentioned, checklists can be readily automated via digital means and can even be improved upon as data quality and completeness can be managed as mandatory.

In addition to ensuring quality, digitally shared data can be aggregated, aligned and pooled or even co-located to produce rich new sources of knowledge. A common driver of these efforts is to facilitate knowledge sharing in hopes of preventing future humanitarian crisis. One such effort is underway to pool data from the recent Ebola outbreak in West Africa. The issues of data ownership, control and access all must be settled [60].

Conclusions

While digital health is imperfect and still in its adolescence, the field is rapidly evolving. New digital studies and capabilities are being reported almost daily, and many have the ability to improve the ethical conduct of research in humanitarian settings. By automating chain of custody of data, by using smart metadata and by exploiting the other inherent capabilities of digital technologies, the quality and conduct of research in humanitarian settings can improve. The change will not be easy, but the rewards appear worth the risk.

The decision to conduct research in humanitarian settings is incredibly complex, and a case can often be made against intervention. However, when the decision to intervene is made, that intervention must be thorough and profound, as each clinical interaction happens only once and is irreplaceable.

Abbreviations

*e*Health: Electronic health; *e*ID: Electronic ID; HIPAA: Health insurance portability and accountability act; ID: Identification; IOS: Apple proprietary operating system; *m*Heath: Mobile health; NGO: Non-governmental organizations; PHI: Personal health information; PII: Personally-identifiable information; TMH: Tele-mental health; WiFi: Wireless network

Acknowledgements

The author's primary institutions as listed on title page and Takeda R&D (former employer), Tyler Finkelstein (proof reader).

Authors' contributions

The author read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

This submission was approved by author's primary institutions.

Competing interests

The author declares that he has no competing interests.

Publisher's Note

Author details

[1]Datavant Inc., San Francisco, CA, USA. [2]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA. [3]Manson Unit, Médecins Sans Frontières, London, UK.

Received: 22 September 2017 Accepted: 5 April 2018

References

1. Ford N, Mills EJ, Zachariah R, Upshur R. Ethics of conducting research in conflict settings. Confl Heal. 2009; https://doi.org/10.1186/1752-1505-3-7.
2. Schopper D, Upshur R, Matthys F, Singh JA, Bandewar SS, Ahmad A, van Dongen E. Research ethics review in humanitarian contexts: the experience of the independent ethics review Board of Medecins Sans Frontieres. PLoS Med 2009; doi: https://doi.org/10.1371/journal.pmed.1000115.
3. Schopper D, et al. Innovations in research ethics governance in humanitarian settings. BMC Med Ethics. 2015; https://doi.org/10.1186/s12910-015-0002-3.
4. Waldman RJ, Toole MJ. Where is the science in humanitarian health? Lancet. 2017; https://doi.org/10.1016/S0140-6736(17)31275-8.
5. Blanchett K, et al. Health in humanitarian crisis 1: evidence on public health interventions in humanitarian crisis. Lancet. 2017; doi.org/10.1016/S0140-6736(16)30768-1
6. Checchi F, et al. Health in humanitarian crisis 2: public health information in crisis-affected populations: a review of methods and their use for advocacy and action. Lancet. 2017; doi.org/10.1016/S0140-6736(17)30702-X
7. Cancedda C, et al. Strengthening health systems while responding to a health crisis: lessons learned by a nongovernmental organization during the Ebola virus disease epidemic in Sierra Leone. J Infect Dis. 2016; 214(suppl 3):S153–63.
8. Mesmar S. The impact of digital technology on health of populations affected by humanitarian crisis: recent innovations and current gaps. J Public Health Policy. 2016; doi.org/10.1057/s41271-016-0040-1
9. Jobanputra K. Electronic medical records in humanitarian emergencies - the development of an Ebola clinical information and patient management system. F1000 Res. 2017; https://doi.org/10.12688/f1000research.8287.3.
10. Perakslis ED. Strategies for delivering value from digital technology transformation. Nat Rev Drug Discov. 2017; https://doi.org/10.1038/nrd.2016.265.
11. National Academies of Science. Integrating clinical research into epidemic response: the Ebola experience. Washington DC: National Academies Press; 2017.
12. Provazza A. How mobile health technology aided Ebola response. Tech Target; 2016. https://searchmobilecomputing.techtarget.com/news/450297862/How-mobile-health-technology-aided-Ebola-response.
13. JeFee- Bahloul H, Duchen D, Barkil-Oteo A. Attitudes towards implementation of store-and-forward Telemental health in humanitarian settings: survey of Syrian healthcare providers. Telemed e-Health. 2016; https://doi.org/10.1089/tmj.2015.0021.
14. Wright V, Dalwai M, Smith RV, Jemmy JP. Médecins Sans Frontières' clinical guidance mobile application: analysis of a new electronic health tool. Public Health Action. 2015; https://doi.org/10.5588/pha.15.0057.
15. Vickery M. The surprising place where cash is going extinct. BBC World - Future Untold World. 2017;13:1–12.
16. Tseng J, Samagh S, Fraser D, Landman AB. Catalyzing healthcare ransformation with digital health: performance indicators and lessons learned from a digital health innovation group. Healthc. 2017; https://doi.org/10.1016/j.hjdsi.2017.09.003.
17. Pare G, Sicotte C, Jaana M, Giroud D. Prioritizing the risk factors influencing the success of clinical information system projects. A Delphi study in Canada. Methods Inf Med. 2008;47(3):251–9.
18. Project Management Institute, Executive Engagement: The Role of the Sponsor. Available: https://www.pmi.org/-/media/pmi/documents/public/pdf/business-solutions/executive-engagement.pdf. Accessed 19 Jan 2018.
19. Panyard DJ, Ramly E, Dean SM, Bartels CM. Bridging clinical researcher perceptions and health IT realities: a case study of stakeholder creep. Int J Med Inform. 2018; https://doi.org/10.1016/j.ijmedinf.2017.11.014.
20. Ostrovsky A, Deen N, Simon A, Mate K. A framework for selecting digital health technology. Cambridge, MA: Institute for Healthcare Improvement; 2014. http://www.ihi.org/resources/Pages/Publications/AFrameworkforSelectingDigitalHealthTechnology.aspx. Accessed 15 Aug 2017
21. Behns S, Lau L, Sarrazin H. The new tech talent you need to succeed in digital: McKinsey & Co; 2016. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-new-tech-talent-you-need-to-succeed-in-digital. Accessed 20 Jan 2018.
22. Half R. 6 basic SDLC methodologies: which one is best?. 2017. Available: https://www.roberthalf.com/blog/salaries-and-skills/6-basic-sdlc-methodologies-which-one-is-best. Accessed 20 Jan 2018.
23. Gostin L. Ethical principles for the conduct of human subject research: population-based research and ethics. Law Med Healthc. 1991;19(3–4): 191–201. Fall-Winter.
24. CIOMS Working Group. International ethical guidelines for health-related research involving humans. Geneva, Switzerland: CIOMS; 2016.
25. Hunt M, Pringle J, Christen M, Eckenwiler L, Schwartz L, Dave A. Ethics of emergent information and communication technology applications in humanitarian medical assistance. Int Health. 2016; https://doi.org/10.1093/inthealth/ihw028.
26. Taylor L. The ethics of big data as a public good: which public? Whose good? Phil Trans A Math Phys Eng Sci. 2015; https://doi.org/10.1098/rsta.2016.0126.
27. Woodward A, Sheahan K, Martineau T, Sondrp E. Health systems research in fragile and conflict affected states: a qualitative study of associated challenges. Health Res Policy Syst. 2017; https://doi.org/10.1186/s12961-017-0204-x.
28. Sheather J, et al. A Medecins sans Frontieres ethics framework for humanitarian intervention. PLOS Med. 2016; doi.org/10.1371/journal.pmed.1002111
29. Asi YM, Williams C. The role of diital health in making progress toward Sustainable develop goal (SDG)3 in conflict-affected populations. Int J Med Inform. 2017; https://doi.org/10.1016/j.ijmedinf.2017.11.003.
30. Berens J, Mans U, Verhulst S. Mapping and comparing responsible data approaches. New York: GOVLAB Centre for Innovation; 2016.
31. Kadam RA. Informed consent process: a step towards making it meaningful. Perspect Clin Res. 2017; https://doi.org/10.4103/picr.PICR_147_16.
32. Aguila E, Weidmer BA, Illingworth AR, Martinez H. Culturally competent informed-consent process to evaluate a social policy for older persons with low literacy: the Mexican case. Sage Open. 2016; https://doi.org/10.1177/2158244016665886.
33. Grady C, et al. Informed consent: the changing face of informed consent. NEJM. 2017; https://doi.org/10.1056/NEJMra1603773.
34. Simonds VW, Garroutte EM, Buchwald D. Health literacy and informed consent materials: designed for documentation, not comprehension of Health Research. J Health Commun. 2017; https://doi.org/10.1080/10810730.2017.1341565.
35. Cummings SR, Rowbotham MC. Electronic informed consent and internet-based trials. NEJM. 2017:859–61.
36. Perrenoud B, Velonaki VS, Bodenmann P, Ramelet AS. The effectiveness of health literacy interventions on the informed consent process of health care users: a systematic review protocol. JBI Database Syst Rev Implement Rep. 2015; https://doi.org/10.11124/jbisrir-2015-2304.
37. Moore S, Tasse AM, Thorogood A, Winship I, Zawati M, Doerr M. Consent processes for mobile app mediated research: systematic review. JMIR Health Uhealth. 2017; https://doi.org/10.2196/mhealth.7014.
38. Abay S, Addissie A, Davey G, Farsides B, Addissie T. Rapid ethical assessment on informed consent content and procedure in Hintalo-Wajirat, northern Ethiopia: a qualitative study. PLoS One. 2016; https://doi.org/10.1371/journal.pone.0157056.
39. U.-D. DIVISION FOR SUSTAINABLE DEVELOPMENT. Sustainable development goal 16. 2016. Available: https://sustainabledevelopment.un.org/sdg16. Accessed 15 Sept 2017.
40. Griffith E. How to wipe yourself off the grid. PCMag. 2012:1–3. https://www.pcmag.com/article2/0,2817,2376024,00.asp.
41. Perakslis E. Cybersecurity in health care. NEJM. 2014; https://doi.org/10.1056/NEJMp1404358.
42. Liu V, Musen MA, Chou T. Data breeches of protected health information in the United States. JAMA. 2015; https://doi.org/10.1001/jama.2015.2252.
43. Berkley S. Solving a global identity crisis. MIT Technol Rev. JAMA. 2015; 313(14):1471-3. https://doi.org/10.1001/jama.2015.2252.
44. Merrell RC, Doarn DR. Identity theft, a reprise. Telemedicine and e-health; 2017. https://doi.org/10.1089/tmj.2017.29010.crd.
45. Ambadiyil S, Soorej KS, Mahadevan Pillai VP. Biometric based unique ID generation and one to one verification for security documents. Procedia Comput Sci. 2015; doi.org/10.1016/j.procs.2015.02.075
46. HHS.gov. Guidance regarding methods for De-Identfication of protected health information in accordance with the health insurance portability and accountability ACT (HIPAA) privacy rule, 2012. https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/De-identification/hhs_deid_guidance.pdf. Accessed 5 Sept 2017.
47. Gemalto, Digital identity trends – 5 forces that are shaping 2017, 2017. Available: http://www.gemalto.com/govt/identity/digital-identity-trends. Accessed 14 Sept 2017.
48. NIST. Digital identity guidelines: now available. 2017. https://pages.nist.gov/800-63-3/. Accessed 4 Sept 2017.

49. Clark J, et al. Digital identity: towards shared principles for public and private sector cooperation. Washington DC: World Bank Group; 2016.
50. Sahoo SS, Valdez J, Rueschman M. Scientific reproducibility in biomedical research: provenance metadata ontology for semantic annotation of study description. AMIA Annu Symp Proc. 2016;2016:1070–9. eCollection 2016
51. Imran M, Hlavacs H, Haq IU, Jan B, Khan FA, Ahmad A. Provenance based data integrity checking and verification in cloud environments. PLoS One. 2017; https://doi.org/10.1371/journal.pone.0177576.
52. Kazic T. Ten simple rules for Experiments' provenance. PLoS Comput Biol. 2015; https://doi.org/10.1371/journal.pcbi.1004384.
53. Perakslis ED, Stanley M. A cybersecurity primer for translational research. Sci Transl Med. 2016; https://doi.org/10.1126/scitranslmed.aaa4493.
54. Perakslis E. Cyber threat, protection and resilience in humanitarian relief and response. Lancet Global Health Blog. 2017;
55. Norton, Symantec. Android vs iOS: which is more secure? Available: https://us.norton.com/internetsecurity-mobile-android-vs-ios-which-is-more-secure.html. Accessed 15 Sept 2017.
56. Coughlin SS. Reproducing epidemiologic research and ensuring transparency. Am J Epidemiol. 2017; https://doi.org/10.1093/aje/kwx065.
57. Mandal J, Parija SC. Ethics of sponsored research. Trop Parasitol. 2014; https://doi.org/10.4103/2229-5070.129140.
58. Goldacre B, et al. Pharmaceutical companies' policieson access to trial data, results and methods: audit study. BMJ. 2017; https://doi.org/10.1136/bmj.j3334.
59. Han S, et al. A checklist is associated with increased quality of reporting preclinical biomedical research: a systematic review. PLoS One. 2017; https://doi.org/10.1371/journal.pone.0183591.
60. Maxmen A. Massive Ebola data site planned to combat outbreaks, Nature news; 2017. p. 1–4.

