

You + ME Registry: A Research Platform to Facilitate Clinical and Therapeutic Discoveries in ME/CFS and Related Diseases

Allison Ramiller¹, Kathleen Mudie^{1*}, Elle Seibert¹, Sadie Whittaker¹

*Corresponding author: research@solvecfs.org

¹The Solve ME/CFS Initiative (Solve M.E.)

ABSTRACT

ME/CFS (Myalgic Encephalomyelitis / Chronic Fatigue Syndrome) is a chronic, complex, heterogeneous disease that affects millions and lacks both diagnostics and treatments. Big data, or the collection of vast quantities of data that can be mined for information, has transformed the understanding of many complex illnesses like cancer (1,2) and multiple sclerosis (3,4), by dissecting heterogeneity, identifying subtypes, and enabling the development of personalized treatments. It is possible that big data can reveal the same for ME/CFS.

Solve M.E. developed and launched the You + ME Registry to collect longitudinal health data from people with ME/CFS, people with Long COVID (LC) and control volunteers using rigorous protocols designed to harmonize with other groups collecting data from similar groups of people.

The Registry is an invaluable resource because it integrates with a symptom tracking app, as well as a biorepository, to provide a robust and rich dataset that is available to qualified researchers. Accordingly, it facilitates collaboration that may ultimately uncover causes and help accelerate the development of therapies.

KEYWORDS

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome, Long COVID, data acquisition source, post-infectious, longitudinal cohort study, patient-powered

INTRODUCTION

BACKGROUND TO ME/CFS

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a chronic, complex, systemic disease that affects anywhere from 1.5 to 3.4 million people in the USA (5,6), with an estimated annual economic cost of \$36–\$51 billion (6). The etiology of ME/CFS is unknown but it is associated with infectious triggers or other precipitating events, such as injury, trauma, or exposure to environmental hazards (7). ME/CFS has not been correlated with age, race, or socioeconomic group; however, 3 to 4 times as many women as men present with symptoms (7).

The true impact of this disease remains uncertain largely because ME/CFS is commonly misdiagnosed, with primary reasons referenced as lack of confidence from clinicians and the absence of an established biomarker (8). Clinicians must rely on a detailed evaluation of the presence of characteristic symptoms, health history and a physical examination. While many clinicians are aware of ME/CFS, they often lack essential experience necessary to diagnose this complex disease.

ME/CFS is characterized by debilitating fatigue and other multi-system physical and neurocognitive symptoms, which are exacerbated following minimal physical or mental exertion. Post-exertional malaise (PEM) is the hallmark symptom of ME/CFS and is associated with elevated symptom burden and psychological distress (7). Otherwise, substantial clinical heterogeneity exists among patients, with a range of symptoms that includes orthostatic intolerance (OI), postural tachycardia syndrome (POTS), brain fog, headaches, unrefreshing sleep and other sleep dysfunction, joint pain, and muscle pain/fibromyalgia (7).

Although understanding of biological abnormalities in ME/CFS has increased considerably in recent decades (9–14), the lack of a diagnostic biomarker, complex clinical presentation, and patient heterogeneity has severely hampered progress in treatment of ME/CFS, causing many patients to have poor outcomes (7,15). While there are no treatments approved for ME/CFS, a number of pharmacological and non-pharmacological interventions are used to manage symptoms. There have been a few promising clinical trials undertaken; however, they have not resulted in approved therapies for the ME/CFS population. For example, researchers analyzing data from a Phase III trial of rintatolimod (tradename Ampligen®) found that only a subset of ME/CFS patients defined by relatively short duration of disease (symptom onset within 2-8 years) had improved exercise tolerance in response to rintatolimod (16). There is an acute interest in identifying subsets of patients who might be responders (16,17).

Methodological problems, small sample sizes, and/or selection bias towards those less severely affected by ME/CFS (18), also create persistent roadblocks to progress the understanding of the disease. To increase research validity and drive progress, a more complete characterization is needed. Patient registries and biobanks are invaluable resources for collecting real-world clinical data at a large scale that can then be used for research.

In May 2020, Solve M.E., a non-profit organization whose mission is to make ME/CFS and other post-infection diseases widely understood, diagnosable, and treatable, launched the You + ME Registry and Biobank (Registry) (www.youandmeregistry.com). The Registry is a secure online data repository where people with ME/CFS, related diseases, and control volunteers can enter information on their health.

Given strong evidence for viral etiology of ME/CFS and experience from the SARS-CoV-1 pandemic in 2003 (19–21), there is the potential for the current SARS-CoV-2 pandemic to lead to a substantial increase in the number of ME/CFS cases. There is evidence that some people experience long-term effects from COVID-19 (termed Long COVID) with a constellation of symptoms reported that are strikingly similar to those reported for ME/CFS (22–27). The COVID-19 pandemic and emergence of long-term symptoms in some individuals presents a unique scientific opportunity to understand factors of resistance and susceptibility to long-term, post-viral impacts. In December 2020, in response to the increasing number of individuals with Long COVID (LC), the Registry was adapted and opened up to those who are suffering from the long-term effects of COVID-19 and ‘control’ individuals who had COVID-19 but do not have long-term effects.

The Registry is designed to be a foundational resource for research and it is unique for several reasons: 1) it integrates a symptom tracking app that can provide more data points for dynamic/cyclic chronic diseases; 2) it is a rigorous, systematic infrastructure for collecting data that was co-created with the ME/CFS and Long COVID communities; 3) the data and patient cohorts are available to all qualified researchers, supporting numerous scientific studies and engaging a network of expertise; 4) it is centered around principles of data harmonization and collaboration with other research studies to accelerate the search for causes and therapies.

METHODS

Technical Infrastructure

The You + ME Registry data is securely stored in a mongoDB Atlas cloud database (<https://www.mongodb.com/>). The database is encrypted and participant data is encrypted in transit using industry-standard TLS/SSL to protect sensitive information when it's transmitted to and from the front end apps and backend database. Amazon Cognito is used for user authentication and access control.

