

BIO 512

Digital Epidemiology

Digital Cohorts & Trials

Learning Objectives

- Understand what problems digital cohorts and trials are trying to solve, and how
- Understand what problems emerge when running cohorts and trials digitally, and attempts to solve them
- Be familiar with real-world examples of digital cohorts and trials

Digital Cohorts & Trials

A Recent Development

- Digital Public Health Surveillance can be exploratory and experimental.
- Cohorts and trials are too resource-intensive for that.

Digital Cohorts & Trials

Issues With Conventional Studies

- Clinical site - everybody has to travel there (time, transportation costs, etc.)
- Measurements on site, in “artificial” environment, at one time point.



Data collection & coordination

Digital Cohorts & Trials

Issues With Conventional Studies

- Many studies have participation rates of $< 10\%$
- Participation rates in long-term cohorts has dropped over time
- Problem with representation

Digital Cohorts & Trials

Strength of Digital Studies

- Measurements happen “in situ” (ecological measurements), continuously.
- Cost shifts to devices, coordination



Data collection

Data collection



Data collection



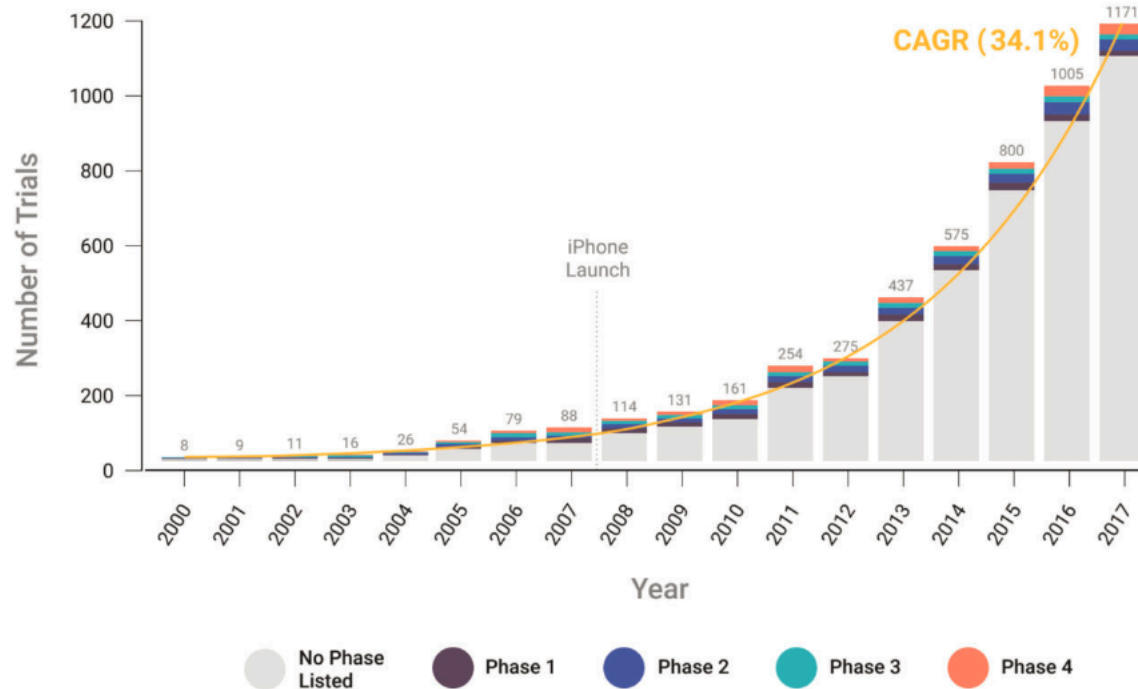
Data collection



Coordination

Digital Cohorts & Trials

Growth of Digital Studies



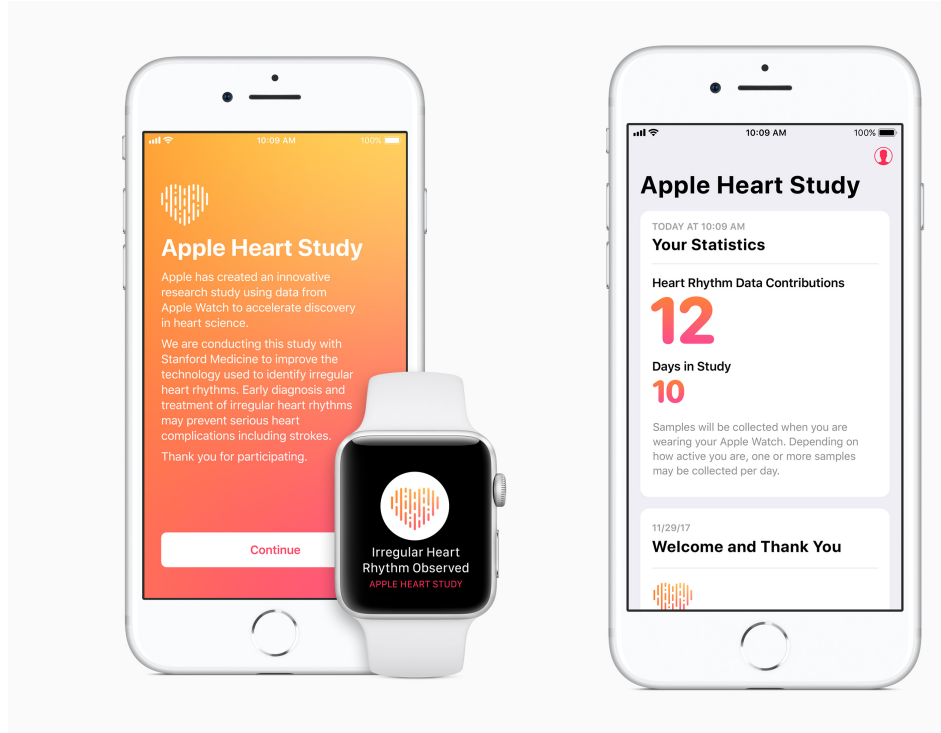
Digital Cohorts & Trials

Emerging Terminology

- As in every emerging area, there is a diversity of terms that refer to almost the same thing.
- Decentralized trials
- Siteless trials
- mHealth

Digital Cohorts & Trials

Digital Cohort Examples



- Apple Heart Study, launched in 2017
- Once the watch identifies possible aFib, participant gets an ECG
- ECG then interpreted by clinicians

Digital Cohorts & Trials

Digital Cohort Examples

THE NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation

Marco V. Perez, M.D., Kenneth W. Mahaffey, M.D., Haley Hedlin, Ph.D., John S. Rumsfeld, M.D., Ph.D., Ariadna Garcia, M.S., Todd Ferris, M.D., Vidhya Balasubramanian, M.S., Andrea M. Russo, M.D., Amol Rajmane, M.D., Lauren Cheung, M.D., Grace Hung, M.S., Justin Lee, M.P.H., Peter Kowey, M.D., Nisha Talati, M.B.A., Divya Nag, Santosh E. Gummidipundi, M.S., Alexis Beatty, M.D., M.A.S., Mellanie True Hills, B.S., Sumbul Desai, M.D., Christopher B. Granger, M.D., Manisha Desai, Ph.D., and Mintu P. Turakhia, M.D., M.A.S., for the Apple Heart Study Investigators*

ABSTRACT

BACKGROUND

Optical sensors on wearable devices can detect irregular pulses. The ability of a smartwatch application (app) to identify atrial fibrillation during typical use is unknown.

METHODS

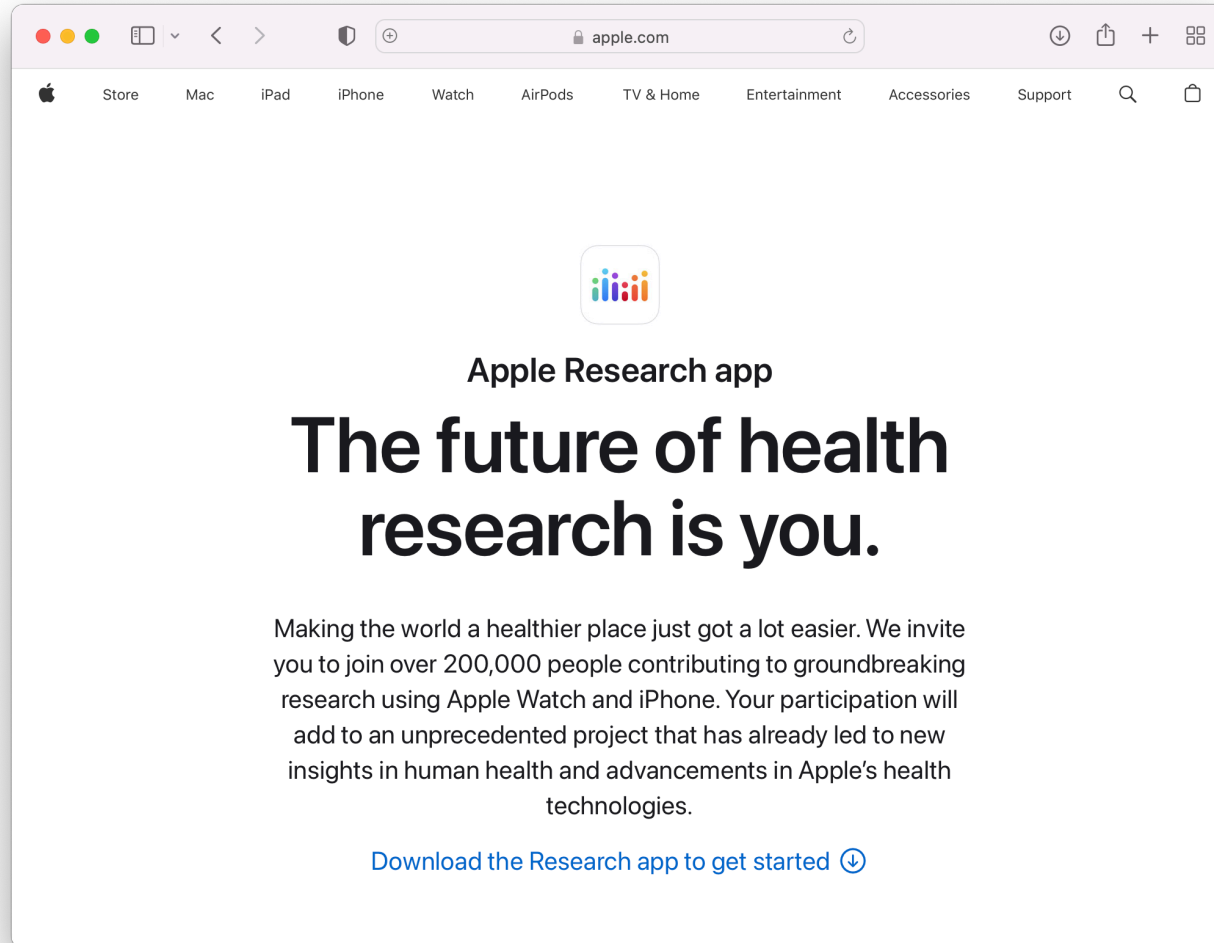
From the Division of Cardiovascular Medicine (M.V.P.), Stanford Center for Clinical Research (K.W.M., A.R., N.T.), the Quantitative Sciences Unit (H.H., A.C., V.B., L.S.E., M.D.), Information