Vital Signs: Digital Health Law Update | Winter 2020 – Targeted Excerpts

Jones Day

Welcome to the first issue of "Vital Signs"—a refreshed and reframed revival of Jones Day's Digital Health Law Update, given our long-standing commitment, experience, and leadership in digital health. We plan to issue Vital Signs near the beginning of each quarter to cut through all the noise and provide access to a curated resource on "must know" legal and regulatory topics impacting digital health.

As you will see in this first issue, our team of contributors from around the globe culled through various information across legal specialty areas (cybersecurity, health regulatory, IP, litigation, tax, and transactional) to present complex issues in a user-friendly fashion. Hot links to underlying legal and regulatory materials are included within each section, enabling Vital Signs to serve as your ongoing quick reference for all the notable recent legal and regulatory materials.

In this issue, we start with "Industry Insights"—dedicated in each issue to an overarching topic of industry interest—covering the evolving nature of how disparate industry stakeholders are advancing their vision, collaboration, and utilization of digital health, especially in the decade ahead. You'll also learn about: (i) U.S. federal efforts focused on streamlining FDA's digital health approach and advancing Medicare's telehealth reimbursement, while enforcement attention and methods evolve; (ii) U.S. states' legislative attention around telehealth and data privacy; (iii) Europe's flurry of digital health policy covering privacy, cross-border services, device development, and interoperability; (iv) Mexico's adoption of a novel tax on digital intermediaries supporting telemedicine; and (v) Japan's adoption of significant modifications to its Pharmaceuticals and Medical Devices Act.

INDUSTRY INSIGHTS

Worlds Colliding—A New Decade for Digital Health

Over the past decade, unprecedented advances in technology and know-how have resulted in an ever-expanding and diversified "digital health" universe. From the growing use of communication technologies to deliver clinical care, to the rise in fitness and healthy living apps, to improvements in diagnostic medical devices, digital innovation has increasingly defined many of our health care experiences over the past 10 years. As we enter a new decade, it is certain this trend will continue.

We expect the next 10 years, however, to encompass something more than continued vertical growth. If the last decade was defined by innovation and forward advancement in numerous yet siloed sectors of the health care industry, the next decade is likely to be defined by horizontal growth—a growing overlap in disparate digital health worlds to create one, integrated digital health ecosystem. We began to see some of this overlap at the tail end of the 2010s.

For example, traditional telehealth services are starting to incorporate medical devices (particularly diagnostic and triaging software) as part of patient-facing platforms. Software has long been used for administrative functions in health care settings. In the last few years, however, we began to see new types of clinical decision support ("CDS") software being utilized for things like patient screening, as well as for assisting providers in identifying the most appropriate treatment plan for a patient's disease or condition. To date, these functionalities have been relatively simplistic, but with the rise of machine-learning algorithms and AI, it seems certain that more sophisticated medical software will soon be upon us. We anticipate that CDS software will become a ubiquitous and indispensable aspect of patient care in the 2020s.

Digital health is also being utilized to bring therapies to market. Sponsors of clinical trials increasingly use digital means, such as data mining through claims and electronic health records, to more rapidly and comprehensively identify, recruit, and engage with clinical trial participants. The use of wearable technologies has also improved the means of collecting and transmitting clinical data to trial investigators. Together with the more traditional means of communicating patient reported outcomes, the use of data from wearables enhances a sponsor's ability to demonstrate the achievement of applicable clinical trial endpoints. In addition, technological advances are being further deployed to support the development of "decentralized" or "virtual" clinical trial protocols, in which an increasing percentage of trial procedures are administered from a patient's home or other remote locations.

The promise of digital health in each of these contexts is significant. These developments could result in less expense, more expedient and more accurate outcomes, and better engagement with subjects, patients, and consumers. But these developments, of course, also give rise to attendant practical and legal challenges, particularly given the different rules and regulators involved. For example, the use of sophisticated medical software and data mining gives rise to privacy and security concerns and may require the use of special informed consents. Likewise, while clinical trials are typically regulated at the federal level, the rise in "virtual" clinical trial protocols may require careful consideration of state telehealth and practice of medicine rules. While the development of technology invariably lurches ahead of regulation, all stakeholders will need to work together to advocate for continued evolution of the regulatory process to allow for a responsible deployment of innovative technologies to improve patient care.

UNITED STATES DEVELOPMENTS

Federal

FDA Proposes New Guidance on Clinical Decision Support ("CDS") Software

In 2016, the 21st Century Cures Act ("Cures Act") explicitly excluded CDS software from FDA regulatory authority. Due to the impressive diversity of digital products that provide some type of clinical decision support (including many that incorporate artificial intelligence and machine learning algorithms), industry stakeholders have since grappled with significant uncertainty about whether certain applications might qualify as CDS software or, alternatively, fall under FDA's regulatory authority. FDA initially issued draft guidance in 2017 to help resolve some of this uncertainty but later withdrew that draft guidance in response to public feedback. FDA recently issued new draft guidance in September 2019 to again address the lingering uncertainty. The September 2019 draft guidance identifies three distinct categories of software: (i) that which meets the Cures Act definition of "CDS software" and, therefore, falls outside of FDA regulatory authority; (ii) that which falls under FDA's regulatory authority but over which FDA intends to exercise "enforcement discretion" due to the lower risk of the applications; and (iii) that which falls under FDA's regulatory authority and over which FDA intends to focus its regulatory oversight. Industry should closely monitor FDA's oversight activities with regard to CDS software in 2020 to gain further insight into FDA's interpretation of its regulatory reach.

FDA Modernizes Regulations to Facilitate Utilization of Digital Technologies in Clinical Trials

In 2019, FDA continued its initiatives focused on modernizing the clinical trial process by issuing numerous draft and final industry guidance supporting the use of adaptive and complex innovative trial designs, novel trial endpoints and enrichment strategies, and duly qualified biomarkers and other drug development tools. Certain of these initiatives are pursuant to mandates under the Cures Act and incorporate evolving technologies such as complex algorithms and digital tools. FDA also recently articulated policies intended to provide patients with a more pronounced voice in clinical investigation design and to permit industry to mine and utilize certain real-world evidence in the drug development process. FDA provides additional information regarding real-world evidence and patient engagement on its website.

Proposed CONNECT Act Transfers Telehealth Reimbursement Eligibility Decisions to Regulators

The CONNECT for Health Act of 2019 ("CONNECT Act") was introduced into the U.S. Congress House and Senate in October 2019. If adopted, the Connect Act could significantly expand reimbursement for telehealth services provided to Medicare beneficiaries. Specifically, the CONNECT Act would empower the U.S. Department of Health and Human Services ("HHS") to waive the specific geographic, modality, provider type, and other limitations currently in effect under §1834(m) of the Social Security Act. HHS may grant such waivers where proposed outlays: (i) reduce spending without reducing quality; (ii) improve quality without increasing spending; or (iii) serve health shortage areas. The CONNECT Act would also eliminate the rural geographic limitation for tele-mental health services and tele-specialist consultations in the emergency department, all §1834 limitations for hospice recertifications following an in-person visit, and the originating site limitation for certain Indian Health Service and Native Hawaiian Health Care facilities.

Federal Criminal and Civil Investigations Touch on Digital Health Industry

Two major criminal investigations by the FBI, HHS-OIG, and DOJ involved the digital health industry in 2019: Operation "Double Helix" and the related Operation "Brace Yourself." These investigations were unprecedented in the digital space for their size and scope. Operation Double Helix involved $2.1 billion in alleged losses resulting from fraudulent billing of genetic tests, orthotic braces, and pain creams. Notably, CMS issued its own press release highlighting its involvement in this enforcement activity through the Center for Program Integrity, publicly flexing its enforcement muscle. Operation Brace Yourself involved $1.2 billion in alleged telemedicine fraud among companies that sell durable medical equipment ("DME") across the country. Charges were brought against at least 24 defendants including corporate entities, health care providers, and corporate executives, and they involved allegations of commercial/nongovernment program fraud in addition to federal program fraud. These investigations, along with the HealthRight investigation announced in 2018, highlight the government's continuing focus on areas of long-standing concern (compound drugs, DME, and lab testing), but in the digital space.

Government civil enforcement activity in the digital health space appears to be on the uptick, with a continued focus on compounding prescriptions reimbursed by TRICARE and enforcement against electronic health record ("EHR") companies. Diabetic Care Rx LLC, two executives, and private equity firm Riordan, Lewis & Haden Inc. settled a qui tam lawsuit for $21.36 million in September 2019, following up on the DOJ's intervention decision in early 2018. In February 2019, EHR vendor Greenway Health LLC entered into a settlement for $57 million that the DOJ independently pursued, and a few months later IBM and Cúram Software entered into a settlement for $14.8 million related to misrepresentations made during the development of Maryland's Health Insurance exchange website and IT platform.

Proposed Stark and Anti-Kickback Regulatory Reforms Notable for Potential Broad Health Industry Implications

In October 2019, OIG and CMS published large packages of proposed reforms to modernize the regulations that interpret the federal Anti-Kickback Statute and the federal Stark Law, respectively. Jones Day is in the process of publishing a series of *Commentaries* to summarize the more significant proposals within the packages. Though only a narrow set of the proposed reforms specifically impact the digital health industry, all proposed reforms impact the health care space generally and may have implications for various digital health initiatives. See our *Commentaries* on the New Exception for Dialysis-Related Telehealth Technologies, Changes to Valuation Terms Under the Federal Stark Law, and Newly Proposed Protections for Cybersecurity Technology Under the AKS and Stark Law.

Express Scripts Unveils Stand-Alone Digital Health Formulary

Express Scripts, one of the nation's largest pharmacy benefit managers, unveiled the industry's first stand-alone digital health "formulary" in December 2019. The initial cohort of digital health solutions on the formulary are designed to aid in the management of common chronic conditions, including diabetes, hypertension, asthma, and depression. This initial cohort includes specific "preferred" and "alternative" vendors for each of these various chronic conditions, with solutions that engage with patients in a variety of ways (e.g., remote monitoring of a patient's use of a drug therapy, synching a diagnostic device to provide vital information to patients and alert them of potential concerns, virtual personal coaching from a health or wellness professional, and virtual support from a patient's peer). A team of physicians, pharmacists, and experts in health research assisted Express Scripts in developing the formulary, with the goal of ensuring that each digital health solution on the formulary is cost-effective, demonstrates therapeutic value, and meets privacy and security standards. In an era where digital health solutions are rapidly expanding and the global digital health market is expected to reach a valuation of more than $234 billion in 2023, the Express Scripts digital health formulary seeks to: (i) provide patients with greater access to validated digital health solutions for their chronic conditions; (ii) reduce the administrative burden for plan sponsors in contracting with digital health companies; and (iii) provide a pathway for digital health companies to get their products covered and to patients.

GLOBAL DEVELOPMENTS

Europe

Multiple Advances for Electronic Health Records and Digital Health Interoperability

On February 6, 2019, the European Commission adopted Recommendation (EU) 2019/243 (available in English), which sets out a framework for the development of an EU electronic health record exchange format to enable the secure, interoperable, and cross-border access to and exchange of electronic health data. The framework includes a series of principles on access and exchange of electronic health records and a set of common technical specifications, and it encourages Member States to ensure secure access to electronic health record systems at the national level.

On July 10, 2019, MedTech Europe—the European trade association representing the medical technology industries—released a position paper welcoming the European Commission's Recommendation and calling on European health authorities, payers, and procurers to "engage and contribute to an interoperable data ecosystem" (available in English). MedTech Europe also asked for more public investment in digital health infrastructures, including electronic health record systems, and requested the European Union to provide economic resources for this purpose.

On June 12, 2019 the eHealth Network—a network set up under Article 14 of Directive 2011/24/EU and composed of members of EU national authorities responsible for eHealth—adopted a document providing "Guidelines on an interoperable eco-system for digital health and investment programmes for a new/updated generation of digital infrastructure in Europe" (available in English). The document aims to guide the European Commission, Member States, and industry toward funding and purchasing choices that foster interoperability in digital health.

New Guidelines Issued Regarding Software as a Medical Device

On October 11, 2019, the European Commission issued guidelines regarding software under the new EU Regulations on medical devices: "Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745–MDR and Regulation (EU) 2017/746–IVDR" (available in English). The Guidance provides manufacturers with useful guidance on both qualification (i.e., when a software is considered to be a medical device) and classification (under which category the product falls depending on the risk of harm the device poses to patients (Class I, IIa, IIb, or III)). The Guidance contains a list of examples of software currently used in the health care industry that qualify as medical devices. The document also clarifies the rules on classification laid down in the new EU Regulations. Finally, the Guidance underlines the importance of gathering solid clinical evidence in order to support devices' claims, as well as collecting post-market data for demonstrating software's efficacy and safety over time.