Human-Centered Design

The creation of the Registry platform and data collection process integrated community input and human-centered design (HCD) methodologies (28). One-on-one phone interviews were conducted with stakeholders and experts, including people with ME/CFS across the disease severity spectrum, caregivers, researchers, clinicians, and experts in informatics and human-centered design.

Qualitative and quantitative research methods were used to understand the health tracking priorities of people with ME/CFS and their preferred data collection frequency. A community survey collected perspectives on a symptom tracking app and these results were compared to results from a survey deployed by collaborators at Columbia University. This generated responses from over 1,200 people with ME/CFS. Over 30 individuals tested a beta version of the app on their Apple and Android phones—including people with ME/CFS, care partners, researchers, and clinicians—leading to the creation of Version 2.0.

While the current capabilities of the Registry can support expansive data collection, the platform is also built with the capacity to integrate new data types (e.g. passive monitoring from wearables). The user experience and data collection are continuously improved based on feedback from patients and researchers.

Data Harmonization

The validated data-collection instruments within the Registry (see Data Collection section) are aligned with those used by other researchers and clinicians studying ME/CFS patients. They include National Institutes of Health (NIH) National Institute of Neurological Disorders and Stroke (NINDS) Common Data Elements (29) to facilitate aggregation of data across studies. After consulting with community members, additional data fields were included to build a richer understanding of each participant's health history. The Registry also integrates the NINDS Centralized Globally Unique

Identifier (GUID) solution, a secure tool that generates unique IDs without exposing personally identifiable information (PII) to allow data sharing and collaboration across research groups.

Participant Recruitment

The Registry is open to all individuals with ME/CFS, LC, and other populations, including individuals with other chronic diseases, and individuals considered to be “healthy” controls. The aim is to enroll a diverse global cohort of participants who are representative of the broader ME/CFS and LC communities, and to create the largest possible global dataset to interrogate.

The Registry is promoted via Solve M.E.’s and You + ME’s dedicated @youmeregistry social media channels (Facebook, Twitter, Instagram), on a dedicated online informational website (<https://youandmeregistry.com/>), printed newsletters, and via email to the Solve M.E. listserv. It is also promoted in webinars and conference presentations.

Additional recruitment is conducted through partners in both ME/CFS and LC by leveraging their social media channels and email listservs. This is particularly important for recruitment to our ME/CFS and LC cohorts, given that many in these communities are connected via robust social media groups (30). Our LC recruitment strategy specifically is multi-pronged (Table 1). Outreach is focused on areas with historically high incidences of COVID-19 cases, e.g., New York, New Orleans and Los Angeles. Future plans for recruitment include referrals from clinicians and health systems.

Table 1. Overview of current and planned recruitment strategies for the You + ME Registry

Type of Outreach	Target Organizations/Partners and Strategy
Solve ME communication channels	Promote the Registry to our established network via: 1) a database of over 34,000 active contacts; 2) organizational and Registry social media accounts with a combined following of 6,772 on Twitter, 34,114 on Facebook, and 2,256 on Instagram; 3) our educational webinar series for researchers, clinicians, and patients.
COVID Survivor PASC patient groups	Partner with established groups serving COVID-19 survivors and individuals with PASC, including online forums and support groups on social media, to promote the Registry to their networks.
Long COVID Alliance	Partner with members to create a referral pipeline to the Registry from over 50 science, post-viral disease and patient advocacy and research organizations working together to find answers for Long COVID and post-viral illness.
Internet & social media advertising	Google ads and social media posts directed towards individuals who have experience COVID-19 and primary care providers who may be treating those with persistent symptoms of COVID-19.
Clinics/Health systems	Partner with health systems, clinics, and hospitals serving our populations of interest to provide a postcard that will be handed out to their COVID-19

	patients. The postcard will ask about the development of persistent post-viral symptoms, and direct patients to the Registry for voluntary sign up.
Membership organizations/Trade associations	Partner with healthcare workers and EMS unions; other unionized or non-unionized essential workers, such as large grocery/drug store chains, transit workers, and delivery services; university-based countrywide student organizations, athletic associations and student health networks; and medical specialty associations to share the recruitment notices to their membership.

Criteria for Selection

The Registry is open to individuals of all genders, with an anticipated gender split between males and females reflective of the gender prevalence of ME/CFS and LC. Adults (aged 18 years and above) are eligible for participation. All races and ethnic origins are included. Although we don't limit enrollment from control volunteers, we will be making every effort to ensure controls are adequately matched to patients by age, sex, race, and other key demographic indicators.

People with ME/CFS

People with ME/CFS self-diagnosed or diagnosed by a clinician are eligible to enroll. The method of diagnosis is recorded for each participant.

People with Long COVID

People who had COVID, whether confirmed by a lab test or not, are eligible to enroll. The method of initial COVID-19 diagnosis is recorded for each participant.

Control Volunteers

Control volunteers are made up of individuals without ME/CFS or LC, including individuals considered to be "healthy" controls and those with other chronic illnesses (e.g. fibromyalgia).

Symptoms Assessment and Algorithm for ME/CFS Case Criteria

The Registry algorithmically scores participant responses to the U.K. ME/CFS Biobank (UKMEB) Symptoms Assessment Questionnaire to determine fulfillment of distinct ME/CFS case criteria (31). The algorithm is licensed from the UKMEB and coded into the online platform in a separate endpoint that securely runs the data for scoring.

Data Collection

Participants sign up for the Registry via a secure website interface. Participants are asked to complete an electronic informed consent form for collection of data and to be

re-contacted for optional biosample collection or other study opportunities. After consenting to the Registry, they are guided through a series of baseline surveys (Table 2), including medical history, diagnosed conditions, symptoms and quality of life, and medication history.