Digital Cohorts & Trials

Digital Cohort Examples

RESULTS

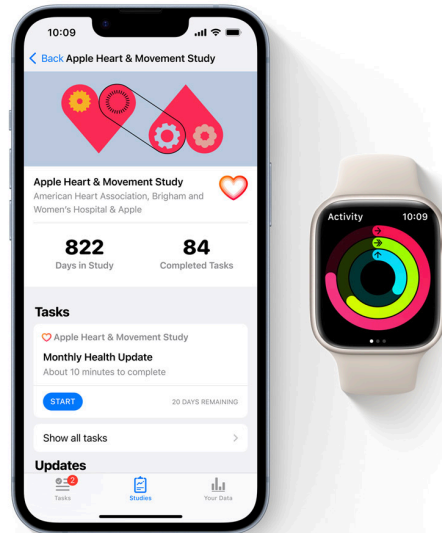
We recruited 419,297 participants over 8 months. Over a median of 117 days of monitoring, 2161 participants (0.52%) received notifications of irregular pulse. Among the 450 participants who returned ECG patches containing data that could be analyzed — which had been applied, on average, 13 days after notification — atrial fibrillation was present in 34% (97.5% confidence interval [CI], 29 to 39) overall and in 35% (97.5% CI, 27 to 43) of participants 65 years of age or older. Among participants who were notified of an irregular pulse, the positive predictive value was 0.84 (95% CI, 0.76 to 0.92) for observing atrial fibrillation on the ECG simultaneously with a subsequent irregular pulse notification and 0.71 (97.5% CI, 0.69 to 0.74) for observing atrial fibrillation on the ECG simultaneously with a subsequent irregular tachogram. Of 1376 notified participants who returned a 90-day survey, 57% contacted health care providers outside the study. There were no reports of serious app-related adverse events.

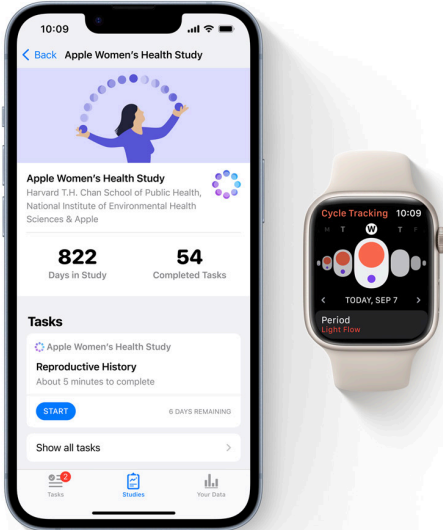


Get the app. Join a study.

Apple Heart and Movement Study

A collaboration between the American Heart Association and Brigham and Women's Hospital, a Harvard Medical School affiliate. This study uses Apple Watch and iPhone to provide insight into the links between physical activity and heart health over time.² Findings have already contributed to advancements in fall detection on Apple Watch and Walking Steadiness — a first-of-its-kind metric that provides insights into fall risk on iPhone. Using these two devices, researchers can gain a better understanding of potential early warning signs to create interventions and products that may help us all lead longer, healthier, and more active lives.

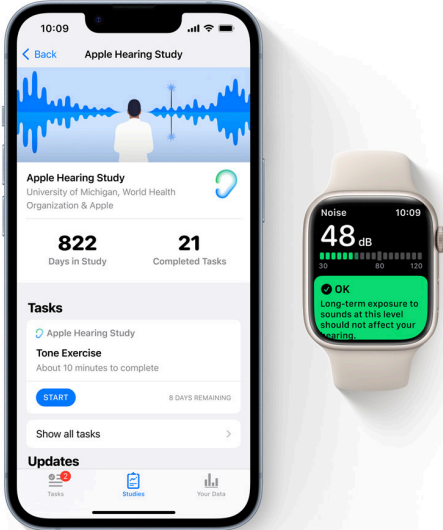




The image shows a promotional graphic for the Apple Women's Health Study. On the left, an iPhone displays the app's interface. At the top, it says 'Back Apple Women's Health Study' with a back arrow. Below is an illustration of a woman holding a globe. The text 'Apple Women's Health Study' is followed by 'Harvard T.H. Chan School of Public Health, National Institute of Environmental Health Sciences & Apple'. It shows '822 Days in Study' and '54 Completed Tasks'. Under 'Tasks', there is a section for 'Reproductive History' with a 'START' button and '6 DAYS REMAINING'. At the bottom, there are icons for 'Tasks', 'Studies', and 'Your Data'. To the right of the iPhone is an Apple Watch displaying the 'Cycle Tracking' app with a circular progress indicator and the text 'Cycle Tracking 10:09'. On the right side of the graphic, the title 'Apple Women's Health Study' is written in a large, bold font. Below the title, a paragraph explains the partnership with Harvard and the NIH to study factors affecting menstrual cycles and gynecologic conditions, mentioning the study's role in understanding reproductive behavior during the COVID-19 pandemic and improving predictions in the Cycle Tracking app. It concludes by stating that data from Apple Watch and iPhone, along with survey responses, will help inform the development of innovative products for menstrual cycles and risk assessment.

Apple Women's Health Study

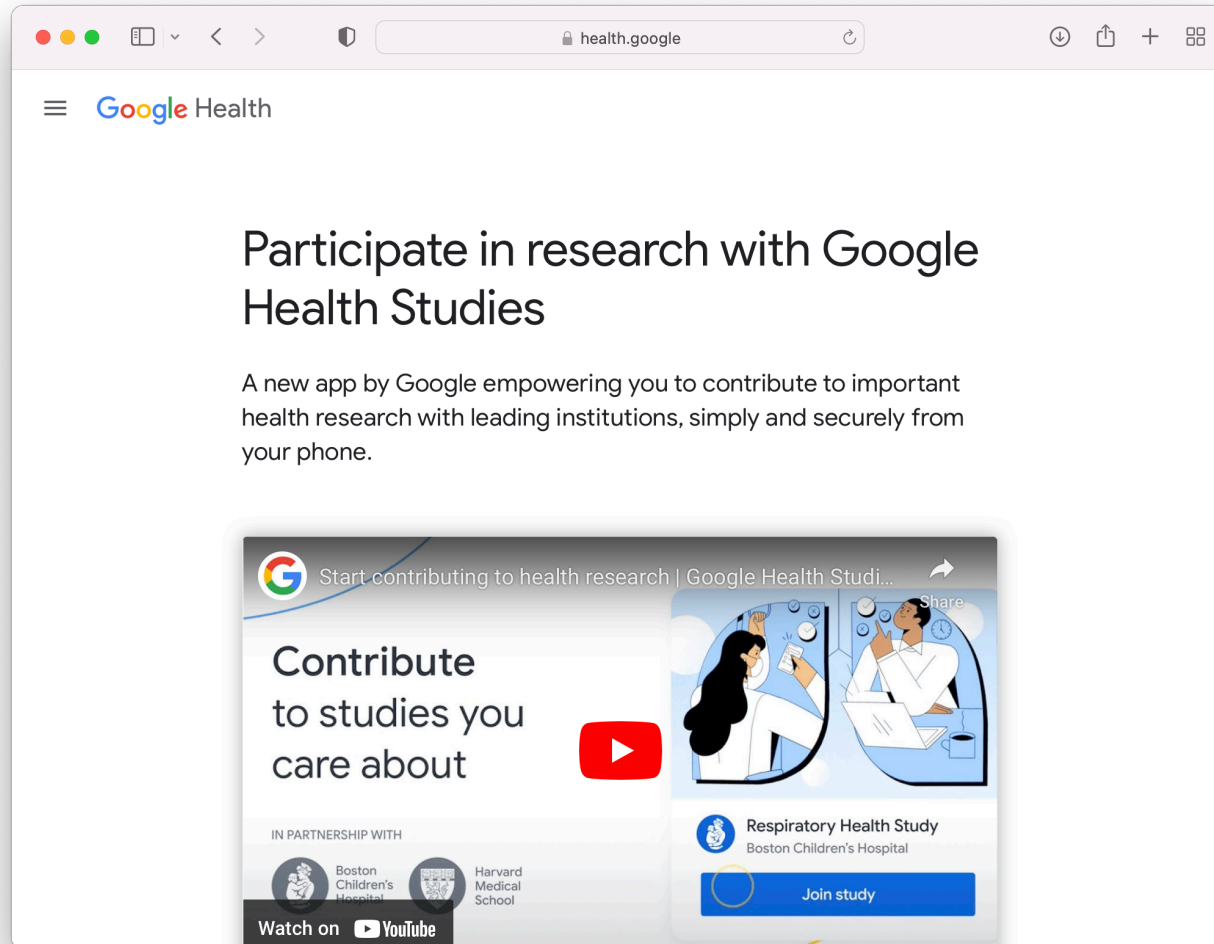
Apple is teaming up with the Harvard T.H. Chan School of Public Health and the National Institute of Environmental Health Sciences to gain a deeper understanding of how certain demographic and lifestyle factors affect menstrual cycles and gynecologic conditions including infertility, menopause, and PCOS. The study has already helped researchers better understand reproductive behavior during the COVID-19 pandemic, and improved period and fertile window predictions in the Cycle Tracking app. Data from Apple Watch and iPhone, along with participants' survey responses, will help inform the development of innovative products for menstrual cycles, as well as risk assessment and early screening of gynecologic conditions.