New Guidance Available Regarding the Cybersecurity of Medical Devices

In December 2019, the Medical Device Coordination Group—a group composed of representatives of all Member States and chaired by a representative of the European Commission—adopted a new document "Guidance on Cybersecurity for medical devices" (available in English). The document provides assistance to manufacturers regarding how to meet the cybersecurity requirements of Annex I to the new Medical Device (available in English) and In-Vitro Medical Device Regulations (available in English), to ensure that devices are protected against unauthorized access.

ENISA Issues a Status Report on Incident Response Development

In November 2019, the European Union Agency for Cybersecurity ("ENISA") issued a report on the status of EU Member States incident response development. The report provides an analysis of the current operational Incident Response setup within the Directive on Security of Network and Information Systems sectors, including the health care sector.

Accelerating Precision Health by Applying the Lessons Learned from Direct-to-Consumer Genomics to Digital Health Technologies

Joyce Y. Tung, PhD, 23andMe; **Ryan J. Shaw, PhD, RN,** Duke University School of Nursing; **Jill M. Hagenkord, MD, FACP,** Color Genomics; **Meredith Hackmann,** National Academies of Sciences, Engineering, and Medicine; **Marion Muller,** Emory University; **Sarah H. Beachy, PhD,** National Academies of Sciences, Engineering, and Medicine; **Victoria M. Pratt, PhD,** Indiana University; **Sharon F. Terry, MA,** Genetic Alliance; **Ann K. Cashion, PhD,** National Institute of Nursing Research; **Geoffrey S. Ginsburg, MD, PhD,** Duke University School of Medicine

March 19, 2018

Emerging Opportunities for Digital Health

Digital health technologies have emerged in recent years, providing opportunities for engaging with consumers and integrating with the health care system to realize precision medicine, defined as “an evidence-based approach to the care of people and patients that uses innovative tools and data science to customize disease prevention, detection, and treatment; improve the effectiveness and quality of care; and sustain health” [1]. These technologies—including smartphone applications (apps), wearables (e.g., Fitbit, Apple Watch), and mobile-device-based tools—can make health care and management of health and wellness more efficient for patients and clinicians and can support research. Digital health technologies offer consumers, patients, and clinicians the opportunity to interact with personal health data and the health care system in a way that has not been available previously.

One such tool—the fertility app, Glow—helps users who are trying to conceive track data such as ovulation, physical symptoms, and diet and exercise [2]. The company has amassed a wealth of user-reported data and, in a study, claimed that conception rates are significantly higher in the population that frequently uses the app compared to those who do not [3]. Other health tracking apps increasingly popular among consumers track user-reported food and water intake, as well as steps walked, heart rate, and sleeping patterns by pairing with a wearable device. From 2014 to 2016, wearable use increased by 12 percent, reflecting use by 21 percent of consumers surveyed across seven countries [4]. One wearable company, Fitbit, increased its active user base from 16.9 million in 2015 to 23.2 million by the end of 2016 [5].

While tracking devices are popular with physically active individuals, they can also be useful for helping consumers with chronic diseases such as chronic heart failure. For example, a wireless scale can be used to track weight and a diet app to record nutritional information—behaviors important for controlling their condition. Other digital health apps have shown potential to improve health care decision making and management, including glucometers that can connect with smartphones to help those with type 1 diabetes and sensors that can attach to inhalers to help those with asthma or chronic obstructive pulmonary disease. According to a 2016 survey, 76 percent of patients who were instructed by a physician to use a wearable to track lifestyle, fitness, or vitals complied with the doctor’s request [4].

In the field of genomics and precision health, digital health apps are also making their way onto the market. For instance, Helix launched a DNA app store in July 2017 that acts as a “one stop shop” for interested individuals to interact with their genomic data [6]. The Scripps Translational Science Institute’s MyGeneRank study is using genetic information received from users of 23andMe, a personal genetics company [7], to assess genetic risk for coronary heart disease and reporting it back through a mobile app [8]. Other research efforts, such as the All of Us Research Program, hope to use wearables and other home-health technologies to gather environmental and personal health measurements to correlate with patient outcomes data [9,10].

Given the short amount of time that mobile health apps have been on the market, and their growth to date, it is not unreasonable to assume that these will continue to be adopted and grow in popularity and use among consumers as other technologies have done. While there is potential for these technologies to improve health, there are still many challenges that must be addressed, including data quality, consumer access to data, others' access to data, data use, and regulatory oversight. The direct-to-consumer (DTC) genomics industry, which has been increasingly moving into the clinical health space over the last 10 years, offers insights for digital health technologies that may help companies navigate integration within health systems. Like digital apps, DTC genomics companies like 23andMe and Helix have become increasingly common, offering consumers the opportunity to gain insights into their genetics. By providing health-related data with potential applications in health care and encouraging patients to take a more active approach in their health decision making, these DTC genomics companies represent a new hybrid business model combining a traditional medical device and consumer product. As DTC genomics has evolved, the industry has had to grapple with many of the same issues now facing the digital health industry. Here we review a number of these key issues and identify recommendations for those working in the digital health space.

Key Challenges

Regulatory Uncertainty

The regulatory uncertainty surrounding laboratory developed tests (LDTs) created a challenge for DTC genomic testing companies from the outset. Although LDTs first came under FDA authority in 1976, the agency had for decades exercised "enforcement discretion," meaning that the laboratories developing and performing the tests were not required to obtain FDA clearance or approval for marketing [20].

Many companies incorrectly assumed that DTC genetic tests, which often were also LDTs, would fall under this same enforcement discretion and thus were unprepared for the regulatory challenges that followed [13] (*see Box 1*).

A key lesson for app developers is the importance of understanding the regulatory landscape, the differences in regulatory categories (i.e., drugs, medical devices, and diagnostics), and their place in it in order to assess what requirements they will need to fulfill to market their products and services. For example, it may be important to understand and plan for what is needed to demonstrate analytical and clinical validity at the beginning of product development. The FDA's 2013 warning letter to 23andMe cited the company's failure to provide sufficient information to support the analytical and clinical validity of the test. This requirement was predicated on the FDA's assessment that the 23andMe product was a medical device and thus required to meet certain standards for that type of regulatory category, which the FDA has subsequently published.

Many parallels can be seen between the regulation of DTC tests and mobile health tools, the latter referred to by the FDA in three categories: mobile platforms, mobile apps, and mobile medical apps. Mobile medical apps meet the definition of a device; mobile platforms are handheld commercial off-the-shelf computing platforms, such as smartphones and tablets; and mobile apps are software applications that are run on mobile platforms. Current FDA regulations explain that the FDA uses a tiered approach to oversight [21]. The agency's focus is on apps and devices that meet the definition of a medical device and that present a significant risk to patients if the apps or devices do not work as intended. The FDA does not apply regulatory requirements to mobile health tools that are not considered mobile medical apps. For example, a mobile app may potentially be used as a tool to help an individual be aware of their risk for skin cancer based on an image provided by the end user; however, the app would have to meet much stricter regulatory requirements if it were to diagnose that individual with skin cancer and/or provide a treatment plan, rather than give general guidance that the user should consult a physician. Regulation of digital health products will necessitate developers being aware of appropriate regulatory pathways, specific regulatory requirements (or lack thereof), and what will be needed to demonstrate performance.

Consumers versus the Health Care System

Analytical Validity, Clinical Validity, and Clinical Utility

Central to the question of how DTC health information should be integrated into health care are issues

Box 1 | **The Rise of Direct-to-Consumer Genomics**

Companies began offering direct-to-consumer (DTC) genetic testing in or around the early 2000s, primarily focusing on nutrigenetics. This drew the scrutiny of multiple federal agencies, leading to a 2006 US Government Accountability Office (GAO) report that raised serious concerns about the accuracy and interpretation of DTC genetic tests, as well as their lack of regulatory oversight [11]. In an effort to educate consumers about the potential risks and benefits of purchasing at-home genetic tests, the Federal Trade Commission (FTC) published a general guidance, drawing on information from the US Food and Drug Administration (FDA) and Centers for Disease Control and Prevention. The guidance advised consumers to approach these tests with skepticism and encouraged consumers to involve physicians when interpreting their test results [12]. The guidance highlighted some of the potential concerns for consumers engaging with these tests, including analytical and clinical validity and clinical utility. The points on which this guidance touched remain at the forefront of the DTC genomics discussion.

By 2007, companies such as deCODE Genetics and 23andMe began to market more complex DTC genetic tests to the public [13], sparking much interest among consumers and concern among regulators about the potential benefits and harms. Companies provided health reports that linked a participant's genotype results to associations with disease risks, as well as physical and personality traits. In 2008, the DTC genetic testing service from 23andMe was named *Time*'s Invention of the Year [14]. The company 23andMe's "$399 saliva test that estimates your predisposition for more than 90 traits and conditions" was considered a pioneer of retail genetic testing [14]. By late 2009, 23andMe and another company, Pathway Genomics, had increased their offerings by adding information that was considered "actionable" for health care decision making, such as testing for rare variants, drug response based on genetics (pharmacogenetics), and carrier screening [15]. New DTC genetic testing companies continued to join the field, despite ongoing regulatory scrutiny. With minimal federal regulations from the Clinical Laboratory Improvement Amendments (CLIA) of 1988 already in place [16], discussion regarding the best way to proceed quickly gained attention and continued for many years.

At the same time, the clinical genetic testing regulatory system continued to face scrutiny. The Secretary's Advisory Committee on Genetics, Health, and Society at the National Institutes of Health released an extensive report related to clinical genetic testing in 2008 that pointed out major gaps in the current regulations and offered solutions to many of these problems [17]. The GAO issued a second report in 2010. Several professional societies also issued statements in opposition to DTC genetic testing, maintaining that such testing should be performed and interpreted by medical professionals. In 2010, the FDA held a public workshop on DTC genetic testing to gather more feedback to inform developing policy. Some claimed that DTC tests did not fall under the FDA's jurisdiction because they reported educational information, however, the FDA's position was always that DTC tests were medical devices. In 2010, the FDA sent letters to several DTC genomics companies informing them that their tests were medical devices that provided data on "genetic predispositions for important health conditions and medication sensitivities," as well as pharmacogenetics results that might influence drug treatment. The letters also indicated that these tests did not have FDA clearance or approval. In 2013, the FDA issued a warning letter to 23andMe that prevented it from marketing any health-related genetic reports until the FDA authorization was obtained. In the letter, the FDA stated that it had given the company various opportunities to comply with feedback on analytical and clinical validity requirements for its tests and potential regulatory pathways to pursue, but the agency had received little indication that the company was taking the necessary steps toward compliance [18].

Though these DTC companies have faced some practical hurdles in recent years, the field continues to learn from early challenges and evolve. For example, 23andMe's Bloom syndrome carrier test report was authorized in 2015, and the FDA exempted this kind of device from premarket review moving forward. Alberto Gutierrez, then director of the Office of In Vitro Diagnostics and Radiological Health at the FDA said, "The FDA believes that in many circumstances it is not necessary for consumers to go through a licensed practitioner to have direct access to their personal genetic information." This was followed, in April 2017, by the FDA's decision to authorize 23andMe to market DTC tests that provide information on an individual's genetic predisposition to certain medical diseases or conditions, such as Parkinson's disease and late-onset Alzheimer's disease [19]. While 23andMe currently has the only DTC genetic test to receive FDA authorization, these recent innovative regulatory decisions define a clear path to market for additional DTC genetic tests and are likely to pave the way for expansion of the DTC genetic test industry.

SOURCE: Tung et al., "Accelerating precision health by applying the lessons learned from direct-to-consumer genomics to digital health technologies," National Academy of Medicine.

concerning the analytical validity, clinical validity, and clinical utility of the information. Briefly, an assay is analytically valid if it accurately and reproducibly measures what it intends to. A test is clinically valid if what it measures is clinically meaningful [22]. Clinical utility describes the ability of a test to lead to a clinical decision with evidence of improved outcomes, and for a digital device that is manufactured to replace an existing device, this is usually measured by direct comparison to the current standard of care (e.g., a device that attaches to a smartphone to measure blood oxygen saturation versus a battery-operated pulse oximeter). For genetic tests, which may assay up to millions of endpoints, identifying the right studies to demonstrate analytical and clinical validity has been challenging [23], and showing clinical utility is an even higher bar that generally requires more complex clinical studies. Studies of health care professionals show that they consistently have concerns about the quality, reliability, and clinical utility of information from DTC genetic testing [24].

Consumers, and oftentimes their clinicians, are generally not able to assess analytic validity and clinical validity. Consumers might misunderstand clinical utility for usefulness in a real-world context. Thus, a key challenge for the DTC genomics and digital health industries will be navigating these differences in understanding of definitions around validity and utility.

Physicians use health information to help diagnose disease and guide treatment, and generally require the accuracy demonstrated by medical-grade devices, which have strong analytical and clinical validity. For example, mobile medical apps that require FDA approval, such as a Bluetooth-enabled glucometer, may be viewed by physicians as valid tools. Consumers, on the other hand, may see the apps as helpful to maintain or improve their health, manage a condition, or gain insights into their personal habits. For these purposes, their standards for the validity of the information they receive may be less stringent. In addition, a consumer may see personal utility in information that has little clinical utility. For instance, a survey by Rock Health found that 47 percent of consumers were willing to pay $50 for a test to identify a predisposition for Alzheimer's disease, which has no cure and for which there are early clinical interventions only in clinical trials, whereas only 28 percent were willing to pay the same amount for a genetic test of metabolism [25].