Longitudinal characterization is imperative to understanding chronic illness given its evolving and cyclical nature. ME/CFS and control volunteers are sent email reminders to complete an abbreviated set of surveys every 90 days following registration (Figures 1A & 1B). Given that the disease state can be more dynamic in the first six months after symptom onset, the LC cohort is sent follow-up surveys every 30 days in the initial six months and then every 90 days for longer-term follow-up. We hypothesize that biologically and/or clinically distinct factors drive divergent outcomes in post-COVID-19 patients toward complete recovery or long-term sequelae. So, the COVID cohort is sent follow-up surveys every 30 days in the initial six months and then every 90 days for longer-term follow-up (Figure 1C). Through more frequent data collection during this critical period following recent infection, we aim to identify factors driving and/or predictive of this bifurcation in outcomes.

One-time questionnaires are also deployed through You + ME in addition to the regular longitudinal assessments. These questionnaires collect cross-sectional data from unique instruments that are not part of the routine longitudinal assessments, and currently include surveys on family health history, joint hypermobility (self-report Beighton), and COVID-19 vaccination experience.

Table 2: Baseline surveys used in the You + ME Registry and what they measure

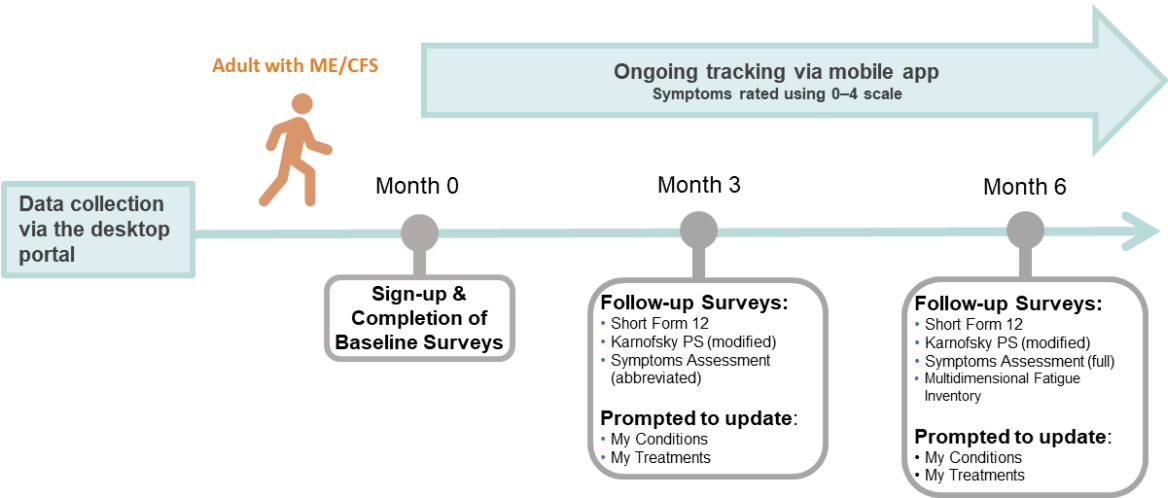
Survey	Measures
ME/CFS Disease History ^d	Disease-specific history (triggers, onset, disease course)
COVID-19 History ^a	Infection status, acute illness, clinical course
UK ME Biobank Symptoms Assessment (31) ^{a, b}	Symptom experience and fulfillment of ME/CFS case definitions
Short Form-36 (SF-36) (32) ^{a, b}	Health-related quality of life
Karnofsky Performance Status (modified) (33) ^a	Functional status
Multidimensional Fatigue Inventory (MFI) (34,35)	General fatigue, physical fatigue, reduced motivation, reduced activity, and mental fatigue

Demographics ^b	Basic demographic information including: age, race, ethnicity, income, education, employment status
My Conditions ^c	Diagnosed conditions
My Treatments ^c	Medications, supplements, and other treatments

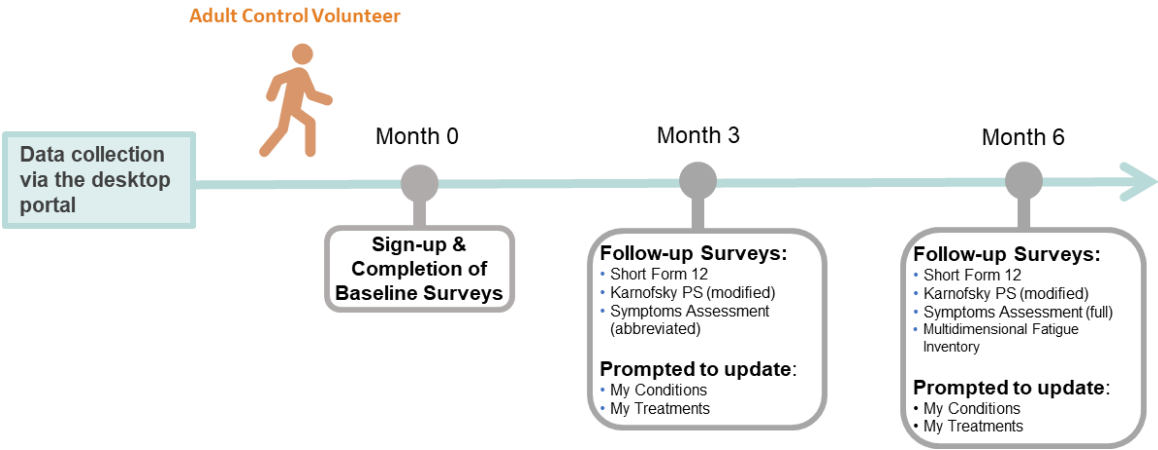
a = Exact same questionnaire asked at follow-up timepoints
b = Abbreviated/modified version of questionnaire asked at follow-up timepoints
c = Form that can be revised/added to on an ongoing basis
d = Survey is only presented to those who indicate they have ME/CFS

Figure 1: Overview of the first 6 months of longitudinal Registry data collection, which includes electronic surveys administered at enrollment (baseline) and follow-up time intervals in: (A) Adults with ME/CFS; (B) Adult control volunteers; and (C) Adults post-COVID (with long-term symptoms or fully recovered). Individuals with ME/CFS or Long COVID can opt to track their symptoms using a numerical scale from 0 (symptom absent) to 4 (very severe) in a mobile app.