The image shows a screenshot of the Apple Hearing Study app interface. On the left, a smartphone displays the app's main screen. At the top, it says 'Apple Hearing Study' with a 'Back' button. Below this is a blue waveform graphic. The main section shows '822 Days in Study' and '21 Completed Tasks'. Under 'Tasks', there is a 'Tone Exercise' section with a 'START' button and '8 DAYS REMAINING'. At the bottom, there is an 'Updates' section. On the right, a smartwatch displays a 'Noise' level of '48 dB' with a green bar indicating the level. Below the noise level, it says 'OK Long-term exposure to sounds at this level should not affect your hearing.'

Apple Hearing Study

Conducted in conjunction with the University of Michigan, this groundbreaking study will advance the understanding of how hearing could be impacted over time by exposure to sound at certain levels. Researchers have already collected exposure data from 400 million hours of environmental sounds and supplemented this research with lifestyle surveys to analyze how sound exposure affects hearing, stress levels, and cardiovascular health. The study data will also be shared with the World Health Organization as a contribution to its Make Listening Safe initiative.



The screenshot shows a web browser window with the address bar displaying "health.google". The page header includes the Google Health logo. The main heading is "Participate in research with Google Health Studies". Below this, a paragraph states: "A new app by Google empowering you to contribute to important health research with leading institutions, simply and securely from your phone." The central content is a video player. The video title is "Start contributing to health research | Google Health Studi...". The video content includes the text "Contribute to studies you care about" and "IN PARTNERSHIP WITH" followed by logos for Boston Children's Hospital and Harvard Medical School. A "Watch on YouTube" button is at the bottom left of the video player. On the right side of the video player, there is a "Share" button and a "Join study" button for the "Respiratory Health Study" by Boston Children's Hospital.

health.google

Google Health

Participate in research with Google Health Studies

A new app by Google empowering you to contribute to important health research with leading institutions, simply and securely from your phone.

Start contributing to health research | Google Health Studi...

Contribute to studies you care about

IN PARTNERSHIP WITH

Boston Children's Hospital Harvard Medical School

Watch on YouTube

Respiratory Health Study
Boston Children's Hospital

Join study

The screenshot shows a web browser window with the address bar displaying 'health.google'. The page title is 'Participate in Health Studies - Google Health'. The Google Health logo is at the top left. Below the logo, there are two tabs: 'Digital wellbeing study' and 'Respiratory health study', with the latter being selected and highlighted in blue. The main heading is 'Help researchers better understand respiratory diseases'. Below this, a paragraph states: 'The first study available is a respiratory health study conducted by Boston Children's Hospital and Harvard Medical School. If you participate in this study, you'll provide data to help researchers understand how demographics, health history, behavior, and mobility patterns contribute to the spread of respiratory illnesses.' A button labeled 'Join respiratory health study' is positioned below the text. At the bottom, there is a circular profile picture of Dr. John Brownstein, followed by his name 'Dr. John Brownstein' and a short bio: 'Dr. John Brownstein, professor at Harvard Medical School and Chief Innovation Officer of Boston Children's Hospital.'

health.google

Participate in Health Studies - Google Health

Participate in Health Studies - Google Health

Google Health


Digital wellbeing study

Respiratory health study

Help researchers better understand respiratory diseases

The first study available is a respiratory health study conducted by Boston Children's Hospital and Harvard Medical School. If you participate in this study, you'll provide data to help researchers understand how demographics, health history, behavior, and mobility patterns contribute to the spread of respiratory illnesses.

[Join respiratory health study](#)



Dr. John Brownstein

Dr. John Brownstein, professor at Harvard Medical School and Chief Innovation Officer of Boston Children's Hospital.

health.google

Participate in Health Studies - Google Health

Participate in Health Studies - Google Health


Google Health

Digital wellbeing study Respiratory health study

Contribute to an understanding of digital wellbeing

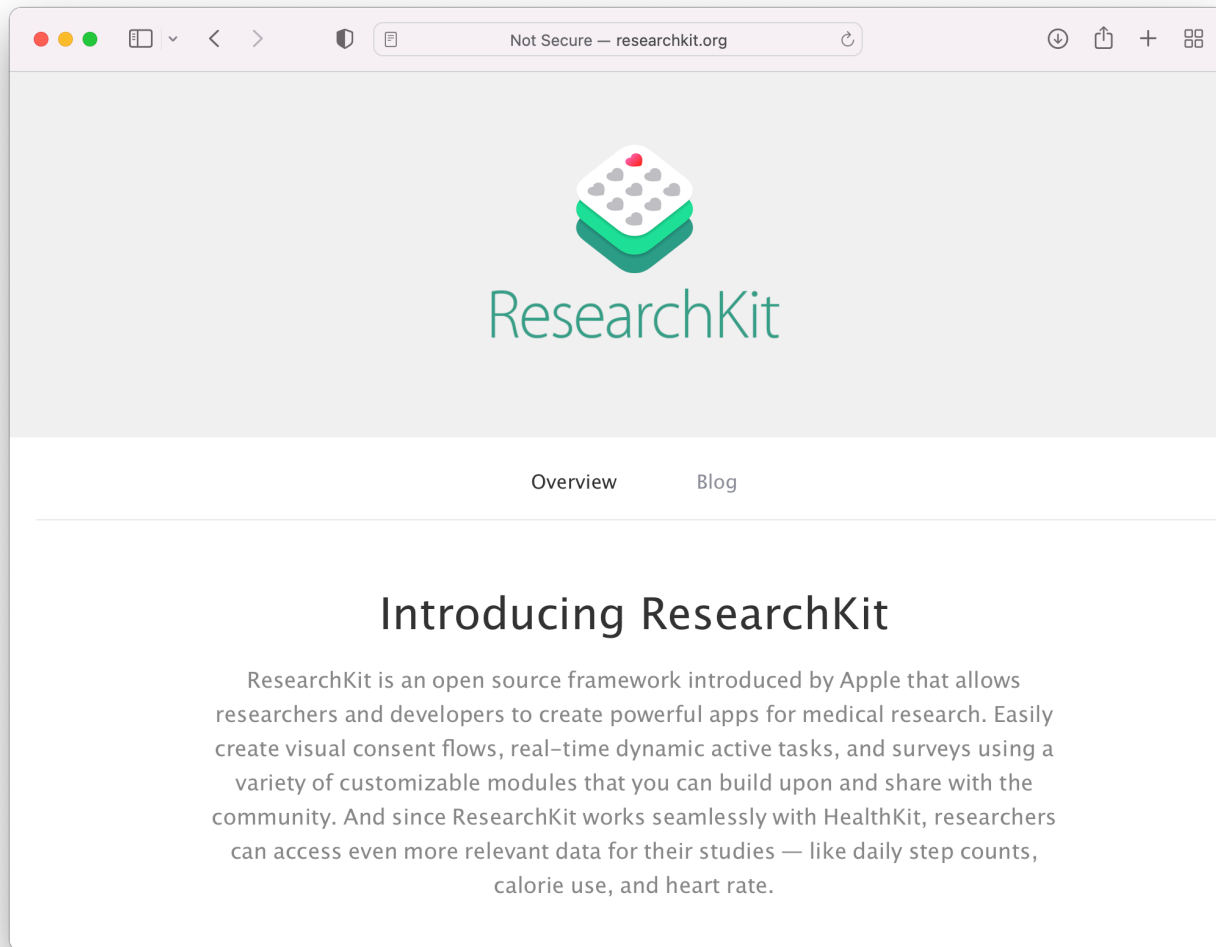
The second study available is a digital wellbeing study conducted by the Center for Digital Mental Health at the University of Oregon. If you participate in this study, you'll provide data to help researchers understand how patterns of smartphone use are associated with mental and physical wellbeing.

[Join digital wellbeing study](#)



Dr. Nicholas Allen

Dr. Nicholas Allen, Ann Swindells Professor of Clinical Psychology and Director of the Center for Digital Mental Health at the University of Oregon.