This gap in clinical utility and consumer expectations can create challenges when consumers bring information from DTC devices and apps to their physicians. Many consumer devices, including many wearables, are not intended for clinical use and thus may have performance that is not validated. For example, a study by Wang et al. found that wrist-based wearable devices used for heart rate monitoring varied widely in terms of accuracy compared to the standard chest strap electrode-based monitors prescribed by many physicians for their cardiac patients [26]. Similarly, while accelerometers such as those found in smartphones may be useful to understand a "gestalt" view of activity and serve as motivational tools for health behavior change when coupled with apps, their accuracy as a clinical-grade tool has not been well described. These distinctions may not be clear to many consumers. In addition, many providers, health systems, and electronic health records are not adequately prepared to handle the large amounts of data from DTC genetic testing or digital health apps. When patients bring this information to their providers, they may expect answers or guidance on how it should influence their health behaviors, even though in many cases evidence-backed clinical guidelines do not exist for this type of information. If, instead, providers are not able to respond to the data in a way that consumers expect, they may become frustrated or disappointed, which may lessen the value of the device or app to the consumer.

Accessibility and Reimbursement

When DTC genomics companies first appeared on the market, the cost of some full screening tests ran upwards of $2,000. Thus, the price point was well out of range for the average consumer [15]. As technological improvements were made, these tests decreased in price, allowing more consumers into the market. While the question of whether these technologies will exacerbate health disparities is still being debated, the need for people of all backgrounds to have access to DTC genomics is important [27]. While much genetic variation is shared among all people, there are important differences across populations, and having greater diversity among consumers of genetic testing will help the industry improve the services it provides to everyone. Digital health technology may offer a significant opportunity to reduce health disparities by affordably increasing access to care. According to the Pew Research Center, 77 percent of Americans own smartphones [28], creating a potential outlet to reach a majority of consumers across socioeconomic and urban and rural divides.

A secondary concern in this area is whether and how these technologies may be reimbursed by third-party payers. While strictly DTC genetic tests have generally stayed out of the reimbursement space, there are physician-prescribed genetic test providers that have taken a hybrid approach by helping people seek reimbursement but limiting an individual's out-of-pocket expense to a certain amount [29]. In addition, as McGuire et al. discussed, in the early stages of DTC genetic testing, reimbursement systems were not ready to handle the challenges of DTC genomics, both in terms of covering the time spent by health care providers reviewing DTC genetic test results, and additional testing prompted by the DTC reported results [30]. Engaging with payers and health care providers may be of value for digital health stakeholders, as it will set expectations for requirements to fully realize the technology in a clinical setting.

Recommendations and Guidelines from Professional Societies

Early on, DTC genetic testing companies were met with caution from professional societies and agencies, including the American Society of Human Genetics (ASHG). ASHG recognized the potential for increased access to testing and greater consumer awareness, but also recommended actions around transparency, provider education, and test and laboratory quality if the field wanted to move forward [31]. For their part, several DTC companies indicated a willingness to work with stakeholder groups to define a set of standards that could help companies in the field self-regulate [13]. As consumers increasingly showed interest in these tests and the evidence base grew, some groups updated their position statements in support of tests that provided meaningful information to patients and proper education to act on the results [32,33].

Because digital health is a relatively young field that is rapidly evolving, gaps exist in guidance and recommendations from professional societies, but key leaders can be engaged to offer direction and encourage innovation. Some groups have sought to help lead the way by creating frameworks and resources for the field, among them the Health App Decision Tree from the Children's Health Fund and an interactive tool created by the FTC. The Health App Decision Tree starts with the question "Is the app selling a product or therapy?" and then goes through a detailed decision-making process of yes-or-no questions until an app developer either gets to an inappropriate app or appropriate app end point [34]. The questions pertain to privacy issues, data settings, and the type of information provided in the app, among others. Another tool created for app developers is the FTC's Mobile Health Apps Interactive Tool, which helps determine which federal laws apply to the app [35]. Based on 10 yes-or-no questions, the tool provides information on the Health Insurance Portability and Accountability Act (HIPAA), and FDA and FTC laws that may or may not apply based on whether the tool is considered a medical device and who the intended user is. For those areas where HIPAA does not apply, Dzau et al. recommend industry implementation of a "digital Hippocratic oath" to help strengthen data security and privacy for consumers using digital health apps [9].

Data Privacy and Data Accessibility

Many digital health apps collect a large amount of private data about their users, which presents issues of data privacy and security. In the DTC genetics world, the Genetic Information Nondiscrimination Act (GINA) and HIPAA both provide privacy protections to patients. Despite these important pieces of legislation, there are still gaps—HIPAA may not apply to all players in the DTC genetic testing or digital health space [35,36], and GINA protects against discrimination in health insurance and employment, but not life insurance. Being the steward of users' personal data may also have other legal ramifications. In 2015, a man named Michael Usry was taken into questioning for an open murder case from 1996, based on a familial DNA search in a publicly searchable database owned by the company Ancestry, which provides customers with genealogical information. When police found similarities between the DNA found at the crime scene and Usry's father, Ancestry was compelled by a court order to provide the name of Usry's father [37].

In addition, as it is not always easy to understand what data are being used by apps and devices, consumers may not be aware of how their information is being used. For example, GPS information could be continuously collected in the background of the app without a consumer realizing or remembering that it is still happening. Even if this information on data collection was given to a consumer when downloading the app or activating it at baseline, the consumer still may not understand what is being presented to them due to digital literacy challenges [38,39]. Furthermore, data collected from sensors connected to mobile devices (e.g., wearables) could be analyzed to predict certain health behaviors such as dental or smoking habits, highlighting potential privacy and insurance implications [40].

Having access to their own data is also important for consumers. A 2016 survey of consumers and physicians showed that 92 percent of patients believe that they should have full access to their health records, while only 18 percent of physicians agree [4]. Similarly, a study by Kaufman et al. showed that the top incentives for people to get involved in a large cohort study included receiving their own lab results, genetic results, and medical records [41]. Government regulations are starting to favor consumers in this regard. The US Department of Health and Human Services states that an individual has the right to access his or her lab results, including the underlying data. This includes, for example, the full underlying gene variant data generated by a lab test that uses next generation sequencing [42]. Increasingly, consumers expect to control their own data [43].

Recommendations and Opportunities

Based on the issues outlined above, we offer the following recommendations for groups developing digital health apps and technologies:

1. Engage early with regulators and professional societies about product development.
2. Identify and articulate well the intended use and audience.
3. Proactively address issues of data privacy and data accessibility.

Engage Early with Regulators and Professional Societies about Product Development

As described above, there remains uncertainty around regulation of apps and services in the digital health space. It is thus valuable to engage with potential regulators early and often in the product development process to determine the appropriate regulatory path, particularly as it relates to generating data to demonstrate the performance of the product. Furthermore, companies may find value in aggregating evidence to substantiate prior claims. Interestingly, in July 2017, the FDA announced new steps to empower consumers and advance digital health care [44]. As part of the Medical Innovation Access Plan, the agency created a new component focused on digital health innovation that included a pre-certification for a software pilot program. The pre-certification would allow companies to submit less information than currently required before marketing new digital health tools.

Still, digital health that goes beyond the FDA classification of a mobile medical app is considered the "wild west" [45]. There is an absence of peer review of apps in health care, which makes ascertaining the quality of an app and verifying its claims challenging for both consumers and health care providers. There is a need for a peer review body or a set of guidelines that clinicians and consumers can look to for guidance. Peer review of apps may also be needed during the peer review process of manuscripts that focus on digital health tools used in studies. Engaging with professional societies can help drive the development of consensus guidelines that will make it easier to develop products and drive the industry forward more quickly.

Identify and Articulate Well the Intended Use and Audience

As consumers and health care providers have different requirements and expectations for digital health products, developers should identify the intended use and audience for their products. A product that is aimed at clinicians, or is intended to produce information a consumer would share with a clinician, may need to demonstrate a certain level of analytical validity, clinical validity, and/or clinical utility before a clinician would find the information valuable. A product that is intended for a consumer's educational or personal use, however, may not need the same level of validation. Identifying who will use the product and how will help define what types of studies need to be conducted to support the product's intended uses.

In addition, identifying who will use the product can help define who should pay for the product and, as appropriate, how it can be integrated in the health system. A product aimed at consumers will involve considerations of cost in terms of accessibility to many consumers. There may be different considerations if a developer wants to seek reimbursement from third-party payers, so engagement with those groups during the development of the product can provide guidance. Given the already limited time and capacity of health care providers, a new model for incorporating data from digital health technologies into health care may be needed, both in terms of clinical workflow and reimbursement models, which, again, may require a multistakeholder approach from those involved in the industry.

Proactively Address Issues of Data Privacy and Data Accessibility

The privacy of their personal health information is important to many consumers, even early adopters [46]. Digital health app developers should proactively address privacy challenges by evaluating potential privacy issues and creating a consumer-centric privacy policy that describes how customer data are used. Some companies even publish a transparency report describing government requests for data [47,48], which many see as a welcome additional step.

To encourage open data-sharing policies and allow consumers to be stewards of their own data, digital health technology developers should build application programming interfaces to make it easier for users to access and share their data with other services. Developers should also consider working together to come up with data standards that will make it easier to share, exchange, and understand the data. Making the data more portable will support the growth of the industry as a whole. Giving consumers dynamic and granular data-sharing, access, and privacy preference tools could accelerate knowledge gain and benefits for health. With the larger amount of personal health data that will be stored, we also note that certain security issues will need to be addressed. In the health care setting, recent data breaches have highlighted some of the vulnerabilities in data security, which may have serious consequences for patients and consumers. Health care organizations have indicated that they lack the infrastructure and expertise to ensure necessary data protections for patients [40]. As new technologies come to the market, it will be important for developers to build security protections within their devices to thwart potential cyberattacks aimed at disabling or manipulating devices.

Future Potential

When the internet was in its infancy, many people believed it would serve as a tool for basic education and research — or, in other words, be a simpler way to communicate. User growth of more than 500 percent between 2000 and 2012 has led to the evolution of an incomprehensibly influential tool that has molded the world — and will continue to do so [49]. Many people believe that precision health has the potential to follow a similar trajectory of growth, and if used effectively, the field of digital health could be a catalyst for accelerating its rise. Similar in complexity to the genome, the internet and smartphones feature a majority of users who are not familiar with how they work. As knowledge and technology advance to widen the range of people participating in genetic testing, an extremely large and influential dataset could be created, providing new opportunities to accelerate research, drug development, and precision health. The DTC genomics industry has had to navigate the challenges of regulatory issues, health care provider and consumer expectations, and privacy issues. Digital health, by learning from some of the lessons learned by DTC genomics, may be able to more efficiently integrate with the health care system and, in turn, help realize the full potential of precision health for the larger population.

References

1. National Academies of Sciences, Engineering, and Medicine. 2015. *Roundtable on Translating Genomic-Based Research for Health: 2015 annual report.* Washington, DC: The National Academies Press.
2. Glow. 2017. *Modern care for your fertility.* https://glowing.com (accessed December 13, 2017).
3. Hartz, A., P. Chenette, E. Scott, and R. Ye. 2015. *Monthly fecundity and the benefits of ovulation tracking in the United States.* What big data tells us about the true nature of fertility. https://assets.documentcloud.org/documents/2488193/asrm-benefit-of-ovulation-tracking-amp-big-data.pdf (accessed October 24, 2016).
4. Safavi, K., and R. Ratliff. 2016. *Patients want a heavy dose of digital.* https://www.accenture.com/t20160629T045303__w__/us-en/_acnmedia/PDF-6/Accenture-Patients-Want-A-Heavy-Dose-of-Digital-Infographic.pdf#zoom=50 (accessed October 24, 2016).
5. FitBit reports $574m Q416 and $2.17b FY16 revenue, sells 6.5m devices in Q416 and 22.3m devices in FY16. 2017. *BusinessWire,* February 22, http://www.businesswire.com/news/home/20170222006497/en/Fitbit-Reports-574M-Q416-2.17B-FY16-Revenue (accessed August 17, 2017).
6. Helix. 2017. *Discover your DNA story*. https://www.helix.com (accessed December 13, 2017).