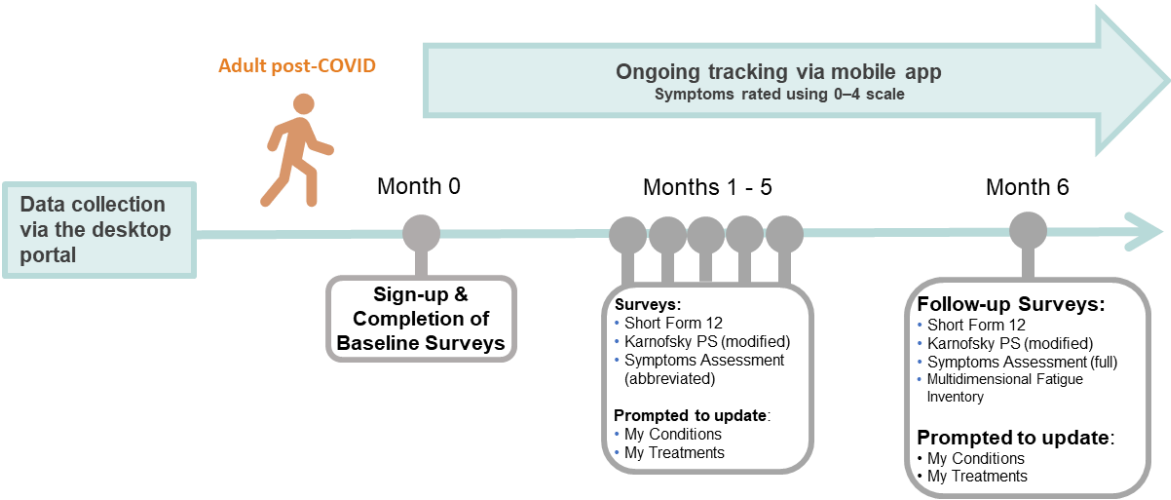
A.



B.



C.

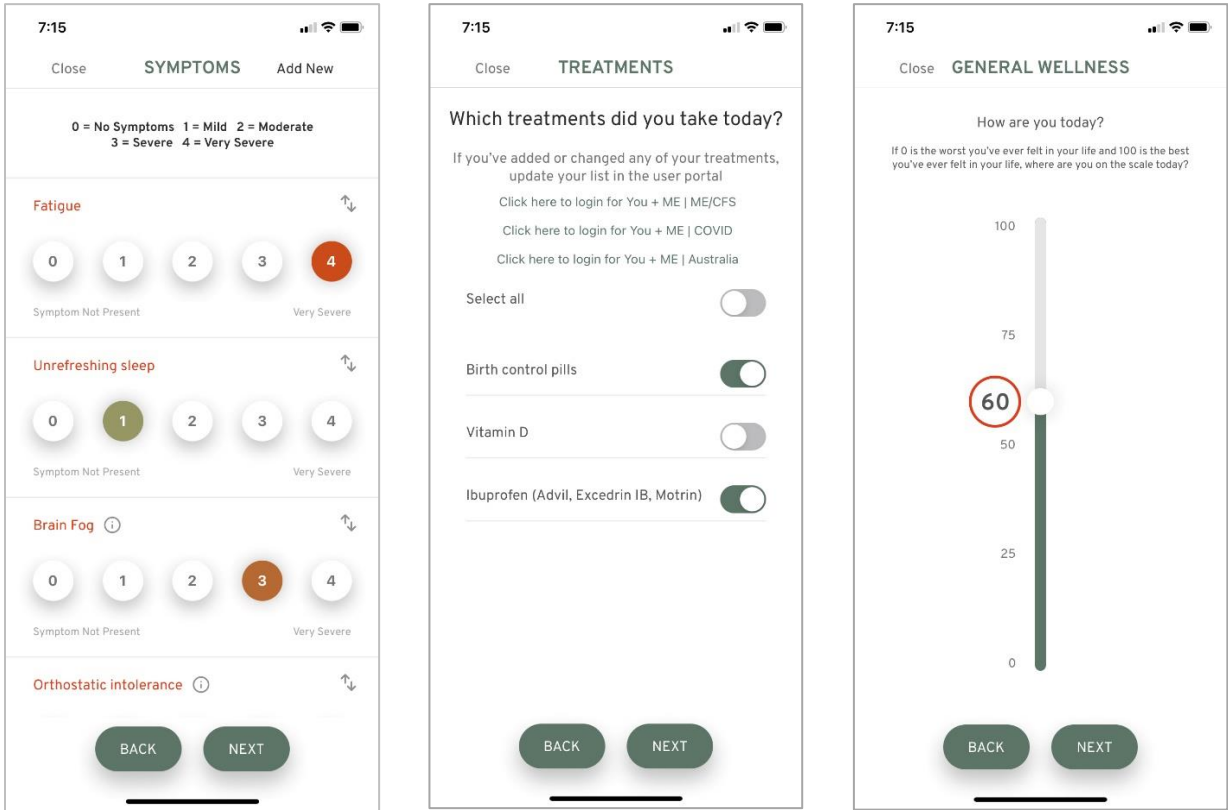


Severity scores are defined as: 0 = symptom not present, 1 = mild, 2 = moderate, 3 = severe, 4 = very severe, according to the DePaul Symptom Questionnaire (36)

Symptom Tracking

Upon completion of the first set of surveys, participants are sent an email with a link to download the You + ME mobile symptom tracking app. The symptom tracking app allows individuals with ME/CFS, LC and other chronic diseases to record symptoms, lifestyle factors, life events, and any activities on an ongoing basis. Because participants have varying levels of disease severity, the frequency that these can be recorded is chosen by the participant and ranges from every day to once a week. Our intention is to capture the data as frequently as possible without overburdening our participants.