Digital Cohorts & Trials

Digital Cohort Examples

npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

Assessment of menstrual health status and evolution through mobile apps for fertility awareness

Laura Symul^{1,2}, Katarzyna Wac^{1,3,4}, Paula Hillard⁵ and Marcel Salathé²

For most women of reproductive age, assessing menstrual health and fertility typically involves regular visits to a gynecologist or another clinician. While these evaluations provide critical information on an individual's reproductive health status, they typically rely on memory-based self-reports, and the results are rarely, if ever, assessed at the population level. In recent years, mobile apps for menstrual tracking have become very popular, allowing us to evaluate the reliability and tracking frequency of millions of self-observations, thereby providing an unparalleled view, both in detail and scale, on menstrual health and its evolution for large populations. In particular, the primary aim of this study was to describe the tracking behavior of the app users and their overall observation patterns in an effort to understand if they were consistent with previous small-scale medical studies. The secondary aim was to investigate whether their precision allowed the detection and estimation of ovulation timing, which is critical for reproductive and menstrual health. Retrospective self-observation data were acquired from two mobile apps dedicated to the application of the sympto-thermal fertility awareness method, resulting in a dataset of more than 30 million days of observations from over 2.7 million cycles for two hundred thousand users. The analysis of the data showed that up to 40% of the cycles in which users were seeking pregnancy had recordings every single day. With a modeling approach using Hidden Markov Models to describe the collected data and estimate ovulation timing, it was found that follicular phases average duration and range were larger than previously reported, with only 24% of ovulations occurring at cycle days 14 to 15, while the luteal phase duration and range were in line with previous reports, although short luteal phases (10 days or less) were more frequently observed (in up to 20% of cycles). The digital epidemiology approach presented here can help to lead to a better understanding of menstrual health and its connection to women's health overall, which has historically been severely understudied.

npj Digital Medicine (2019)2:64; <https://doi.org/10.1038/s41746-019-0139-4>

Digital Cohorts & Trials

Digital Cohort Examples

nature
human behaviour

ARTICLES

<https://doi.org/10.1038/s41562-020-01046-9>



Daily, weekly, seasonal and menstrual cycles in women's mood, behaviour and vital signs

Emma Pierson^{1,2}, Tim Althoff³, Daniel Thomas⁴, Paula Hillard⁵ and Jure Leskovec^{1,6}

Dimensions of human mood, behaviour and vital signs cycle over multiple timescales. However, it remains unclear which dimensions are most cyclical, and how daily, weekly, seasonal and menstrual cycles compare in magnitude. The menstrual cycle remains particularly understudied because, not being synchronized across the population, it will be averaged out unless menstrual cycles can be aligned before analysis. Here, we analyse 241 million observations from 3.3 million women across 109 countries, tracking 15 dimensions of mood, behaviour and vital signs using a women's health mobile app. Out of the daily, weekly, seasonal and menstrual cycles, the menstrual cycle had the greatest magnitude for most of the measured dimensions of mood, behaviour and vital signs. Mood, vital signs and sexual behaviour vary most substantially over the course of the menstrual cycle, while sleep and exercise behaviour remain more constant. Menstrual cycle effects are directionally consistent across countries.

Daily, weekly, seasonal and menstrual cycles in human behaviour, health and vital signs affect health and happiness. Daily cycles are implicated in sleep¹ and obesity²; seasonal cycles in mood disorders³ and the menstrual cycle in fertility⁴

studying individual-specific cycles requires an additional piece of data: where each person is in their cycle at each timepoint (cycle phase). Failing to account for this renders individual-specific cycles invisible because cycles do not begin on the same day for each per-

Digital Cohorts & Trials

Digital Trial Examples

An Internet-Based Randomized, Placebo-Controlled Trial of Kava and Valerian for Anxiety and Insomnia

Bradly P. Jacobs, MD, MPH, Stephen Bent, MD, Jeffrey A. Tice, MD, Terri Blackwell, MA, and Steven R. Cummings, MD, FACP

Abstract: The herbal extracts kava and valerian are the leading dietary supplements used in the self-management of anxiety and insomnia, respectively. There is limited evidence to support their effectiveness for these common symptoms. The Internet has been used to a limited extent for research, but it is not known whether randomized controlled trials can be conducted entirely using Internet technology.

We performed a randomized, double-blind, placebo-controlled trial using a novel Internet-based design to determine if kava is effective for reducing anxiety and if valerian is effective for improving sleep quality. E-mail recruitment letters and banner advertisements on websites were used to recruit a large pool of interested participants (1551) from 45 states over an 8-week period.

Participants receiving placebo had a 14.4 point decrease in anxiety symptoms on the STAI-State score and an 8.3 point decrease in insomnia symptoms on the ISI. Those receiving kava had similar reductions in STAI-State score (2.7 point greater reduction in placebo compared with kava; 95% confidence interval [CI], -0.8 to $+6.2$). Those receiving valerian and placebo had similar improvements in sleep (0.4 point greater reduction in the placebo than the valerian group; 95% CI, -1.3 to $+2.1$). Results were similar when limited to the 83% of participants who adhered to study compounds for all 4 weeks.

Neither kava nor valerian relieved anxiety or insomnia more than placebo. This trial demonstrates the feasibility of conducting randomized, blinded trials entirely via the Internet.

Digital Cohorts & Trials

Digital Trial Examples

- 2005!

treatment is particularly prone to publication bias, with negative trials remaining unpublished^{11,13}.

In conclusion, we found that neither kava nor valerian was superior to placebo for the self-management of anxiety and sleep problems. This trial also demonstrates the feasibility of conducting rigorous randomized, blinded trials directly with participants entirely via the Internet.

Digital Cohorts & Trials

Digital Trial Examples

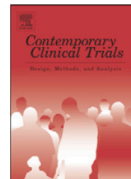
Contemporary Clinical Trials 38 (2014) 190–197



Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial



Web-based trial to evaluate the efficacy and safety of tolterodine ER 4 mg in participants with overactive bladder: REMOTE trial



Miguel Orri^{a,*}, Craig H. Lipset^b, Bradly P. Jacobs^{c,d}, Anthony J. Costello^c, Steven R. Cummings^{c,d,e}

^a Pfizer Ltd, Tadworth, United Kingdom

^b Pfizer Inc, New York, NY, USA

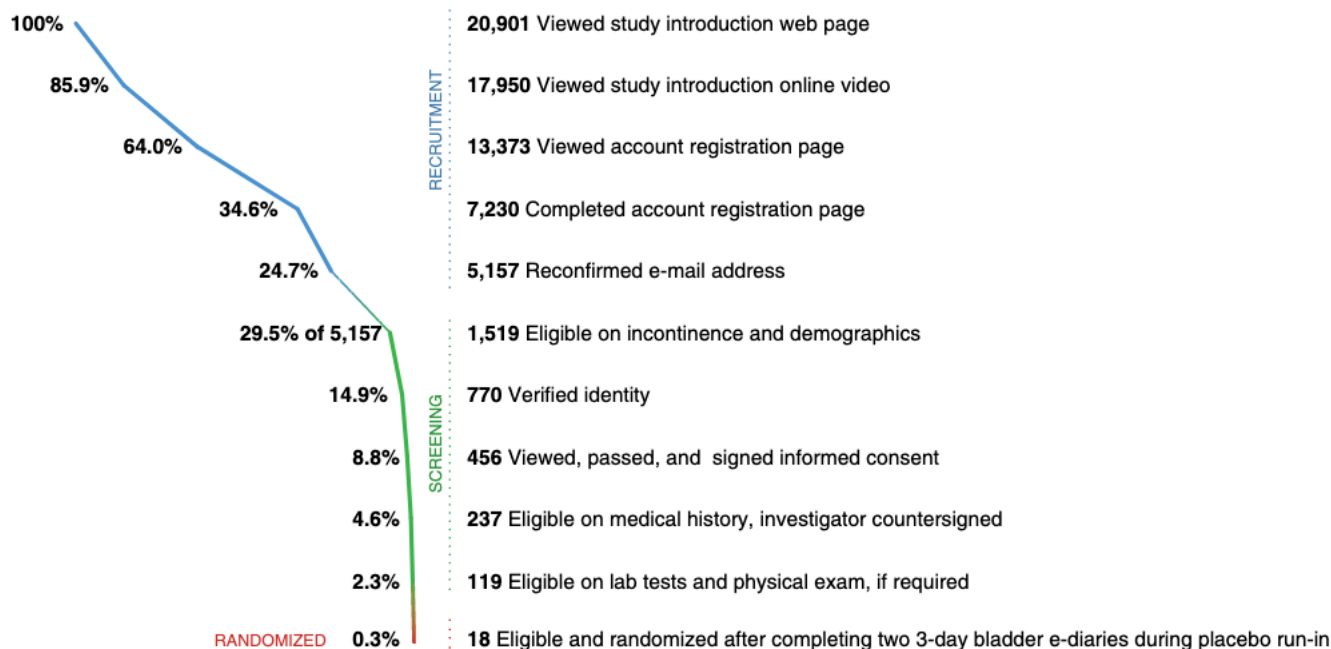
^c University of California-San Francisco, San Francisco, CA, USA

^d Mytrus, San Francisco, CA, USA

^e California Pacific Medical Center Research Institute, San Francisco, CA, USA

Digital Cohorts & Trials

Digital Trial Examples



Digital Cohorts & Trials

Digital Trial Examples

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

AUGUST 6, 2020

VOL. 383 NO. 6

A Randomized Trial of Hydroxychloroquine as Postexposure Prophylaxis for Covid-19

D.R. Boulware, M.F. Pullen, A.S. Bangdiwala, K.A. Pastick, S.M. Lofgren, E.C. Okafor, C.P. Skipper, A.A. Nascene, M.R. Nicol, M. Abassi, N.W. Engen, M.P. Cheng, D. LaBar, S.A. Lothar, L.J. MacKenzie, G. Drobot, N. Marten, R. Zarychanski, L.E. Kelly, I.S. Schwartz, E.G. McDonald, R. Rajasingham, T.C. Lee, and K.H. Hullsiek

ABSTRACT

BACKGROUND

Coronavirus disease 2019 (Covid-19) occurs after exposure to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). For persons who are exposed, the standard of care is observation and quarantine. Whether hydroxychloroquine can prevent symptomatic infection after SARS-CoV-2 exposure is unknown.

METHODS

We conducted a randomized, double-blind, placebo-controlled trial across the United States and parts of Canada testing hydroxychloroquine as postexposure prophylaxis. We enrolled adults who had household or occupational exposure to someone with confirmed Covid-19 at a distance of less than 6 ft for more than 10 minutes while wearing neither a face mask nor an eye shield (high-risk exposure) or while

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Boulware at the University of Minnesota, 689 23rd Ave., Minneapolis, MN 55455, or at boulw001@umn.edu.