7. National Institutes of Health (NIH). 2017. Mygenerank: A digital platform for next-generation genetic studies. In: ClinicalTrials.gov. Bethesda, MD: National Library of Medicine. Available from: https://clinicaltrials.gov/ct2/show/NCT03277365 (accessed December 13, 2017).
8. Muse, E. D., N. E. Wineinger, B. Schrader, B. Molparia, E. G. Spencer, D. L. Bodian, A. Torkamani, and E. J. Topol. 2017. Moving beyond clinical risk scores with a mobile app for the genomic risk of coronary artery disease. *bioRxiv*, https://doi.org/10.1101/101519 (accessed August 17, 2017).
9. Dzau, V. J., G. Ginsburg, A. Chopra, D. Goldman, E. D. Green, D. G. B. Leonard, M. McClellan, A. Plump, S. F. Terry, and K. Yamamoto. 2016. Realizing the full potential of precision medicine in health and health care. *NAM Perspectives*. Discussion Paper, National Academy of Medicine, Washington, DC. https://nam.edu/wp-content/uploads/2016/09/Realizing-the-Full-Potential-of-Precision-Medicine-in-Health-and-Health-Care.pdf (accessed July 12, 2017).
10. NIH. 2017. *The Dish: Lifestyle data: Fitbit pilot.* https://allofus.nih.gov/news-events-and-media/videos/dish-lifestyle-data-fitbit-pilot (accessed November 9, 2017).
11. *Nutrigenetic testing—tests purchased from four web sites mislead consumers*. 2006. Washington, DC: Government Accountability Office.
12. Federal Trade Commission (FTC). 2006. *At-home genetic tests: A healthy dose of skepticism may be the best prescription*. https://osp.od.nih.gov/wp-content/uploads/2013/11/DTC_Consumer_Alert_Jul06.pdf (accessed September 7, 2017).
13. Vorhaus, D. 2010. The past, present and future of dtc genetic testing regulation. *Genomics Law Report,* http://www.genomicslawreport.com/index.php/2010/08/05/the-past-present-and-future-of-dtc-genetic-testing-regulation/#2007 (accessed October 17, 2016).
14. Hamilton, A. 2008. Best inventions of 2008. *Time,* October 29, http://content.time.com/time/specials/packages/article/0,28804,1852747_1854493,00.html (accessed October 17, 2016).
15. Pollack, A. 2010. Consumers slow to embrace the age of genomics. *The New York Times,* March 19, http://www.nytimes.com/2010/03/20/business/20consumergene.html?action=click&contentCollection=Business%20Day&module=RelatedCoverage&region=Marginalia&pgtype=article&_r=1 (accessed October 24).
16. Centers for Disease Control and Prevention. 2016. *Clinical Laboratory Improvement Amendments.* http://wwwn.cdc.gov/clia (accessed October 17, 2016).
17. Ferreira-Gonzalez, A., S. Teutsch, M. S. Williams, S. M. Au, K. T. Fitzgerald, P. S. Miller, and C. Fomous. 2008. US system of oversight for genetic testing: A report from the Secretary's Advisory Committee on Genetics, Health and Society. *Personalized Medicine* 5(5):521-528.
18. US Food and Drug Administration (FDA). 2013. *Warning letter to 23andMe.* http://www.fda.gov/iceci/enforcementactions/warningletters/2013/ucm376296.htm (accessed April 5, 2017).
19. FDA. 2017. *FDA allows marketing of first direct-to-consumer tests that provide genetic risk information for certain conditions.* https://www.fda.gov/newsevents/newsroom/pressannouncements/ucm551185.htm (accessed December 13, 2017).
20. Centers for Medicare and Medicaid Services. 2013. *CLIA overview.* https://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/Downloads/LDT-and-CLIA_FAQs.pdf (accessed October 24, 2016).
21. FDA. 2015. *Mobile medical applications.* https://www.fda.gov/MedicalDevices/DigitalHealth/MobileMedicalApplications/default.htm (accessed October 2, 2017).
22. Teutsch, S. M., L. A. Bradley, G. E. Palomaki, J. E. Haddow, M. Piper, N. Calonge, W. D. Dotson, M. P. Douglas, and A. O. Berg. 2009. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: Methods of the EGAPP working group. *Genetics in Medicine* 11(1):3-14.
23. Madsen, J. 2015. FDA regulation catches up to next-gen gene sequencing. *Law360,* http://www.law360.com/articles/705682/fda-regulation-catches-up-to-next-gen-gene-sequencing (accessed September 29, 2016).
24. Goldsmith, L., L. Jackson, A. O'Connor, and H. Skirton. 2013. Direct-to-consumer genomic testing from the perspective of the health professional: A systematic review of the literature. *Journal of Community Genetics* 4(2):169-180.
25. Devos, L., T. Wang, and S. Iyer. 2016. *The genomics inflection point: Implications for healthcare.* https://rockhealth.com/reports/the-genomics-inflection-point-implications-for-healthcare (accessed September 30, 2016).

26. Wang, R., G. Blackburn, M. Desai, D. Phelan, L. Gillinov, P. Houghtaling, and M. Gillinov. 2017. Accuracy of wrist-worn heart rate monitors. *JAMA Cardiology* 2(1):104-106, https://jamanetwork.com/journals/jamacardiology/article-abstract/2566167 (accessed April 4, 2017).
27. Smith, C. E., S. M. Fullerton, K. A. Dookeran, H. Hampel, A. Tin, N. M. Maruthur, J. C. Schisler, J. A. Henderson, K. L. Tucker, and J. M. Ordovas. 2016. Using genetic technologies to reduce, rather than widen, health disparities. *Health Affairs* 35(8):1367-1373.
28. Pew Research Center. 2017. *Mobile fact sheet.* http://www.pewinternet.org/fact-sheet/mobile (accessed July 12, 2017).
29. Farr, C. 2016. Why patients are getting hit with surprise bills after genetic testing. *Fast Company,* April 20, https://www.fastcompany.com/3059072/why-patients-are-getting-hit-with-surprise-bills-after-genetic-testing (accessed October 23, 2017).
30. McGuire, A. L., C. M. Diaz, T. Wang, and S. G. Hilsenbeck. 2009. Social networkers' attitudes toward direct-to-consumer personal genome testing. *American Journal of Bioethics* 9(6-7):3-10.
31. ASHG statement on direct-to-consumer genetic testing in the United States. 2007. *American Journal of Human Genetics* 81:635-637, http://www.ashg.org/pdf/dtc_statement.pdf (accessed July 13, 2017).
32. Association for Molecular Pathology. 2015. *Association for molecular pathology position statement: Direct access genetic testing (direct to consumer genetic testing)—February 2015.* http://old.amp.org/publications_resources/position_statements_letters/documents/AMPpositionstatementDTCtesting-FINAL_002.pdf (accessed July 12, 2017).
33. National Society of Genetic Counselors. 2015. *Direct access to genetic testing.* http://www.nsgc.org/p/bl/et/blogaid=370 (accessed July 12, 2017).
34. Children's Health Fund. 2015 (unpublished). *Health app decision tree.*
35. FTC. 2016. *Mobile health apps interactive tool.* https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-apps-interactive-tool#which (accessed October 24, 2016).
36. Institute of Medicine and National Research Council. 2011. *Direct-to-consumer genetic testing: Summary of a workshop.* Washington, DC: The National Academies Press, https://doi.org/10.17226/13021.
37. Koerner, B. 2015. Your relative's DNA could turn you into a suspect. *Wired,* October 13, https://www.wired.com/2015/10/familial-dna-evidence-turns-innocent-people-into-crime-suspects (accessed October 23, 2017).
38. Mackert, M., A. Mabry-Flynn, S. Champlin, E. E. Donovan, and K. Pounders. 2016. Health literacy and health information technology adoption: The potential for a new digital divide. *Journal of Medical Internet Research* 18(10):e264.
39. Tennant, B., M. Stellefson, V. Dodd, B. Chaney, D. Chaney, S. Paige, and J. Alber. 2015. Ehealth literacy and web 2.0 health information seeking behaviors among baby boomers and older adults. *Journal of Medical Internet Research* 17(3):e70.
40. Kotz, D., C. A. Gunter, S. Kumar, and J. P. Weiner. 2016. Privacy and security in mobile health: A research agenda. *Computer* 49(6):22-30.
41. Kaufman, D., J. Murphy, J. Scott, and K. Hudson. 2008. Subjects matter: A survey of public opinions about a large genetic cohort study. *Genetics in Medicine* 10(11):831-839.
42. US Department of Health and Human Services. 2016. *Health information privacy.* https://www.hhs.gov/hipaa/for-professionals/faq/2048/does-an-individual-have-a-right-under-hipaa/index.html (accessed October 23, 2017).
43. Adams, A., M. Shankar, and H. Tecco. 2017. *50 things we now know about digital health consumers.* https://rockhealth.com/reports/digital-health-consumer-adoption-2016 (accessed October 25, 2017).
44. Gottlieb, S. 2017. *FDA announces new steps to empower consumers and advance digital healthcare.* https://blogs.fda.gov/fdavoice/index.php/2017/07/fda-announces-new-steps-to-empower-consumers-and-advance-digital-healthcare (accessed September 20, 2017).
45. Gagnon, L. 2014. Time to rein in the "wild west" of medical apps. **Canadian Medical Association Journal** 186(8):e247.
46. Cheung, C., M. J. Bietz, K. Patrick, and C. S. Bloss. 2016. Privacy attitudes among early adopters of emerging health technologies. *PLOS One* 11(11):e0166389.
47. 23andMe. 2017. *Transparency report.* https://www.23andme.com/transparency-report (accessed December 13, 2017).

48. Apple. 2017. *Report on government and private party requests for customer information.* https://images.apple.com/legal/privacy/transparency/requests-2017-H1-en.pdf (accessed December 13, 2017).
49. Ray, T. 2013. GenomeWeb feature: Will the future of genomics be consumer driven? *GenomeWeb News,* https://www.genomeweb.com/clinical-genomics/genomeweb-feature-will-future-genomics-be-consumer-drivengenomeweb-feature-will (accessed October 18, 2016).

Suggested Citation

Tung, J. Y., R. J. Shaw, J. M. Hagenkord, M. Hackmann, M. Muller, S. H. Beachy, V. M. Pratt, S. F. Terry, A. K. Cashion, and G. S. Ginsburg. 2018. Accelerating precision health by applying the lessons learned from direct-to-consumer genomics to digital health technologies. *NAM Perspectives.* Discussion Paper, National Academy of Medicine, Washington, DC. https://nam.edu/accelerating-precision-health-by-applying-the-lessons-learned-from-direct-to-consumer-genomics-to-digital-health-technologies.

Author Information

The authors are participants in the Roundtable on Genomics and Precision Health at the National Academies of Sciences, Engineering, and Medicine. **Joyce Y. Tung, PhD,** is vice president, research, 23andMe. **Ryan J. Shaw, PhD, RN,** is associate professor, Duke University School of Nursing. **Jill M. Hagenkord, MD, FACP,** is chief medical officer, Color Genomics. **Meredith Hackmann** is research associate, National Academies of Sciences, Engineering, and Medicine. **Marion Muller** is a student at Emory University. **Sarah H. Beachy, PhD,** is senior program officer, National Academies of Sciences, Engineering, and Medicine. **Victoria M. Pratt, PhD**, is director, pharmacogenomics, Department of Medical and Molecular Genetics, Indiana University. **Sharon F. Terry, MA,** is president and chief executive officer, Genetic Alliance. **Ann K. Cashion, PhD,** is scientific director, National Institute of Nursing Research, National Institutes of Health. **Geoffrey S. Ginsburg, MD, PhD,** is director, Duke Center for Applied Genomics and Precision Medicine, Duke University School of Medicine.

Conflict-of-Interest Disclosures

Tung is employed by and owns equity in 23andMe. No other authors reported disclosures.

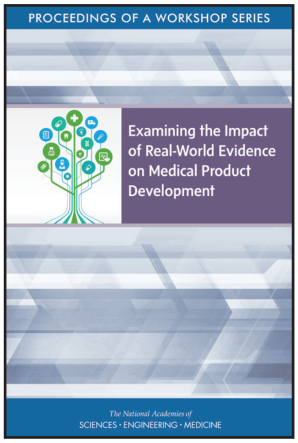
Disclaimer

The views expressed in this paper are those of the authors and not necessarily of the authors' organizations, the National Academy of Medicine (NAM), or the National Academies of Sciences, Engineering, and Medicine (the National Academies). The paper is intended to help inform and stimulate discussion. It is not a report of the NAM or the National Academies.

Workshop Series
HIGHLIGHTS

February 2019

Examining the Impact of Real-World Evidence on Medical Product Development


PROCEEDINGS OF A WORKSHOP SERIES

Examining the Impact of Real-World Evidence on Medical Product Development

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The Forum on Drug Discovery, Development, and Translation of the National Academies of Sciences, Engineering, and Medicine hosted a three-part workshop series in Washington, DC, Examining the Impact of Real-World Evidence (RWE) on Medical Product Development. The series, which was sponsored by the Food and Drug Administration (FDA), was designed to examine how RWE development and uptake could enhance medical product development and evaluation. Workshop participants discussed the current system of evidence generation and its limitations, shared lessons learned from successful initiatives that have incorporated RWE, and explored under what conditions RWE may be appropriate for informing medical product decision-making.

- Workshop 1 (September 2017) focused on how to align incentives to support collection and use of RWE in health product review, payment, and delivery;
- Workshop 2 (March 2018) covered what types of real-world data (RWD) and RWE might be appropriate for specific purposes;
- Workshop 3 (July 2018) examined approaches for operationalizing the collection and use of RWE.

FDA Commissioner Scott Gottlieb spoke at workshop 1, acknowledging that while RWE may not replace data from traditional clinical trials in many cases, FDA is working to develop policies to achieve more appropriate adoption of RWE to support regulatory decision-making, including new indications for approved drugs. He emphasized the importance of expanding the use of RWE in ways that could make medical product development more efficient and cost effective.