Figure 2: Using the Mobile App Tracking Screen, users can: A) Report presence and severity of symptoms felt; B) Log which treatments were taken that day; C) Provide a rating of general wellness



A. B. C.

Data Management

The Principal Investigator is responsible for overseeing data quality and reliability and ensuring safeguards are in place to protect participant privacy and safety. The Registry Data Manager audits the data bi-weekly to improve quality, assure validity and reliability, and guarantee the integrity and credibility of output.

DATA SHARING

Researchers who want to access the data for research apply to Solve M.E. with information required to review their request, including their name, institution, terminal degree, relevant publications, and research interest. The application is reviewed by the You + ME Innovation Council, a group of approximately 12 individuals with deep expertise in ME/CFS and chronic disease research or data science (clinicians, researchers, data scientists, and individuals with ME/CFS and other chronic diseases).

Data is stripped of identifiers and shared with researchers through secure means of transfer. Researchers who use the data are required to sign a Data Use Agreement (DUA) that includes a guarantee to share their methods and findings with the ME/CFS and LC patient communities.

RESULTS

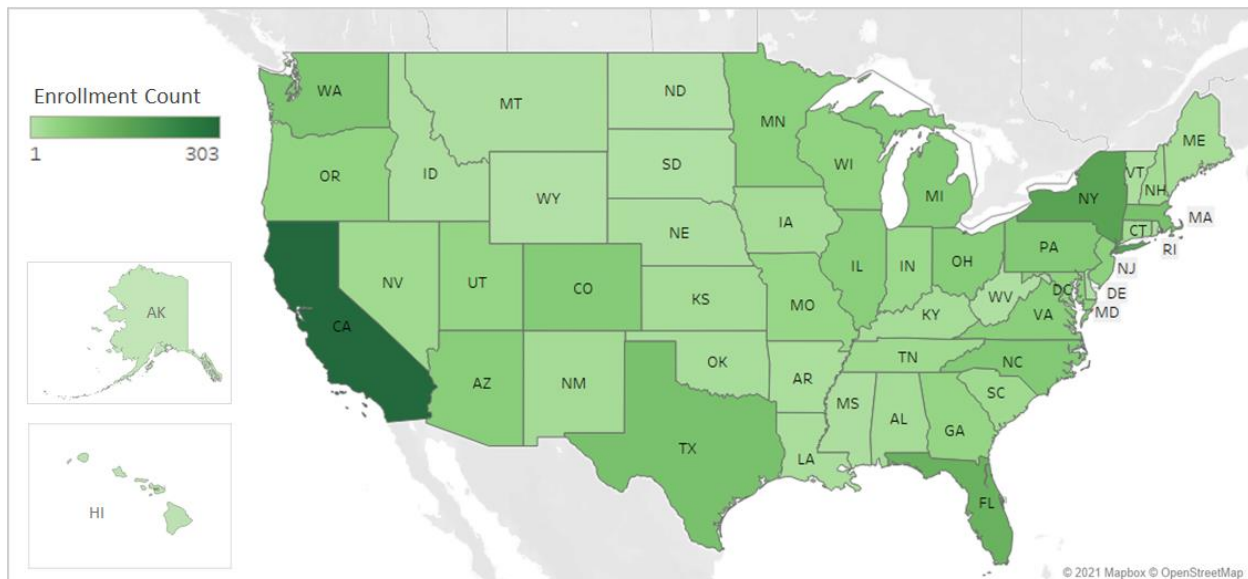
Participant Enrollment

As of September 30, 2021, the Registry has **over 4,200 participants**: 3,033 people with ME/CFS; 833 post-COVID and 473 control volunteers. Recruitment to the Registry continues at a steady pace with **an average of 72 new people registered every week**, responding to social media and other outreach efforts.

The Registry is comprised of a geographically-diverse group, with registrants residing in all 50 states in the United States (see Figures 3A and 3B). The Registry is open to LC registrants internationally. The highest concentration of LC registrants is in the United States (N=651), United Kingdom (N=44), and Canada (N=72); in total, there is representation from 32 countries (see Figure 4).

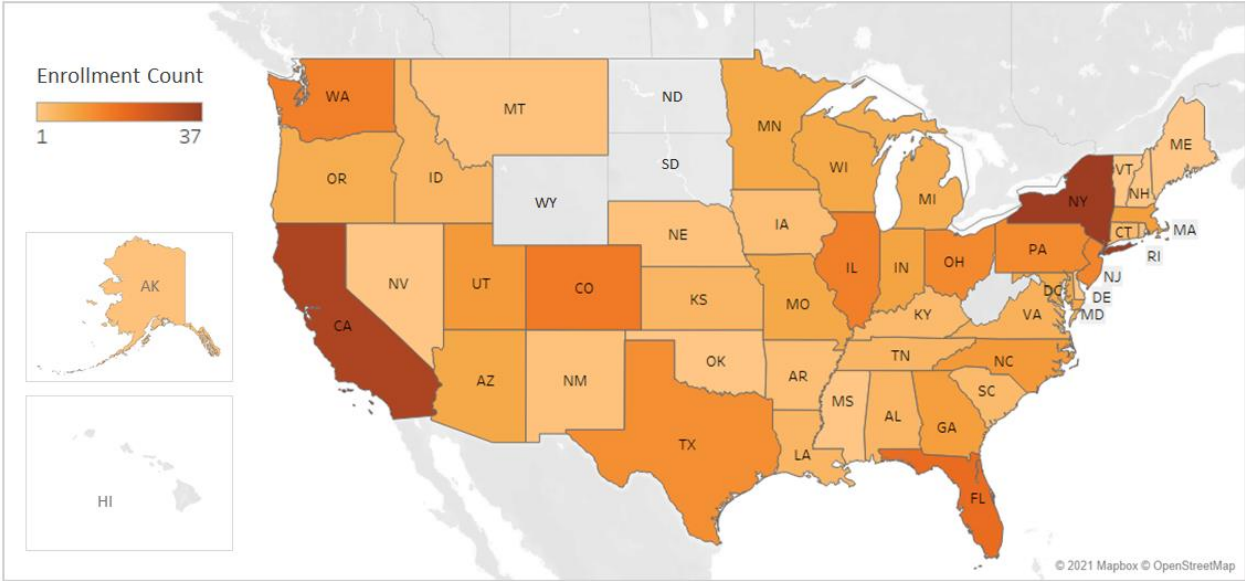
Figure 3: You + ME Registry Enrollment in the U.S. by State (as of September 30, 2021). A: Adults with ME/CFS; B: Adult control volunteers

A.