This article was published on June 3, 2020, at [NEJM.org](https://www.nejm.org).

N Engl J Med 2020;383:517-25.

DOI: 10.1056/NEJMoa2016638

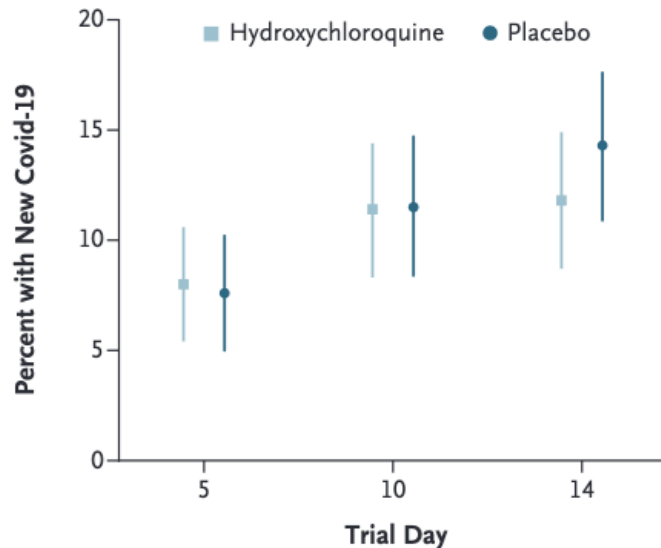
Copyright © 2020 Massachusetts Medical Society.

Digital Cohorts & Trials

Digital Trial Examples

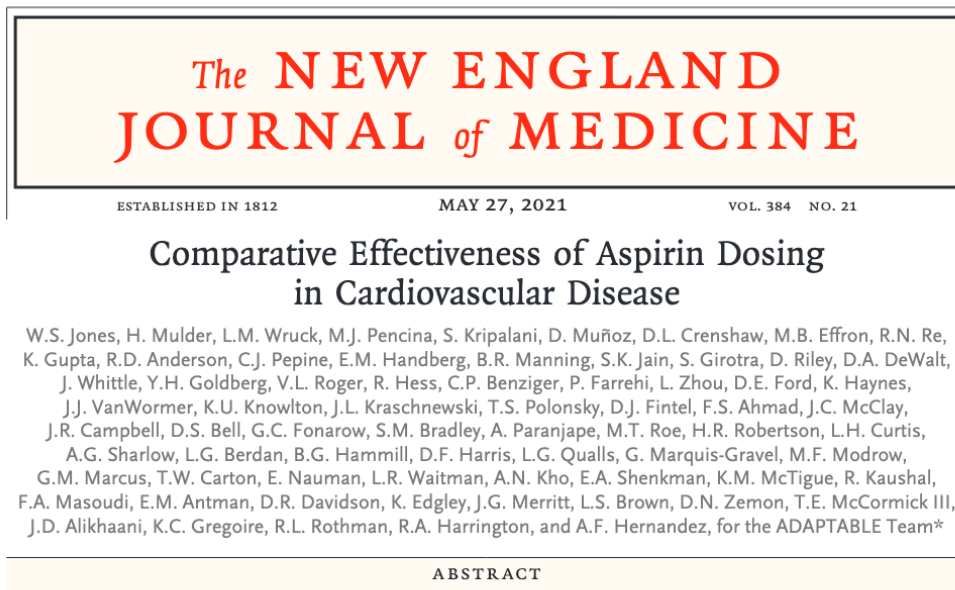
SETTING

Recruitment was performed primarily with the use of social media outreach as well as traditional media platforms. Participants were enrolled nationwide in the United States and in the Canadian provinces of Quebec, Manitoba, and Alberta. Participants enrolled themselves through a secure Internet-based survey using the Research Electronic Data Capture (REDCap) system.¹³ After participants read the consent form, their comprehension of its contents was assessed; participants provided a digitally captured signature to indicate informed consent. We sent follow-up e-mail surveys on days 1, 5, 10, and 14. A survey at 4 to



Digital Cohorts & Trials

Digital Trial Examples



BACKGROUND

The appropriate dose of aspirin to lower the risk of death, myocardial infarction, and stroke and to minimize major bleeding in patients with established atherosclerotic cardiovascular disease is a subject of controversy.

METHODS

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Jones at the Division of Cardiology, Duke University Health System, DUMC 3330,

- 15,076 patients

Digital Cohorts & Trials

Recruitment

- Recruitment can be difficult in non-digital studies: 19% of phase 2 and phase 3 trials fail to reach adequate numbers
- Digital provides massive reach (see. e.g. Apple Heart Study, DETECT, COVID Symptom study, etc.)
- In mixed recruitment, online recruitment typically vastly outnumbers offline recruitment

Digital Cohorts & Trials

Consent

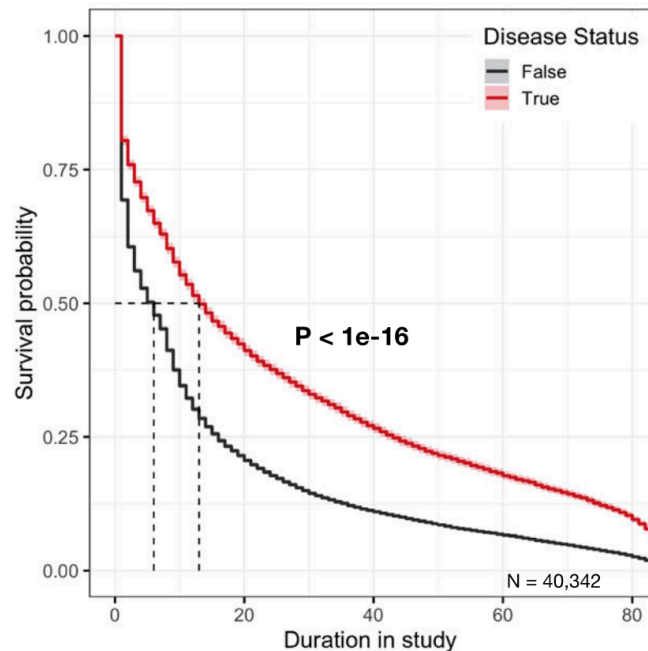
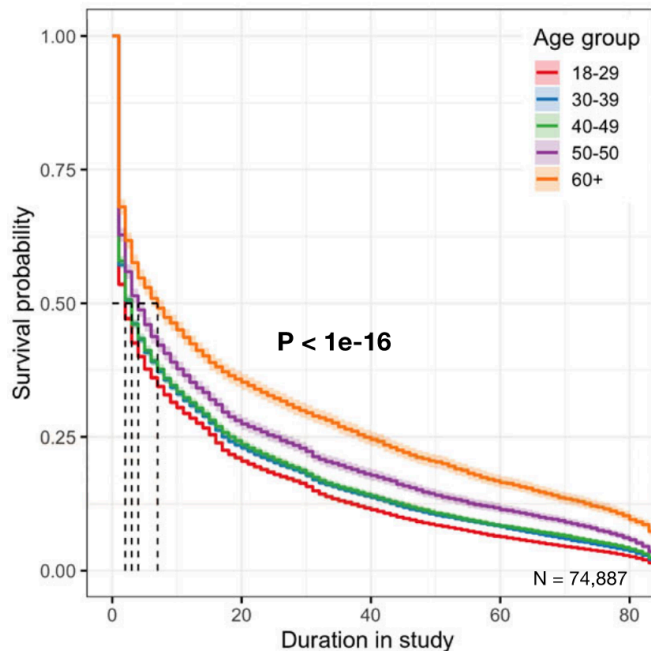
- In any study in which participants need to do something, or non-public data is investigated, people need to give *informed consent*.
- Electronic informed consent (e-consent) now widespread, but initially caused concern (“will people read it?”).
- Today, modern e-consent can have interactive elements
- E-consent is better at avoiding undue influence
- Biggest challenge: verification

Digital Cohorts & Trials

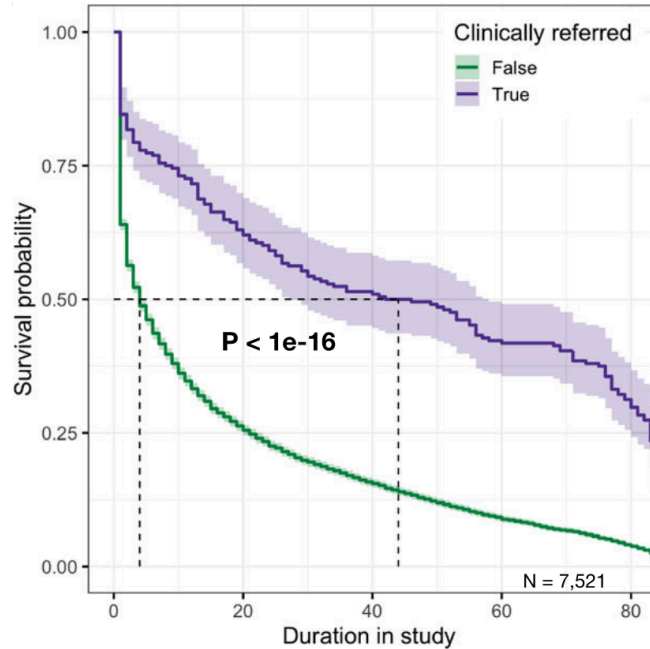
Retention

- Retention in digital studies is a big issue.
- Financial incentives have a big impact.

Digital Cohorts & Trials Retention



Digital Cohorts & Trials Retention



Digital Cohorts & Trials

Data Collection

- Data collection “in situ”: longitudinal, continuous, affordable
- In psychology, approach is known as “experience sampling methods”, and generally as “ecological momentary assessment”
- Active vs. passive data collection
- Data quality? Recall bias, social desirability bias, etc.
- Digital biomarkers need to be continuously assessed

Digital Cohorts & Trials

Data Analysis

- Often highly divers, multimodal (need of multimodal analysis, multimodal AI)
- Ideally analyzed together, as just one modality might give misleading results.
- Missing data: statistical approaches, e.g. imputation.