STAKEHOLDER PERSPECTIVES ON RWE

Several workshop 1 participants, including representatives of payers, health care delivery systems, and patients, presented perspectives on incentives for using RWE. Michael Sherman, Harvard Pilgrim Health Care, highlighted that payers must find a balance between access and affordability while driving innovation. He suggested that in cases for which a product approval may be based on limited evidence, FDA could consider requiring manufacturers to enter into value-based agreements that tie reimbursement to performance and encouraged post-marketing collaborations between payers and pharmaceutical companies.

Michael Horberg, Kaiser Permanente (KP) Mid-Atlantic Permanente Medical Group, and Daniel Ford, Johns Hopkins Health System, described delivery system perspectives. They noted that delivery systems value medical practices that are supported by quality, relevant evidence that demonstrates value to patients and discussed typical evidence generation processes.

Sharon Terry, Genetic Alliance, explained that patient-generated data and community-led registries can be an important source of evidence generation because they focus on patient priorities and lived experiences. These data sources still require rigorous validation, she said, but they should be integrated into clinical decision-making.

LESSONS LEARNED FROM RWE INITIATIVES

In workshop 1, Martin Gibson and Marie Kane, Northwest EHealth, presented on the Salford Lung Studies, which are two late-phase randomized controlled trials (RCTs)—one for asthma and another for chronic obstructive pulmonary disease. These studies were the first to evaluate the effectiveness of a pre-license medication in a real-world setting. Gibson and Kane described broad stakeholder engagement as a reason for the success of the studies and the challenges of developing a suitable data platform.

Richard Platt, Harvard Medical School, described Sentinel, an FDA monitoring system that uses electronic health data to support postmarketing medical product evaluation. He said the distributed system allows external data partners to retain private data prior to curation and can be used on its own or linked to other data sources, such as electronic health records (EHR) or patient-reported data.

Rachael Fleurence, National Evaluation System for Health Technology Coordinating Center, described the use of RWD and RWE for devices. She said that both are crucial for identifying problems with devices early in their use, and reliable RWE could shift device approval timelines and improve surveillance. Fleurence highlighted that registries are used widely for devices, and increased use of RWE could link existing registries to other data sources through Coordinated Registries Networks.

BARRIERS TO IMPLEMENTATION

Brian Bradbury and Elliot Levy, Amgen Inc., described barriers to RWD and RWE implementation, including a lack of: knowledge and awareness about RWE methods; capacity and expertise in relevant areas of research; and systems and processes to support RWE collection and use. Ford and John Doyle, IQVIA, identified RCT–RWE hybrid studies, such as pragmatic trials and cluster randomized designs, as possible approaches that combine advantages of both types of studies. Hui Cao, Novartis, suggested that evidence hierarchies that currently exist in medical product research could be revisited.

Marcus Wilson, HealthCore, described defragmentation as a process to integrate data sources from distinct stakeholders to provide a more complete understanding of a medical product. The process still requires data security and protection of patient privacy and business interests, he said. Anna McCollister-Slipp, Scripps Translational Science Institute, highlighted the lack of urgency around RWE adoption as problematic, as well as the hesitancy to include nontraditional stakeholders in research.

Addressing current evidence generation practices, Robert Califf, Verily Life Sciences, said the system should move past precision to focus on reliability. Potential steps to meet this goal could include the creation of a learning health care system, the use of quality by design, the use of automation, and operating from basic principles of scientific research, he said. Reflecting on the use of observational data networks, Patrick Ryan, Janssen, said analyses that incorporate the entire breadth of data on a particular set of medical products, including those that are not statistically significant, could be used to reflect a fuller understanding of those medical products.

Rory Collins, University of Oxford, focused on methods to improve RCTs, rather than replacing them with observational studies. He said RCTs are good at discovering moderate treatment effects and, while costly, innovative RCT designs that do not create data verification burdens could be useful. Janet Woodcock, FDA, ended the session by stating that the current evidence generation system needs improvement, and said opportunities to test product effectiveness using RWE could arise. She mentioned master protocols as a particular platform of interest.

PRACTICAL APPROACHES AND APPLICATIONS

While workshop 1 explored broad issues concerning barriers and incentives for the use of RWE, workshops 2 and 3 focused on specific questions stakeholders might consider before incorporating RWD and RWE into a study design. Questions raised at the first two workshops were incorporated into draft “decision aids” used to prompt further discussion during workshop 3 (to access the decision aids as well as additional details and resources, please see the Proceedings). The “decision aid” topics included: (1) when a particular real-world data element may be fit to assess study eligibility, treatment exposure, or outcomes; (2) some considerations for controlling or restricting treatment quality in real-world trials; (3) some considerations for obscuring intervention allocation in trials to generate RWE; and (4) potential ways to assess and minimize bias in observational comparisons.

WHEN CAN DECISION MAKERS RELY ON RWD?

At workshop 2, Adrian Hernandez, Duke University School of Medicine, presented on a suite of trials that compared novel oral anticoagulants (NOACs) to warfarin, all of which utilized RWD and consistently showed that NOACs were noninferior to warfarin. He posed a question for consideration: What questions characterize the use of a RWD source and signal reliability before a study is performed? At workshop 3, Jeff Allen, Friends of Cancer Research, presented a pilot project that investigated the performance of real-world endpoints among patients with advanced non-small cell lung cancer treated with immune checkpoint inhibitors. The project demonstrated that several real-world endpoints correlate well with overall survival, and showed that overall survival rates assessed from EHR and claims data were consistent with rates observed in clinical trials.

Aylin Altan, OptumLabs, and Brande Yaist, Eli Lilly and Company, said the usefulness of an RWD source for a particular question depends on whether it has information about the correct population, exposures, and outcomes, and Platt, Yaist, and Robert Temple, FDA, pointed out that it may be acceptable for RWD to be of different quality for different purposes. Cao said accuracy of RWD varies predictably, depending on factors such as treatment administration methods or the outcomes being measured.

Hernandez and Gregory Simon, KP Washington Health Research Institute, said provider-collected RWD is affected by the experience of the provider and the incentives they face. Luca Foschini, Evidation Health, spoke about patient-generated health data, noting that—while it has the potential to answer difficult research questions, facilitate broader participation in health research, and incorporate new data sources—it is subject to different biases than data collected within the health care system.

Other workshop participants discussed issues with the analysis of RWD. Marc Berger, formerly of Pfizer, Grazyna Lieberman, Genentech, and Deven McGraw, Ciitizen, explained that data sharing and transparency in data curation and analysis could be improved to encourage broader use of reliable RWE. Many speakers—including Altan, Berger, Foschini, Simon, and Yaist—pointed out that RWD can be affected by systemic and random bias, and are unique from other data sources because of their dynamic nature. Researchers can compensate for these, they said, but should remain mindful of potential biases when using RWD.

WHEN CAN DECISION MAKERS RELY ON REAL-WORLD TREATMENT?

At workshop 2, Ira Katz, Department of Veterans Affairs (VA), presented on a VA RCT that tested lithium as a treatment for suicide prevention. Katz described key questions that emerged through the study design process and emphasized the difficulty of making the trial generalizable to patients in real-world settings. At workshop 3, Larry Alphs, Newron Pharmaceuticals, presented on two real-world mental health trials (PRIDE and INTERCEPT) that grappled with issues around patient restriction to answer questions about safety and efficacy.

Horberg, Katz, Califf, and Alex London of Carnegie Mellon University discussed inclusion and exclusion criteria in real-world treatment settings. They argued for broadening these criteria in real-world trials to include patients with comorbidities or concomitant treatments to make the results more generalizable.

Alphs and Peter Stein, FDA, discussed a potential approach to choosing real-world trial restrictions, explaining that researchers could consider a specific set of categories that answer the research question while still honoring participant safety and autonomy. W. Benjamin Nowell, Global Healthy Living Foundation, also expressed concern about the role of patients in real-world research, emphasizing that research driven by patients is iterative and considers patients' needs, priorities, and experiences. The purpose of patient-centered research, he said, is to enable patients to make informed decisions about their own health care.

Alphs, Katz, and Simon described the role of researchers in real-world trials: Maintaining the trial protocol and caring for the well-being of patients, with patient safety coming first should the two conflict. Califf, Hernandez, and Stein discussed the importance, and ethical obligation, of setting a standard of care for the control arm of a study when designing real-world trials despite variance in standards across regions and treatment settings.

WHEN CAN DECISION MAKERS LEARN FROM REAL-WORLD TREATMENT ASSIGNMENT?

At workshop 3, Orly Vardeny, University of Minnesota and Minneapolis VA Center for Chronic Disease Outcomes Research, presented on the INVESTED trial, which explored the connection between influenza vaccine and cardiovascular events. The research team hypothesized that a stronger immune response from the high dose flu vaccine would translate into better cardiovascular outcomes, she said; they conducted a double-blinded RCT to prevent systematic biases inevitable in dispensing standard versus high dose vaccines.

Jonathan Watanabe, University of California, San Diego, and London said that blinding allows researchers to study the effects of an intervention without influence from patients or providers, but it may not always be appropriate or feasible. Cathy Critchlow of Amgen, Nancy Dreyer of IQVIA, and James Smith of FDA noted that the appropriateness of blinding is dependent on a study's context and uncertainties. These uncertainties, said London, can be classified along two axes: ensemble efficacy and utilization factors. The interaction of these two categories can indicate the appropriateness of blinding.

Dreyer, Rob Reynolds of Pfizer, and Smith explained that decisions on blinding can also be influenced by practical considerations, such as study cost, feasibility of masking treatment delivery, patient preferences, and data generalizability. Critchlow, Dreyer, John Graham of GlaxoSmithKline, and Smith said patient and provider bias can be difficult to predict, and it may not affect all outcomes, such as quantitative lab readings or all-cause mortality. However, it can affect subjective outcomes or have other effects such as in ascertainment or treatment bias, they said.

GAINING CONFIDENCE IN OBSERVATIONAL COMPARISIONS

At workshop 2, Sebastian Schneeweiss, Harvard Medical School, presented on the use of health care databases for regulatory decision-making. He explained that confidence in database studies is related to the type of effect being detected, and said such studies may be more appropriate when the outcomes and exposures are measurable in the data, when two active treatments are compared, and when the key confounding variables are measurable.

At workshop 3, Hector Izurieta, FDA, described a real-world study using Medicare Part D beneficiary data on the effectiveness and duration of effectiveness of the shingles vaccine, Zostavax. Izurieta explained how the investigators achieved balance between the treatment cohorts using propensity score matching and Mahalanobis metric matching, and conducted a secondary analysis to account for unmeasured confounders.

During discussion, David Madigan, Columbia University, noted that in disease areas for which RCTs are impractical, evidence from observational studies could be particularly valuable. Several participants discussed methods for observational data analysis. Madigan and Schneeweiss said transparent reporting of study methods can promote replicability and aid in assessing study validity. Speaking to a project currently under way, Jessica Franklin, Harvard Medical School, said replication of RCT results using observational databases can help establish criteria for conducting such studies more widely. Looking toward the future of observational studies, Javier Jiminez, Sanofi, and Mark van der Laan, University of California, Berkeley, said new methods such as predictive analytics and machine learning can potentially be used to predict outcomes for individual patients or identify associations.

Nicole Gormley and Heng Li, FDA, spoke from a regulatory perspective. Gormley described FDA's regulatory criteria for evaluating observational evidence: the data's relevance for a product's proposed indication; well-assessed outcomes; methods used to minimize bias; and rigorous statistical analysis.

REGULATORY PERSPECTIVES AND FUTURE OPPORTUNITIES

At workshop 3, Pall Jonsson, National Center for Health and Care Excellence (UK), described the Innovative Medicines Initiative GetReal project, explaining that health technology assessment relies on understanding the comparative effectiveness of new treatments. He noted that RWE can play a role in supplementing evidence from RCTs.

Komathi Stem, monARC Bionetworks, noted that using RWE can potentially engage patients more deeply in their care and in research, particularly with increases in usage of mobile technology and patients' ability to aggregate and store data about their own health. She explained that supporting a patient-centric shift in health research and care may require rethinking legislation, incentives, and partnerships. Levy described how new methods—such as adaptive designs, platform trials, or greater incorporation of RWE—have the potential to significantly reduce cost and time investments required for medical product development.

Concluding the workshop series, a panel of FDA leaders reacted to the workshop discussions. Jacqueline Corrigan-Curay, FDA Center for Drug Evaluation and Research (CDER), said CDER routinely uses RWE to support postmarketing safety evaluation and, to a limited extent, to evaluate effectiveness in certain rare diseases (including oncology). She emphasized that CDER's experience with Sentinel and other demonstration projects can inform policies going forward. Steve Anderson, FDA Center for Biologics Evaluation and Research (CBER), noted that CBER uses population-based data systems to conduct RWE safety and effectiveness studies, including the Biologics Effectiveness and Safety Sentinel Initiative to expand CBER's capabilities by providing data infrastructure, tools, and expertise.

Last, Jeffrey Shuren, FDA Center for Devices and Radiological Health (CDRH), said CDRH uses RWE in its product evaluations in pre- and postmarket decisions; it has started two programs combining registry data with other RWD to address regulatory needs. CDRH's 2017 RWE guidance, Shuren said, highlighted relevance and reliability as two critical considerations in evaluating RWE. All three FDA representatives said their Centers are interested in continuing to use RWE, but acknowledged that evidence used for regulatory purposes is necessarily different.