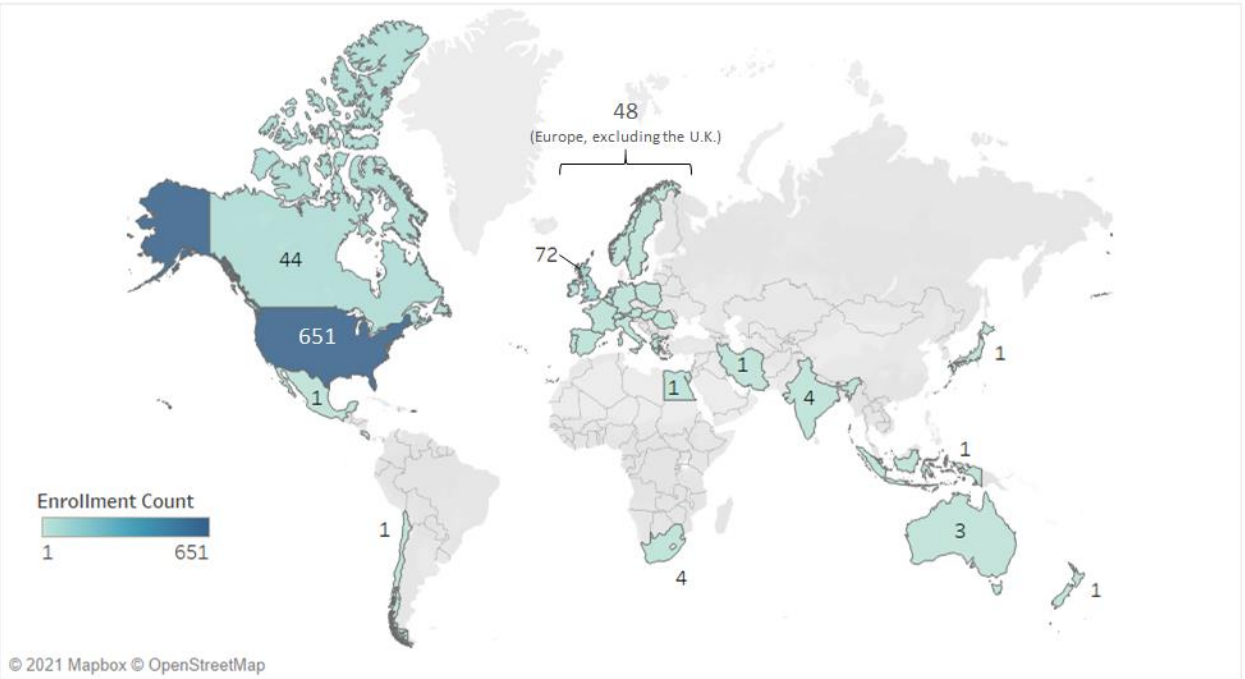
Map based on longitude (generated) and latitude (generated). Each U.S. state with ME/CFS and control volunteers participants enrolled shows color corresponding to enrollment count aggregated from zip code data provided by participants (n = 2085).

B.



Map based on longitude (generated) and latitude (generated). Each U.S. state with participants enrolled shows color corresponding to post-COVID enrollment count aggregated from zip code data provided by participants (n = 387).

Figure 4: You + ME Registry enrollment of adults post-COVID by Country (as of September 30, 2021)



Map based on longitude (generated) and latitude (generated). Each country with participants enrolled shows color corresponding to post-COVID enrollment count from country of residence data provided by participants (n = 836).

Although a major goal of the Registry is to open up participation in research to underrepresented groups, enrollment to date is largely consistent with previous ME/CFS research cohorts. The male to female sex ratio for the entire Registry cohort is 23:100, meaning for roughly every 23 males there are 100 females (intersex individuals make up .001% of participants). Registry participants are predominantly non-Hispanic white. An area of divergence from most previous studies is the significant proportion of individuals in the Registry who are severely to very severely ill, including 33.1% of our ME/CFS cohort and 13.8% of our COVID cohort. These patients, who are house- or bed-bound, are underrepresented in traditional research settings.

Overall enrollment targets by cohort for the first three years of the You + ME Registry are summarized in Table 3. The goal is for 30% of the ME/CFS and COVID cohorts to be controls (in the COVID cohort, controls are individuals who had COVID but fully recovered and never experienced LC). Another goal is for 25% of the entire Registry cohort to be based outside the United States by the end of Year 3.