The screenshot shows the homepage of the **FOOD & YOU** website. The browser address bar displays **foodandyou.org**. The navigation menu includes **Home**, **About**, **FAQs**, and **Get started**, along with a language selector set to **en** and a **Log in** link. The main heading reads: **Track your individual response to food and help advance science**. Below this is a line graph with a grey line and orange data points, accompanied by a downward arrow icon. At the bottom, five categories are listed with corresponding icons: **Citizen Science** (network icon), **Nutrition** (fork and knife icon), **Blood Sugar** (bar chart icon), **Microbiome** (microbes icon), and **Lifestyle** (person icon).

PLOS DIGITAL HEALTH

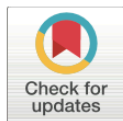
RESEARCH ARTICLE

Food & You: A digital cohort on personalized nutrition

Harris Héritier¹, Chloé Allémann¹, Oleksandr Balakiriev¹, Victor Boulanger¹, Sean F. Carroll¹, Noé Froidevaux¹, Germain Hugon¹, Yannis Jaquet¹, Djilani Kebaili¹, Sandra Riccardi¹, Geneviève Rousseau-Leupin¹, Rahel M. Salathé¹, Talia Salzman¹, Rohan Singh¹, Laura Symul^{1,2}, Elif Ugurlu-Baud¹, Peter de Verteuil¹, Marcel Salathé¹*

1 Digital Epidemiology Lab, School of Life Sciences, School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland, **2** Department of Statistics, Stanford University, Stanford, California, United States of America

* marcel.salathe@epfl.ch



Abstract

Nutrition is a key contributor to health. Recently, several studies have identified associations between factors such as microbiota composition and health-related responses to dietary intake, raising the potential of personalized nutritional recommendations. To further our understanding of personalized nutrition, detailed individual data must be collected from participants in their day-to-day lives. However, this is challenging in conventional studies that require clinical measurements and site visits. So-called digital or remote cohorts allow *in situ* data collection on a daily basis through mobile applications, online services, and wearable sensors, but they raise questions about study retention and data quality. “Food & You” is a personalized nutrition study implemented as a digital cohort in which participants track food intake, physical activity, gut microbiota, glycemia, and other data for two to four weeks.

OPEN ACCESS

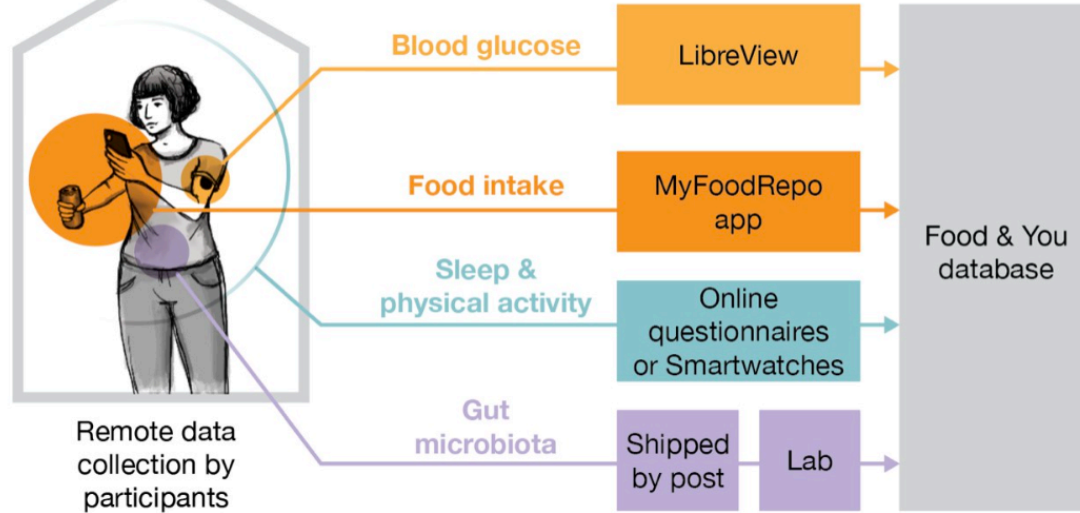
Citation: Héritier H, Allémann C, Balakiriev O, Boulanger V, Carroll SF, Froidevaux N, et al. (2023) Food & You: A digital cohort on personalized nutrition. PLOS Digit Health 2(11): e0000389. <https://doi.org/10.1371/journal.pdig.0000389>

Editor: Lais Duarte Batista, Harvard University, UNITED STATES

Digital Cohorts & Trials

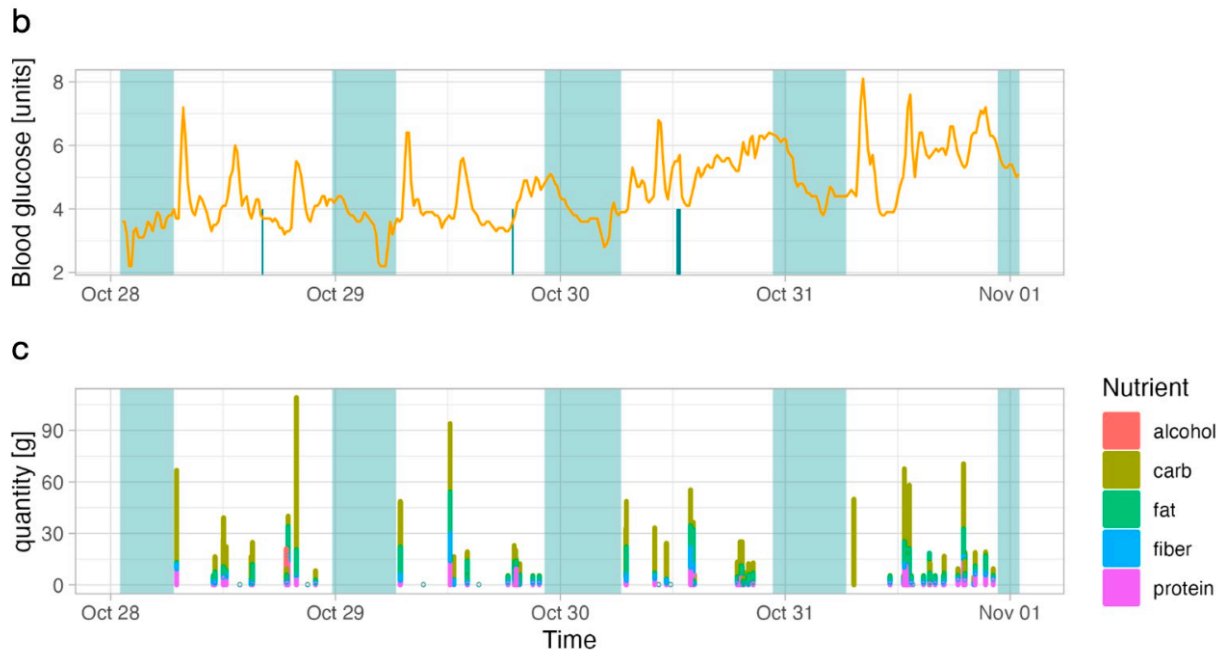
Food & You

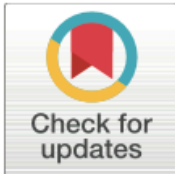
a



Digital Cohorts & Trials

Food & You





OPEN ACCESS

Citation: Héritier H, Allémann C, Balakiriev O, Boulanger V, Carroll SF, Froidevaux N, et al. (2023) Food & You: A digital cohort on personalized nutrition. PLOS Digit Health 2(11): e0000389. <https://doi.org/10.1371/journal.pdig.0000389>