Mark McClellan, Duke-Margolis Center for Health Policy, touched on the idea of fit-for-purpose RWE in an environment with more readily available tools. He noted that clarity and specificity about when RWE is appropriate—and which data sources and methods are appropriate to address different types of questions—is the key to developing a framework for generating relevant evidence. Simon explained that, ultimately, delivering better health care to patients is the goal of using RWE.

DISCLAIMER: This Workshop Highlights was prepared by Erin Hammers Forstag, Benjamin Kahn, Amanda Wagner Gee, and Carolyn Shore as a factual summary of what occurred at the workshop. The statements made are those of the rapporteur or individual workshop participants and do not necessarily represent the views of all workshop participants; the planning committee; or the National Academies of Sciences, Engineering, and Medicine.

SPONSORS: These workshops were supported by AbbVie, Inc.; American Diabetes Association; Amgen Inc.; Association of American Medical Colleges; AstraZeneca; Burroughs Wellcome Fund; Critical Path Institute; Eli Lilly & Co.; FasterCures; Foundation for the National Institutes of Health; Friends of Cancer Research; GlaxoSmithKline; Johnson & Johnson; Merck & Co., Inc.; National Institutes of Health: National Cancer Institute, National Center for Advancing Translational Sciences, National Institute of Allergy and Infectious Diseases, National Institute of Mental Health, National Institute of Neurological Disorders and Stroke, Office of the Director; *New England Journal of Medicine;* Pfizer Inc.; Sanofi; Takeda Pharmaceuticals; and U.S. Food and Drug Administration: Center for Drug Evaluation and Research, Office of the Commissioner.

Examining the impact of real-world evidence on medical product development: Proceedings of a workshop series can be purchased or downloaded from the National Academies Press, 500 Fifth Street, NW, Washington, DC 20001; (800) 624-6242; www.nap.edu.

For more information, visit www.nationalacademies.org/RWEseries.

Health and Medicine Division

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

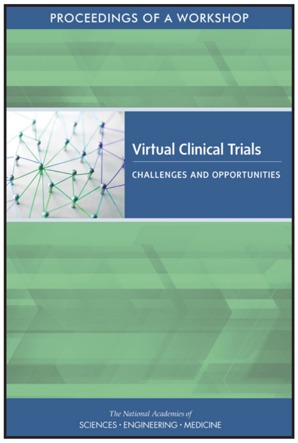
The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.
www.national-academies.org

Workshop
HIGHLIGHTS

September 2019

Virtual Clinical Trials: Challenges and Opportunities

The Forum on Drug Discovery, Development, and Translation of the National Academies of Sciences, Engineering, and Medicine hosted a workshop in Washington, DC, Virtual Clinical Trials: Challenges and Opportunities, held November 28–29, 2018. This workshop examined the current clinical trials infrastructure and potential opportunities for supporting the practical implementation of virtual clinical trials. Workshop participants discussed inefficiencies of the current clinical trial enterprise; the boundaries of what might be considered a virtual clinical trial; the opportunities to expand access for patients; perspectives of using digital health technologies in clinical care and observational and interventional studies; the impact of using digital health technologies on access and equity to clinical trials; the policy landscape governing clinical trials; and possible opportunities for future action.



PROCEEDINGS OF A WORKSHOP

Virtual Clinical Trials

CHALLENGES AND OPPORTUNITIES

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

DEFINING ‘VIRTUAL CLINICAL TRIALS’

Workshop participants used a variety of terms to refer to clinical trials in which all or part of the study incorporates digital health technologies and enables remote participation. Clay Johnston, Dell Medical School—The University of Texas at Austin, observed that an adequate umbrella term is not easy to identify. He commented that terms such as “decentralized,” “remote,” or “site–agnostic” may only describe some types of trials that incorporate digital health technologies. Ray Dorsey, University of Rochester, added that many study activities still require a centralized location. Additional terms referred to by workshop participants include “direct-to-participant,” “location–variable,” and “mobile.” Each of these terms highlights different aspects of how digital health technologies may be incorporated into study design.

OPPORTUNITIES TO IMPROVE CLINICAL TRIALS

According to Dorsey, clinical trials, as currently conducted, are expensive, inefficient, and inaccessible. Furthermore, he added that they fail to represent the patient population adequately. He envisions that in the near term, clinical trials could be conducted centrally (at one trial site), at multiple individual trial sites, and/or remotely (via digital health technologies), depending on the type of data needed. Dorsey provided examples that illustrate how virtual clinical trials can increase participant access and geographic representation, improve participant experience, and enhance recruitment of patient subpopulations.

Donna Cryer, Global Liver Institute, stated that a quality clinical trial is one that generates the minimal amount of credible, replicable, and evaluable data needed to answer meaningful questions with the least time and cost burden on participants. She expressed hope that as more virtual clinical trials are conducted and patient communities are better engaged, the quality of endpoints and outcome measurements will improve such that trials can more effectively address questions about a patient’s quality of life. Cryer emphasized that these types of trials offer opportunities to foster ongoing relationships with participants, better understand clinical conditions longitudinally, and generate new and relevant research questions.

Craig Lipset, Pfizer, Inc., described a 2011 study conducted by Pfizer, *REMOTE*, which was designed to validate available virtual technologies by repeating a standard brick-and-mortar clinical trial that Pfizer had conducted for Detrol, a drug used to treat overactive bladder. *REMOTE* was discontinued due to recruitment issues, but it successfully demonstrated the ability of researchers to remotely screen patients, seek informed consent, monitor safety, and capture required data to demonstrate safety and efficacy. According to Lipset, the study did not operate at the limit of available technology at the time it was conducted, nor did it require new legislation, safe harbor, or guidance from regulators. Lipset emphasized that moving beyond such pilot programs may require a change in will and culture by stakeholders across the clinical trial enterprise.

LESSONS LEARNED FROM VIRTUAL CLINICAL TRIALS

Jenny Bollyky, Livongo, described three challenges associated with remote management of chronic conditions: (1) attributing digitally collected data to the individual under observation, (2) providing meaningful data for the participant being monitored (e.g., instant feedback on blood glucose levels for people living with diabetes), and (3) coordinating care for people with chronic conditions. Leveraging real-world evidence in a research setting may be complicated by the Hawthorne effect—the alteration of behavior by the subjects of a study due to their awareness of being observed, noted Bollyky.

Joshua Denny, Vanderbilt University Medical Center, discussed how longitudinal research on large cohorts that aggregates data from a variety of sources (e.g. electronic health records, participant provided data, collected biospecimens, claims data), can facilitate trials. He focused on three examples: All of Us,[1] U.K. Biobank,[2] and Project Baseline.[3] The combination of data that emerge from cohort studies, noted Denny, could be leveraged to conduct a wide range of clinical research activities, including the identification of new disease targets and pharmacogenomic discovery, testing for adverse drug events, and identifying disease subtypes. Furthermore, Denny highlighted that large, diverse observational cohort studies can enable direct and targeted recruitment of diverse populations, in addition to more intelligent trial design.

Steven Cummings, San Francisco Coordinating Center, discussed what he termed “direct-to-participant” (D2P) trials, which he defined as having no physical clinical sites, and thus no geographic limits on recruitment. He argued that the term ‘D2P trials’ more aptly captures the importance of building relationships with participants and that this approach can simplify trial design (i.e., improve enrollment, screening, data collection and reporting, etc.), thus increasing the likelihood of trial success. He added that recruitment from trusted communities and providers could lend D2P trials more success than solely web-based recruitment.

Kimberly Hawkins, Sanofi Genzyme, highlighted opportunities that decentralized trials provide, including increased geographic flexibility and reduced burden for participants; continuous data collection for faster and more accurate detection of health signals; and improved long-term follow-up with participants. Hawkins also listed challenges that decentralized trials can pose, such as operational (e.g., integration of emerging types of data and maintaining a temperature-controlled supply chain), regulatory (e.g., endpoint and digital health technology validation), and change management (e.g., integration of decentralized trials into medical product development).

Wendy Weber, National Center for Complementary and Integrative Health, discussed the National Institutes of Health’s Health Care Systems Research Collaboratory, and its role in supporting embedded clinical trials, or research that engages health care delivery systems as partners. Such trials can leverage electronic health records (EHRs), which can provide a cost-effective source of information collected during routine care. However, as Weber mentioned, EHRs may not always reflect the schedule of data collection necessary for a clinical trial nor include the type of data needed. As such, embedded pragmatic trials may work best when the outcome of interest is captured in the EHR and when the trial itself does not place undue burden on patients and clinicians during routine health care visits, noted Weber.

Noah Craft, Science 37, emphasized that successful D2P trial design should fit participant needs. Adrian Hernandez, Duke University School of Medicine, added that the field needs to develop the science of patient engagement to better understand patient preferences and participation in clinical research. Jon White, Office of The National Coordinator for Health Information Technology (ONC), stated that ONC and the Centers for Medicare & Medicaid Services have issued proposed rules for how the federal government will regulate health information systems to make EHR data more readily available through application program interfaces. Josh Rose, IQVIA, said that flexibility is key when interacting with participants—some clinical trial participants want more in-person interaction while others prefer to use just the technology.

ACCESS AND EQUITY

Silas Buchanan, Institute of eHealth Equity, emphasized the importance of engaging directly with community members when deploying digital interventions. Building a network of partnerships and leveraging trust brokers within the community can be instrumental in the success of public health campaigns. Using his social impact firm, Institute for eHealth Equity, as an example, Buchanan provided key lessons learned for how virtual clinical trials can be designed and positioned to increase inclusion of underrepresented populations and how to address the unique socioeconomic factors those populations face, if specific trial design considerations call for it. A strong ethos of community engagement is key to the introduction of digital health tools, emphasized Buchanan, as is acknowledgement of histories of discrimination and transparent discussion about power and responsibilities.

Sherine El-Toukhy, National Institute on Minority Health and Health Disparities, discussed the inclusion of minority populations in research and commented on barriers to participation, such as skepticism about the researchers and value of the study. She suggested that while health information technology may reduce health inequities, it can unintentionally exacerbate existing disparities or create new ones. She emphasized that understanding and prioritizing the target population’s needs, values, and preferences is critical for designing culturally and linguistically appropriate clinical trial recruitment material.

Will McIntyre, The Michael J. Fox Foundation, pointed out that trial participation often reflects the urban–rural divide. He then discussed how synergy between the research and technology sectors could equip study participants with

---

[1] Available at https://www.joinallofus.org/en/about (accessed April 16, 2019).

[2] Available at https://www.ukbiobank.ac.uk (accessed April 16, 2019).

[3] Available at https://www.projectbaseline.com accessed April 16, 2019).

technologies that will better enable those living in rural areas to connect with studies.

Sally Okun, PatientsLikeMe, commented on how virtual clinical trials can use unique designs to create new insights and increase participation rates of those who are typically excluded from clinical trials. ALSUntangled, a research consortium, seeks to understand the efficacy of alternative and off-label treatments for amyotrophic lateral sclerosis (ALS) and engages in patient—driven inquiry. In one its studies, ALSUntangled expanded its inclusion criteria such that those with more advanced forms of ALS could participate.

Participant engagement could be improved by returning individual research results to participants—a trend that reflects a broader cultural shift from paternalism to partnership in medicine and research, proposed Kathy Hudson, People-Centered Research Foundation. However, doing so may need to be balanced with the risk of damaging the integrity of the study, added Okun and El-Toukhy. Buchanan highlighted the importance of creating a trustworthy mechanism for returning data to participants. Hudson added that for widespread culture in the pharmaceutical industry to change, there needs to be evidence that involving participants improves outcomes and a continued expectation from the U.S. Food and Drug Administration (FDA) that clinical trials focus on patient partnership.

POLICY CONSIDERATIONS

Leonard Sacks, FDA, highlighted the opportunities to use mobile technologies and engage local providers to promote inclusivity and convenience for trial participants, and for gathering information on real-world patient experience. These opportunities will require policies and regulations to address patient safety, privacy, the integrity of the data produced by remote technologies, and the responsibilities of the investigators involved in technology–enabled decentralized trials. Ensuring participant safety in a decentralized trial is no different than in a traditional clinical trial, said Sacks. However, he noted that digital health technologies are creating opportunities for greater safety oversight by replacing episodic monitoring with continuous monitoring of variables, such as blood glucose levels, heart rate, and rhythm. At the same time, he cautioned, it is important to ensure that technology failure does not jeopardize participant safety or the integrity of the data and that technical support is available for when a digital health technology malfunctions.

Leanne Madre, Clinical Trials Transformation Initiative (CTTI), described lessons learned from CTTI's Decentralized Clinical Trials Project. CTTI has identified multiple benefits to decentralized trials, such as faster trial participant recruitment; improved retention; greater control, convenience, and comfort for participants; and increased participant diversity. To achieve these benefits, CTTI issued six categories of recommendations: (1) Engage early with FDA and those who have already conducted a decentralized trial can be important for developing the trial protocol and trial success; (2) Maintain licensed investigators in each active trial site or use investigators licensed in multiple states, given state-by-state variations in regulations governing physician licensure; (3) Use mobile health care providers to facilitate participant protocol contributions, such as blood draws or administration of the investigational product, as a decentralized trial can cover a wide geographic area; (4) Review laws governing D2P shipment of drugs as these laws can also vary from state to state; (5) Consider differences between a standard and decentralized trial when delegating responsibilities to investigators, sub-investigators, and local providers; (6) Ensure that trial participants and trial staff are aware of procedures related to adverse events.