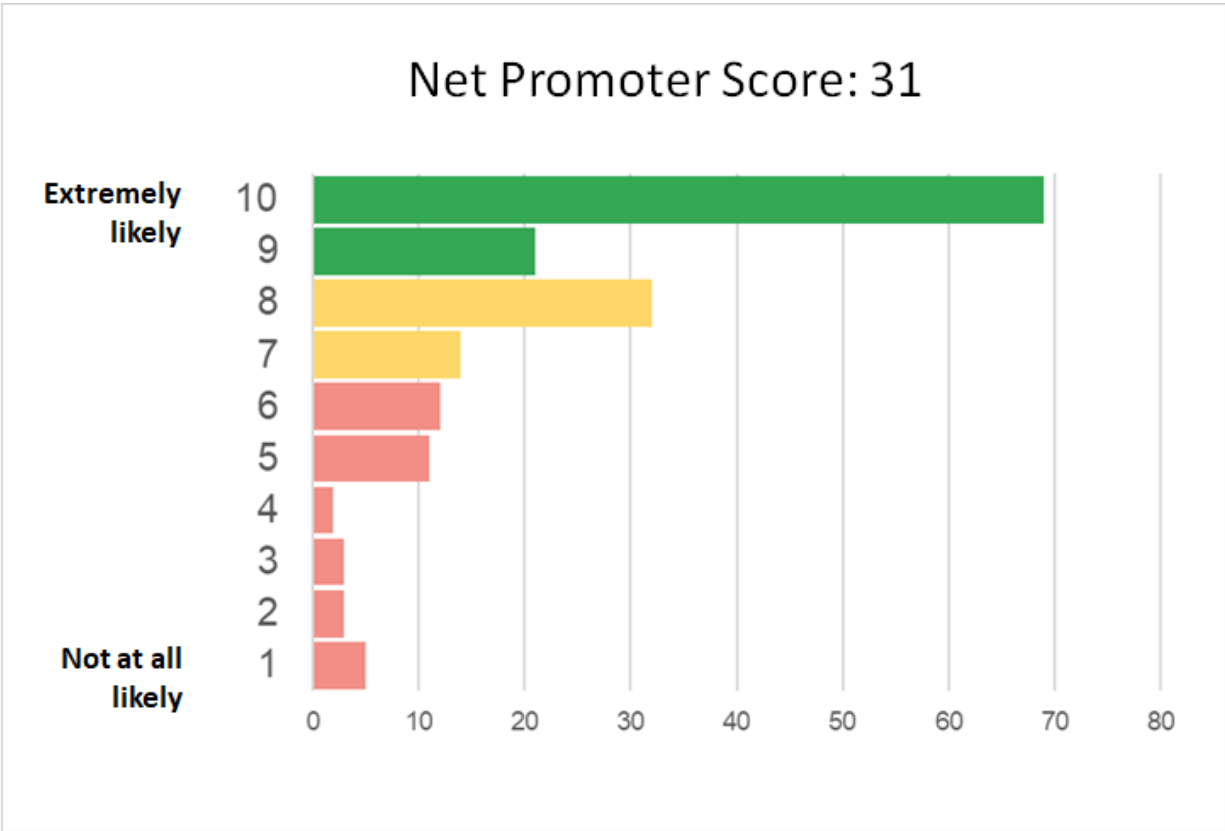
Table 3: You + ME Enrollment targets by cohort for years 1-3

	Year 1	Year 2	Year 3
ME/CFS cohort <i>Incl. control volunteers without ME/CFS</i>	3,000	7,000	10,000
Long COVID cohort <i>Incl. controls (COVID recovered)</i>	800	2,000	5,000
TOTAL	3,800	9,000	15,000

Participant Satisfaction

Of 172 participants who completed our community feedback survey as of September 30, 2021, **71% rated their overall satisfaction with the Registry as a 7 or higher** (on a scale of 1 to 10). The Registry has qualified as “great” using a user satisfaction index measurement called a Net Promoter Score, in which respondents are asked how likely they are to recommend it to a friend (see Figure 5).

Figure 5: You + ME user satisfaction index measured by the Net Promoter Score (as of September 30, 2021)



DISCUSSION

Enrollment has been robust and members of the Registry have reported a high degree of satisfaction with both the data collection process and the digital tools used to capture data. This may be the result of two main factors: 1) the drive of people with ME/CFS and LC to participate in research and fill the gaps in clinical and scientific understanding of what they are experiencing; 2) the creation and ongoing development of the Registry and symptom tracking app with the community’s input. There is a growing body of evidence supporting the benefits of co-creation and human-centered design, particularly as it relates to the development of digital health tools (28,37–39). Facilitating a mechanism for those living with an illness to share the unique insights of their lived experience and partnering with them on the development of tools to capture that experience is a best practice that should be universally adopted to improve our understanding and therapeutic development. The development of the You + ME Registry meets a clear need in the ME/CFS community; participant satisfaction and commitment to the Registry will be key to continued success and further expansion.

Using the Data to Drive Research

The overarching goal of the Registry is to serve as a catalyst for critical research into diagnostics, treatments, and cures for ME/CFS, LC and other post-infection diseases using the power of large cohort with prospectively collected data. Working with the scientific, medical, and pharmaceutical communities, advocating with government agencies, and collaborating with patient groups around the world will lay the foundation for breakthroughs that can improve the lives of millions who suffer from various “long haul” diseases, like ME/CFS and LC.

One of the benefits of this dataset is the ability to look across cohorts to understand the similarities and differences between ME/CFS and LC. The establishment of the Registry during the COVID-19 pandemic, during which case numbers have been highly concentrated over a short period of time, presented an opportunity to track and study trajectories of symptom improvement or worsening in a population with a singular infectious trigger. Clues about susceptibility and resilience to long-term effects of COVID-19 could also benefit the millions of Americans already diagnosed with ME/CFS.

The Registry has started to support data analysis by the internal research team and in partnership with external researchers. One example of active promotion of the Registry is through the Solve M.E. Ramsay Grant Program, an annual peer-reviewed competition for grants in support of pilot studies that first launched in 2016. In 2021, the Ramsay Program opened a new funding mechanism to analyze Registry data, ultimately funding two projects (40). To further increase accessibility and utility of the Registry research dataset, we plan to build an interface that allows vetted researchers to query and use the data for a range of scientific projects. This resource will grow richer over time as more participants enroll and we add new data types, from more expansive digital health data to biological data.

Limitations

Despite these successes, the current Registry study design and dataset has some limitations, including reliance on self-report data, participant drop-off and missing data, a lack of ethnic diversity, and a relatively small number of controls. We are therefore planning a number of actions to improve the representativeness of the Registry cohort and pursuing studies to evaluate our measurements, guide development of the protocol, and examine the quality and usefulness of the data.