Editor: Lais Duarte Batista, Harvard University, UNITED STATES

Received: August 3, 2023

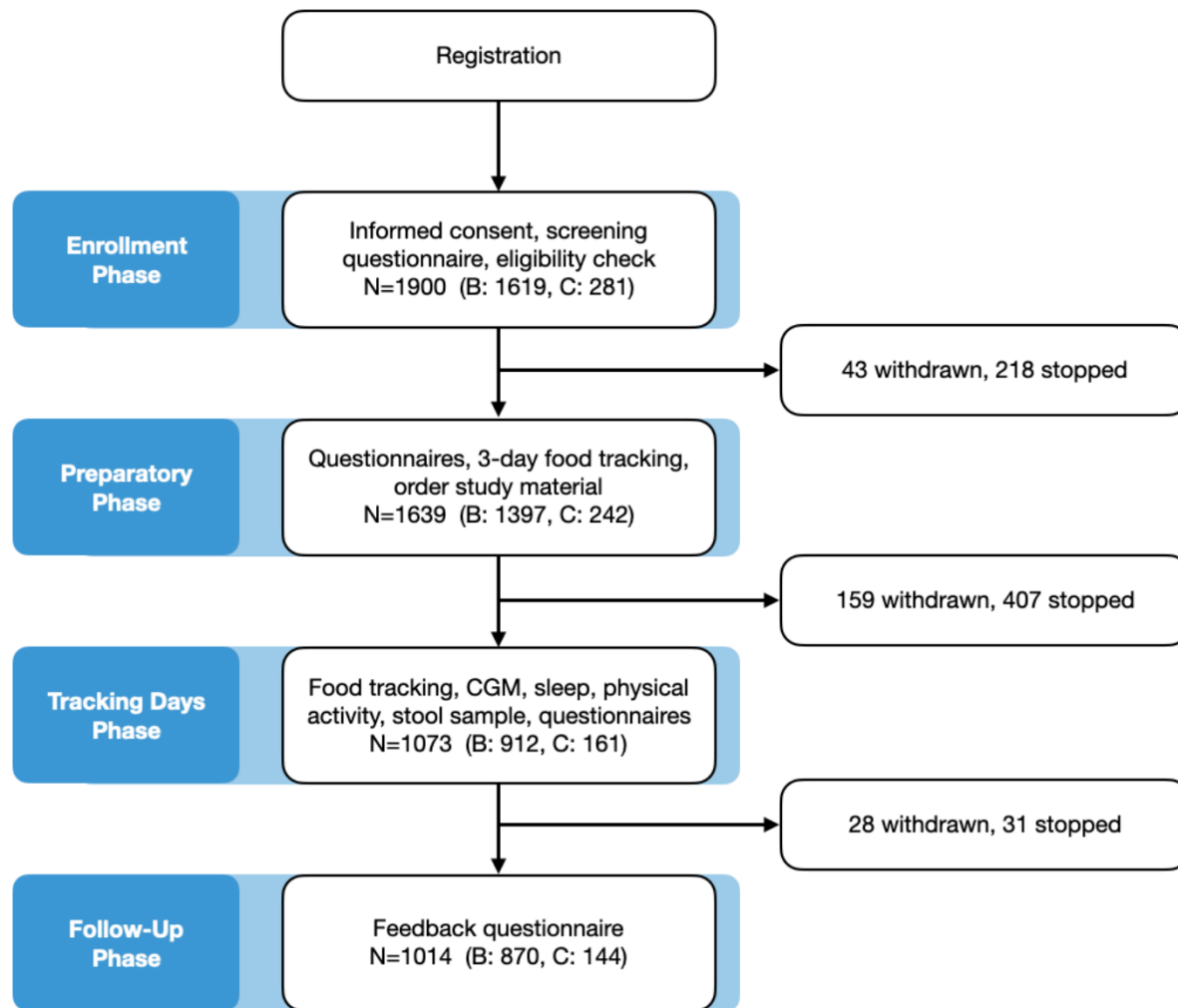
Accepted: October 13, 2023

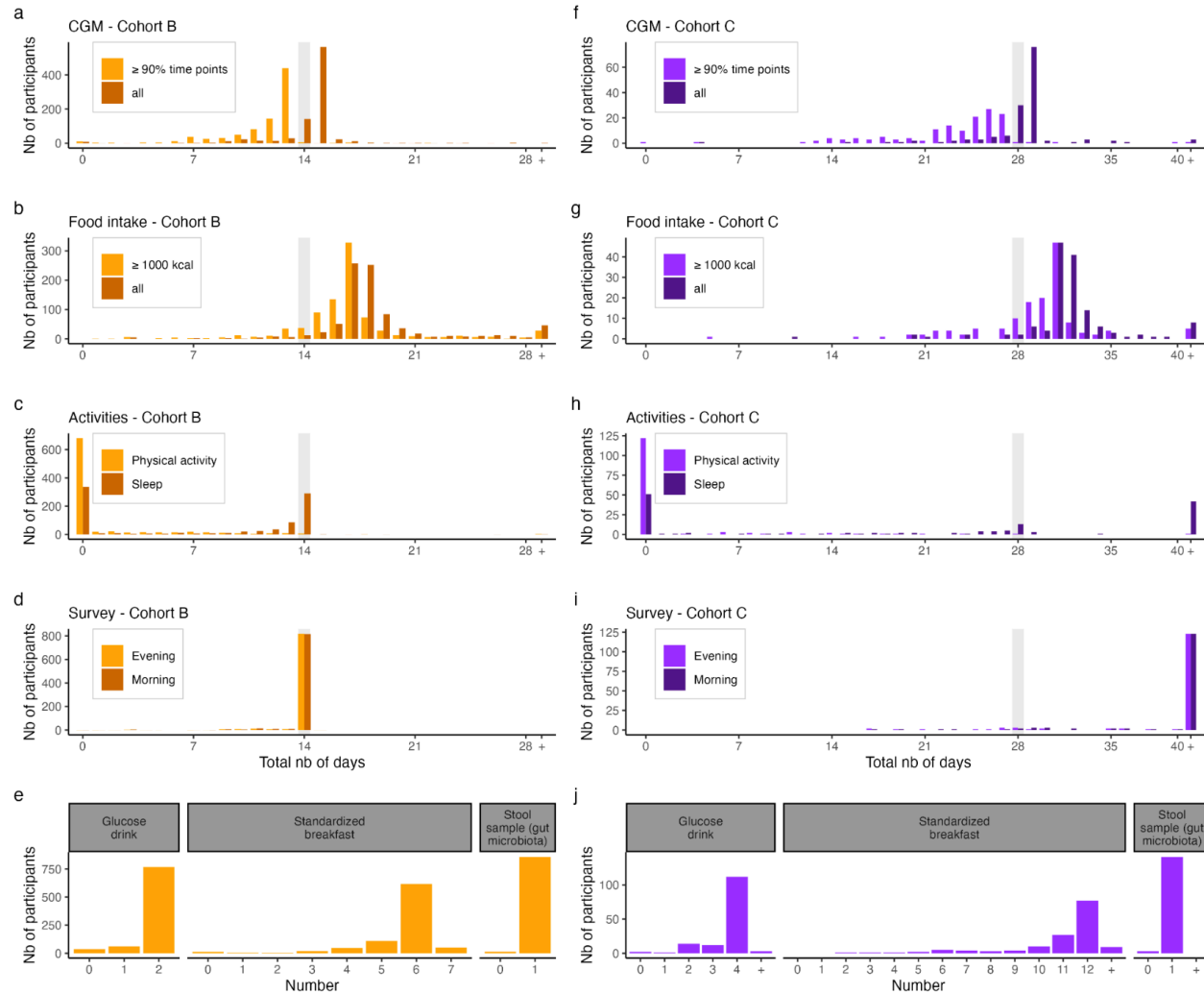
Published: November 30, 2023

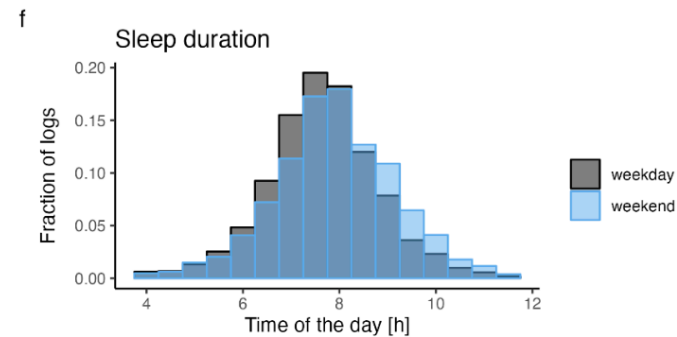
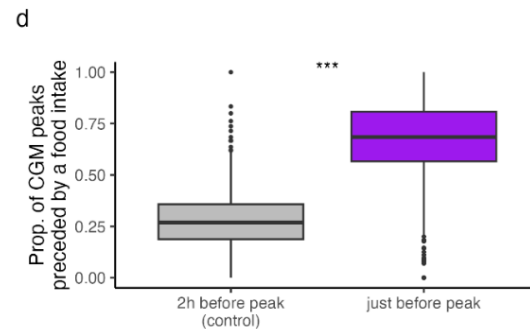
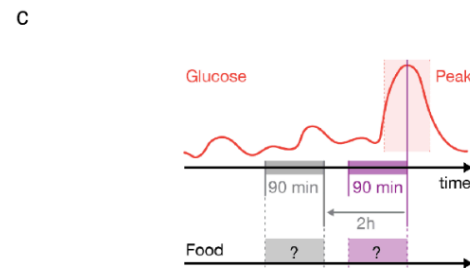
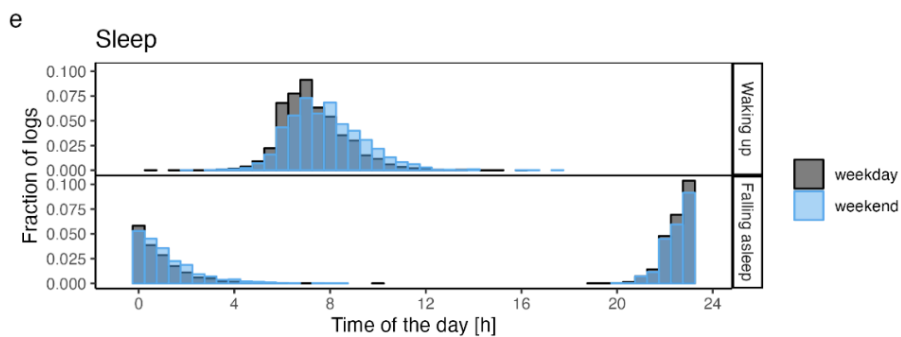
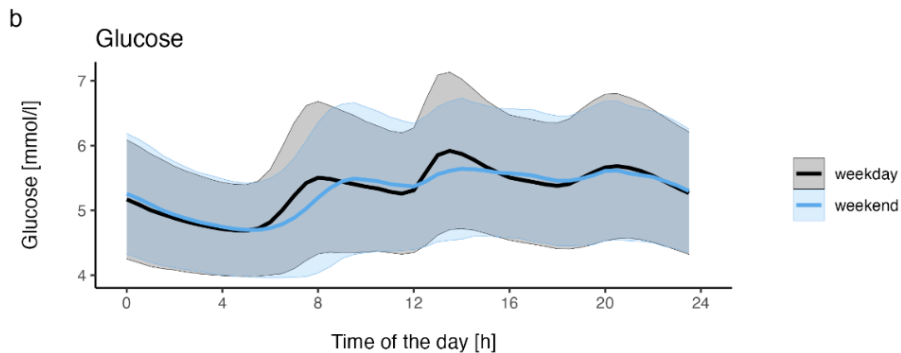
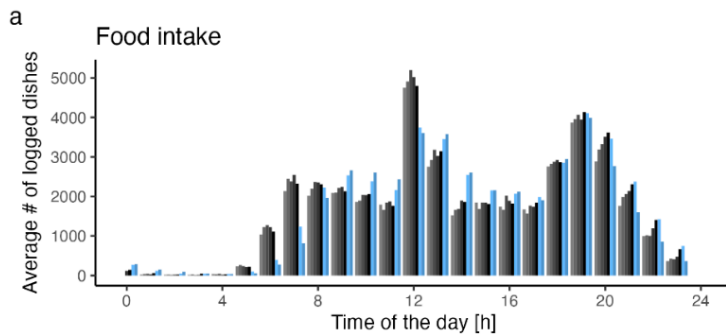
Copyright: © 2023 Héritier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Nutrition is a key contributor to health. Recently, several studies have identified associations between factors such as microbiota composition and health-related responses to dietary intake, raising the potential of personalized nutritional recommendations. To further our understanding of personalized nutrition, detailed individual data must be collected from participants in their day-to-day lives. However, this is challenging in conventional studies that require clinical measurements and site visits. So-called digital or remote cohorts allow *in situ* data collection on a daily basis through mobile applications, online services, and wearable sensors, but they raise questions about study retention and data quality. “Food & You” is a personalized nutrition study implemented as a digital cohort in which participants track food intake, physical activity, gut microbiota, glycemia, and other data for two to four weeks. Here, we describe the study protocol, report on study completion rates, and describe the collected data, focusing on assessing their quality and reliability. Overall, the study collected data from over 1000 participants, including high-resolution data of nutritional intake of more than 46 million kcal collected from 315,126 dishes over 23,335 participant days, 1,470,030 blood glucose measurements, 49,110 survey responses, and 1,024 stool samples for gut microbiota analysis. Retention was high, with over 60% of the enrolled participants completing the study. Various data quality assessment efforts suggest the captured high-resolution nutritional data accurately reflect individual diet patterns, paving the way for digital cohorts as a typical study design for personalized nutrition.