Deven McGraw, Ciitizen Corporation, discussed the importance of protecting participant data and privacy, as well as policy mechanisms in the United States and Europe that govern these protections. Privacy, McGraw stated, is about enabling appropriate data use with good data stewardship that engenders trust among trial participants. She added that investigators must make and keep commitments to trial participants concerning how their data will be used and disclosed, and be transparent about how data will be used. In the context of a virtual or decentralized trial, the Health Insurance Portability and Accountability Act (HIPAA) covers identifiable data collected—even from a mobile health technology—if the investigator is a HIPAA-covered entity. However, it is less clear if HIPAA covers data that reside in consumer mobile devices. McGraw mentioned that recent developments in privacy law, such as the California Consumer Protection Act and the European Union's Global Data Protection Rule, are now requiring more explicit forms of consent for data reuse.

Matthew McIntyre, 23andMe, discussed policy and regulatory challenges in developing informed consent processes for remote studies that involve passive data collection—data collection in which information is automatically gathered, often without the awareness of the research participant. While privacy considerations for passive data collection primarily draw on HIPAA, participants may have concerns that go beyond the risk of de-identification, such as who will have access to their data and how their data will be used. This is true for paradata—additional data collected along with passive data, such as time stamps and geolocation. As the quantity of paradata collected makes de-identification a challenge, McIntyre said that one solution is data minimization. However, this comes at the expense of limiting quality control and oversight of protocol compliance. McIntyre indicated that the incorporation of passive data in trials may require new policies for mixed uses and sources of data, in addition to dynamic ways to acquire informed consent and inform participants about data collection

REFLECTIONS ON THE WORKSHOP AND POTENTIAL FUTURE DIRECTIONS

Concluding the workshop, participants reviewed key themes discussed during the workshop. Johnston emphasized the importance of using human-centered design and seeking input from participants early in the trial design process. Kelly Simcox, Sanofi, pointed out the need to have more visibility regarding current virtual clinical trials so that lessons learned are shared and D2P trials can move beyond the pilot-stage and towards mainstream clinical trial methodology. Rebecca Pentz, Emory University School of Medicine, and Hudson emphasized the importance of relaxing exclusion criteria to increase equitable participation in trials. They also highlighted the importance of community engagement, and how returning data to community members in real time can empower participants and help build trust in the trial process. John Wilbanks, Sage Bionetworks, noted the disconnect between the technology needed to support virtual clinical trials and the current policy environment. He emphasized the importance of understanding consent as an ongoing relationship—not a one-time transaction—especially when digital health technologies are passively collecting data.

Johnston emphasized the need for development of an ontology to describe the variety of studies occurring in the virtual clinical trial space. While regulatory issues were raised as a barrier to conducting virtual clinical trials, Hudson suggested that policy analysis and advocacy could facilitate implementation of reform. Craft went further to propose that the federal government could carve out an exemption to state laws for telemedicine-based clinical research. Johnston and Cummings mentioned that virtual clinical trials may require greater inclusion of participants in study design, governance, as well as drug safety monitoring boards. Resources that provide guidance for meaningful engagement and bidirectional training for a study team will also be important to engage participants meaningfully, noted Hudson.

Virtual clinical trials are not a "one-size-fits-all" model and only a fraction of clinical trials might be considered fully virtual. In the near term, Hawkins suggested that digital health technologies may only be applicable in a few settings, such as disease areas for which telemedicine is already an accepted practice or for evaluating medical products with a known safety profile and endpoints that can be measured remotely. However, in the longer term, Lipset and Dorsey highlighted that virtual clinical trials have the potential to streamline the process of drug development and may offer new opportunities for a modern, more patient-centric clinical trial enterprise.

DISCLAIMER: This Workshop Highlights was prepared by Eeshan Khandekar and Carolyn Shore as a factual summary of what occurred at the workshop. The statements made are those of the rapporteur or individual workshop participants and do not necessarily represent the views of all workshop participants; the planning committee; or the National Academies of Sciences, Engineering, and Medicine.

SPONSORS: This workshop was supported by Abbvie Inc.; American Society for Microbiology; Amgen Inc.; Association of American Medical Colleges; AstraZeneca; Burroughs Wellcome Fund; Critical Path Institute; Eli Lilly and Company; FasterCures; Foundation for the National Institutes of Health; Friends of Cancer Research; GlaxoSmithKline; Johnson & Johnson; Merck & Co.; National Institutes of Health: National Cancer Institute, National Center for Advancing Translational Sciences, National Institute of Allergy and Infectious Diseases, National Institute of Mental Health, National Institute of Neurological Disorders and Stroke, Office of Science Policy; *New England Journal of Medicine*; Pfizer Inc.; Sanofi; Takeda Pharmaceuticals; and U.S. Food and Drug Administration: Center for Drug Evaluation and Research, Office of the Commissioner.

Virtual Clinical Trials: Challenges and Opportunities: Proceedings of a Workshop can be purchased or downloaded from the National Academies Press, 500 Fifth Street, NW, Washington, DC 20001; (800) 624-6242; www.nap.edu.

For more information, visit www.nationalacademies.org/VirtualClinicalTrialsPW.

Health and Medicine Division

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The nation turns to the National Academies
of Sciences, Engineering, and Medicine for
independent, objective advice on issues that
affect people's lives worldwide.
www.national-academies.org

NATIONAL
CANCER POLICY
FORUM

Applying Big Data to Address the Social Determinants of Health in Oncology

Workshop Hosted by the National Cancer Policy Forum
In Collaboration with the Board on Mathematical Sciences and Analytics

October 28-29, 2019
NAS Lecture Room
2101 Constitution Ave., N.W.
Washington, D.C. 20418

Statement of Task

An ad hoc committee will plan and host a 1.5-day public workshop that will examine health equity and the social determinants of health in the context of cancer, and consider opportunities to effectively use big data and convergence science to improve health outcomes and reduce disparities. The workshop will feature invited presentations and panel discussions on topics that may include:

- The impact of social determinants of health on cancer incidence and outcomes
- Ways to capture precise and meaningful data on social determinants of health in oncology (social omics)
- Potential biomarkers of environmental and social factors that may affect cancer incidence and care outcomes (e.g., microbiome, exposome, and epigenome)
- The potential advantages and limitations of various methods for identifying geographic populations and communities at risk (e.g., state, county, zip code, city ward, or congressional district)
- Potential opportunities to reduce bias in capturing big data and applying computational methods to inform cancer care
- Examples of how community engagement and big data could be used to inform policies to promote health equity in oncology

The planning committee will develop the agenda for the workshop sessions, select and invite speakers and discussants, and moderate the discussions. A proceedings of the presentations and discussions at the workshop will be prepared by a designated rapporteur in accordance with institutional guidelines.

Planning Committee

Robert A. Winn, MD (Chair) Professor of Medicine Division of Pulmonary, Critical Care, Sleep, and Allergy University of Illinois at Chicago Associate Vice Chancellor, Community-Based Practice Director, University of Illinois Cancer Center University of Illinois Hospital & Health Sciences System	
Garnet Anderson, PhD Senior Vice President and Director Public Health Sciences Division Fred Hutchinson Cancer Research Center Affiliate Professor, Department of Biostatistics University of Washington	**Kenneth Anderson, MD** Kraft Family Professor of Medicine American Cancer Society Clinical Research Director Jerome Lipper Multiple Myeloma Center Harvard Medical School Dana-Farber Cancer Institute

Karen Basen-Engquist, PhD, MPH
Annie Laurie Howard Research
Distinguished Professor
Director, Center for Energy Balance in Cancer Prevention & Survivorship
Professor of Behavioral Sciences
The University of Texas MD Anderson Cancer Center

Otis Brawley, MD, MACP, FASCO, FACE
Bloomberg Distinguished Professor
Department of Epidemiology
Bloomberg School of Public Health
Department of Oncology
School of Medicine
Johns Hopkins University

Nicole F. Dowling, PhD
Associate Director for Science
Division of Cancer Prevention and Control
Centers for Disease Control and Prevention

Stanton L. Gerson, MD
Director, Case Comprehensive Cancer Center
Professor of Hematological Oncology
Case Western Reserve University
Director, University Hospitals Seidman Cancer Center

Lori Hoffman Högg, MS, RN, CNS, AOCN®
VHA National Program Manager, Prevention Policy
National Center for Health Promotion and Disease Prevention
National Oncology Clinical Advisor
Office of Nursing Services
Department of Veterans Affairs

Nicholas Horton, ScD
Beitzel Professor of Technology and Society
Amherst College

Ronald Kline, MD, FAAP
Medical Officer, Patient Care Models Group
Center for Medicare and Medicaid Innovation
Centers for Medicare & Medicaid Services

Rebecca Miksad, MD
Senior Medical Director
Flatiron Health

Timothy Rebbeck, PhD
Director, Center for Global Cancer Prevention
Vincent L. Gregory Professor of Cancer Prevention
Harvard T.H. Chan School of Public Health
Professor, Division of Population Sciences
Dana-Farber Cancer Institute

Victoria Seewaldt, MD
Professor and Chair, Department of Population Science
City of Hope
Associate Director, Cancer Control
City of Hope Comprehensive Cancer Center

George J. Weiner, MD
C.E. Block Chair of Cancer Research
Professor of Internal Medicine
Director, Holden Comprehensive Cancer Center
University of Iowa

Robin Yabroff, PhD
Senior Scientific Director
Health Services Research
American Cancer Society

Opportunities and Challenges for Using Digital Health Applications in Oncology

A National Cancer Policy Forum Workshop
In Collaboration with the Forum on Cyber Resilience

July 13-14, 2020
Keck 100
500 Fifth St., N.W.
Washington, D.C. 20001

Statement of Task

A planning committee of the National Academies of Sciences, Engineering, and Medicine will plan and host a 1.5-day public workshop that will examine opportunities and challenges, including validation, data security, and patient privacy issues, for the use of digital health applications in oncology. The workshop will feature invited presentations and panel discussions on topics such as:

- An overview of existing and emerging digital health applications and the potential benefits and risks associated with their use
- Strategies to validate digital health applications, regulate their use, and mitigate potential risks associated with their use
- Strategies for protecting the security of data collected using digital health applications
- Patient privacy considerations, especially given the potential for data linkage with data from other sources of personal information
- Best practices and principles for access to and consent for the use of patient data generated by digital health applications
- Ways to integrate patient-generated health data into EHRs and clinical workflow
- Lessons learned from other industries and/or countries that could inform digital health application development and use.

The planning committee will develop the agenda for the workshop sessions, select and invite speakers and discussants, and moderate the discussions. A proceedings of the presentations and discussions at the workshop will be prepared by a designated rapporteur in accordance with institutional guidelines.

Planning Committee

Lawrence N. Shulman, MD, FACP, FASCO (Chair) Deputy Director, Clinical Services Director, Center for Global Cancer Medicine Abramson Cancer Center Professor of Medicine University of Pennsylvania	
Karen Basen-Engquist, PhD, MPH Annie Laurie Howard Research Distinguished Professor Director, Center for Energy Balance in Cancer Prevention & Survivorship Professor of Behavioral Science The University of Texas MD Anderson Cancer Center	**Cathy Bradley, PhD** Associate Director, Cancer Prevention and Control University of Colorado Cancer Center

Deborah Estrin, PhD
Association Dean
Robert V. Tishman ’37 Professor
Cornell Tech

Mia Levy, MD, PhD
Director, Rush University Cancer Center
Associate Professor of Medicine
Division of Hematology and Oncology
System Vice President, Cancer Services
Rush System for Health

J. Leonard Lichtenfeld, MD, MACP
Deputy Chief Medical Officer
American Cancer Society

Bradley Malin, PhD
Vice Chair for Research Affairs in Biomedical Informatics
Professor of Biomedical Informatics, Biostatistics & Computer Science
Co-Director, Center for Genetic Privacy & Identity in Community Settings
Co-Director, Health Data Science Center
Director, Health Information Privacy Laboratory
Vanderbilt University

Deven McGraw, JD, MPH, LLM
Chief Regulatory Officer
Ciitizen

Neal J. Meropol, MD
Vice President
Research Oncology
Flatiron Health

Randall A. Oyer, MD
Medical Director, Oncology
Ann B. Barshinger Cancer Institute
Lancaster General Penn Medicine

Lisa Kennedy Sheldon, PhD, APRN, AOCNP®, FAAN
Chief Clinical Officer
Oncology Nursing Society

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Forum on Drug Discovery, Development, and Translation
Roundtable on Genomics and Precision Health

The Role of Digital Health Technologies in Drug Development

A Workshop

Other Activities of Interest

Genomics-Enabled Learning Health Care Systems: Gathering and Using Genomic Information to Improve Patient Care and Research (2014)

- Workshop Website
- Workshop Proceedings

Harnessing Mobile Technology to Predict, Diagnose, Monitor, and Develop Treatments for Nervous System Disorders (2018)

- Workshop Website
- Workshop Proceedings

Applying Big Data to Address the Social Determinants of Health in Oncology (2018)

- Workshop Website

Neuroscience Data in the Cloud (2019)

- Workshop Website
- Workshop Proceedings