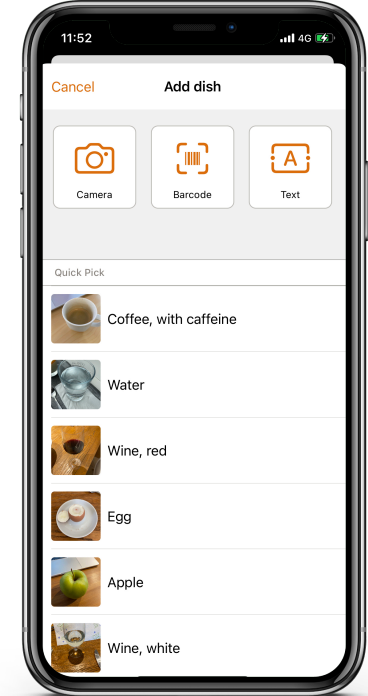
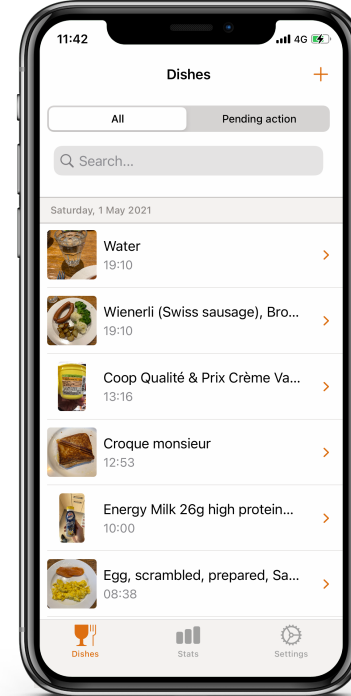




Digital Cohorts & Trials

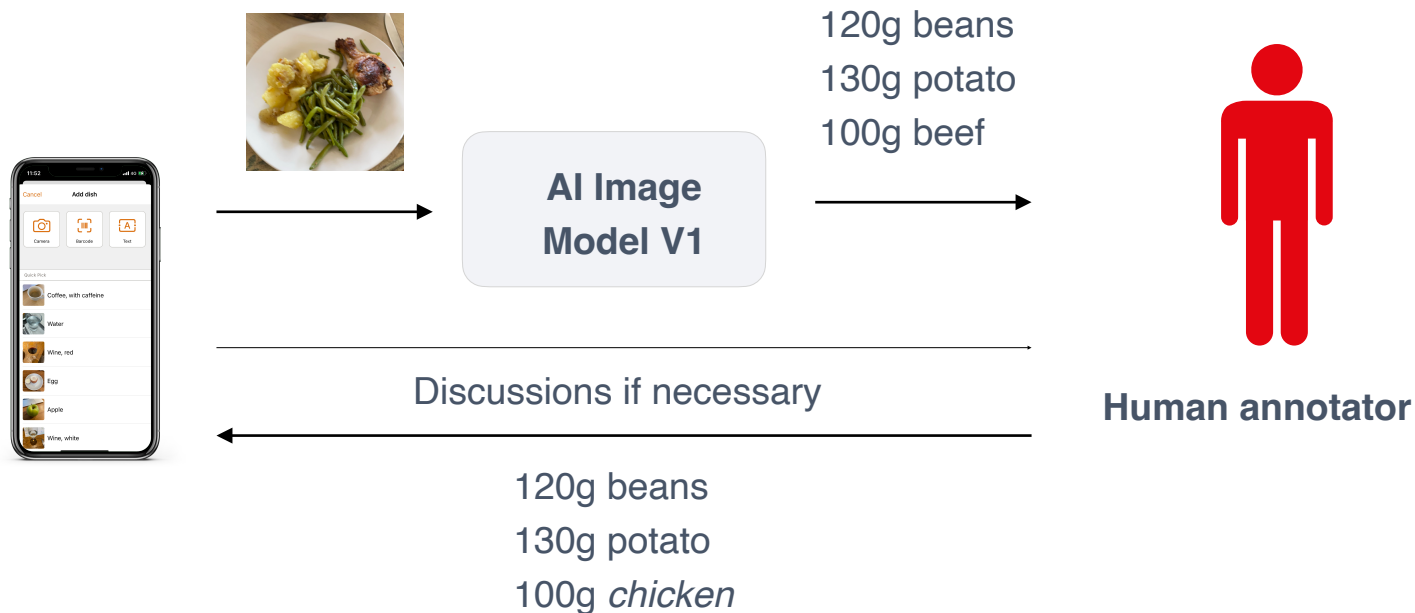
MyFoodRepo

- MyFoodRepo: an app designed collect accurate diet data for the Food & You study
- MyFoodRepo now used in multiple nutritional studies in Switzerland, Germany, Netherland, and US (with interest from France, Italy, India, New Zealand).



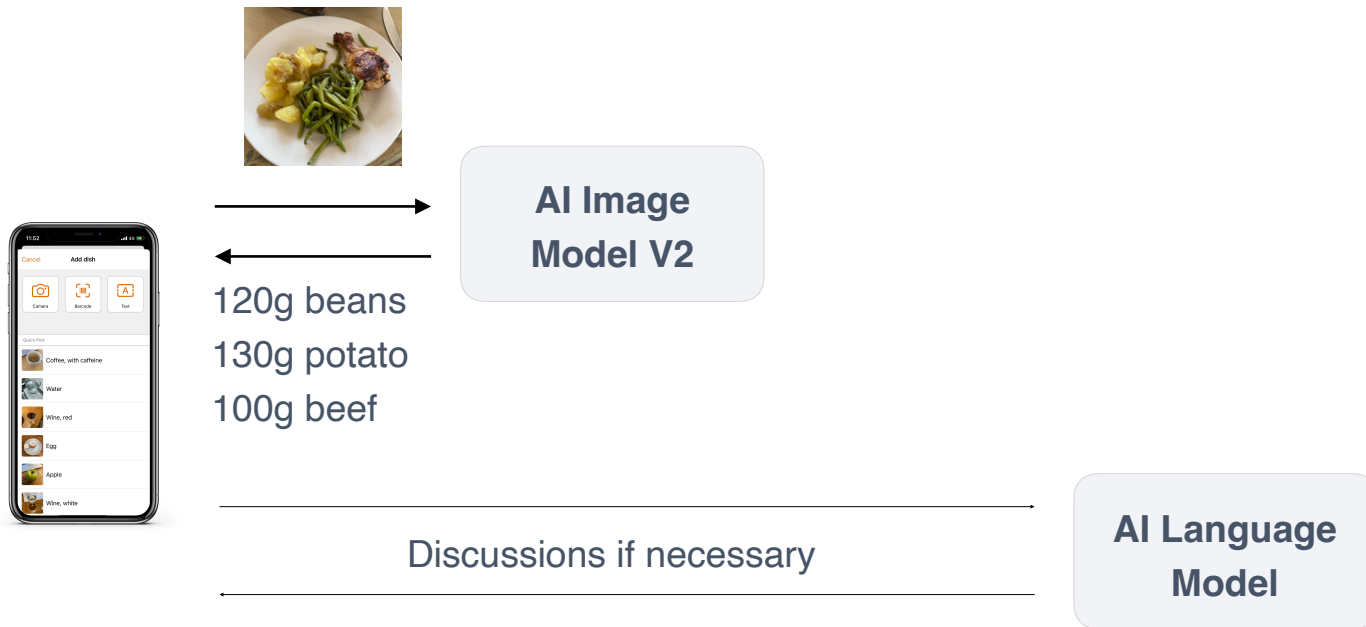
Digital Cohorts & Trials

MyFoodRepo



Digital Cohorts & Trials

MyFoodRepo V2



Digital Cohorts & Trials

Future

- Combining digital cohort studies with digital public health surveillance
- As everything moves digital, security becomes critical: federated learning, secure multi-party computation, and differential privacy
- Large language models (LLMs) might solve retention problem through increased engagement.