Engagement and Data Completeness

Both ME/CFS and LC are illnesses that evolve and change over time. While the Registry and symptom tracking app are specifically designed to capture this, they are dependent on continued engagement from the Registry community, both within visits, and over time. Data completeness, potential bias in completion rates, and data quality will be continuously monitored. Developing novel approaches for engagement and learning from others (37,41) who have successfully achieved this will be an ongoing priority.

Diversifying the You + ME Community

The Registry is predominantly made up of white non-Hispanic individuals. Ethnic minorities are underrepresented as participants in biomedical and public health research, due to a multitude of personal (e.g., cultural distrust and perceptions of research), social (e.g., expense, work and home responsibilities) and research-related (e.g., inaccessibility of study documents and materials, traveling to study locations) factors (42). The use of online surveys that can be completed at home addresses some of these barriers to participation, but there are still reported ethnic and SES differences in web-based research study participation (43). In collaboration with partners, a directed effort will be made to increase Registry inclusivity and participation, and develop strategies to address recruitment bias.

Expanding our Control Cohort

The existing control cohort represents a little more than 10% of overall participants; the target is 30%. To ensure the control cohort is adequately matched on key demographic variables and therefore able to serve as a comparison group to our ME/CFS and LC cohorts, direct, targeted outreach and more innovative approaches, including partnership with other disease Registries to 'share' a control dataset, will be explored.

Next steps for the You + ME Registry

Meeting the Needs of Adolescents

Both ME/CFS and LC affect adolescents; this group is often underrepresented in clinical research (44–46). The symptom clusters experienced by this population are often distinct from the adult population; for example, many adolescents with ME/CFS have orthostatic intolerance as a predominant symptom (47,48).

In early 2022, the Registry will open to adolescents aged 13-17 years of age. Development of this part of the Registry is currently underway and will be designed

specifically for this age group so it includes an appropriate consent and data collection process.

Biosample Collection

To accompany the rich longitudinal phenotypic data collected in the Registry, biological samples will be collected from a subset of the larger Registry cohort to both support specific research projects and to create a biorepository of samples for future research. Samples will include one or more of the following:

1. Dried blood spot (DBS) cards (DNA, RNA, protein expression and metabolomics analyses),
2. Dried urine strips (DUS) (metabolomics),
3. Fecal samples for analyses of microbiome composition and metagenomics (determination of potential microbial metabolites that affect gastrointestinal, immune, metabolic, neurologic and systemic health),
4. Saliva (salivary biomarkers)
5. Venipuncture blood draw and processing and storage of blood components (immunologic, metabolomics, microbiome/virome)

These sample types can support a range of research and make up the immediate biosample collection protocol, but other tissue types, like cerebrospinal fluid, are possible in the future. Samples will be stored by a certified good clinical practice (GCP) provider indefinitely in the Solve ME/CFS Biobank but destroyed upon request if a participant withdraws.

CONCLUSION

This article describes the design and research justification of a patient-powered, longitudinal registry that leverages a dynamic digital platform. This biomedical data resource has the potential to overcome traditional research barriers to understanding the pathophysiology of these heterogeneous conditions, enable the development of diagnostics and treatments, and inform clinical care and improve quality of life for people living with ME/CFS and LC.

The Registry was approved and is overseen by Western IRB under Protocol #20193104 Clinicaltrials.gov identifier: NCT04806620

ACKNOWLEDGEMENTS

Above all, we are incredibly grateful to ME/CFS and Long COVID community members who have continually contributed insights and ideas for the Registry, helped to design the data collection process, spent hours testing the platform and mobile app, and extensive time and energy enrolling and completing surveys.

Special thanks to Rochelle Joslyn (PhD), Rachael Carder, and Beth Mazur for their wisdom and expert advice on numerous aspects of the Registry design and data collection process.

We are deeply appreciative of the scientists and clinicians who provided advice and input, particularly Anthony L. Komaroff (MD), and others in the field who shared their protocols for data collection so we could work to harmonize data - Jarred Younger (PhD), Luis Nacul (MD, PhD), Eliana Lacerda (MD, PhD), Lucinda L. Bateman (MD) and Nancy Klimas (MD), to name a few.

We are incredibly grateful to our partners at Topflight Apps who partnered with us on the development of the Registry platform, with particular thanks to Francesco Mantovani and Alex Lorincz for their dedication and creativity.

Thank you to the UK ME/CFS Biobank and CureME team and the London School of Hygiene and Tropical Medicine for licensing their Symptoms Assessment and scoring algorithm to the Registry.

FUNDING

This work would not be possible without funding from countless individual donors to Solve M.E.

The development of the Registry database and mobile app was supported in part by NIH grant U24-NS-105525 as part of the ME/CFS Collaborative Research Network.

AUTHOR CONTRIBUTIONS

Conceptualization: SW, AR; Data curation: ES, AR; Visualization: KM, AR; Draft of manuscript: KM, AR, SW. All authors approved the final draft of this manuscript.

REFERENCES

1. Chambers DA, Amir E, Saleh RR, Rodin D, Keating NL, Osterman TJ, et al. The

- Impact of Big Data Research on Practice, Policy, and Cancer Care. *Am Soc Clin Oncol Educ B*. 2019 May 17;39:e167–75.
2. Radiological Society of North America. Deep Learning Predicts Woman's Risk for Breast Cancer [Internet]. PR Newswire. 2020 [cited 2021 Oct 27]. Available from: <https://www.prnewswire.com/news-releases/deep-learning-predicts-womans-risk-for-breast-cancer-301180130.html>
 3. Davies N. How Big Data Can Drive Advancements in MS Research [Internet]. Neurology Live. 2021 [cited 2021 Oct 26]. p. 24–7. Available from: <https://www.neurologylive.com/view/big-data-driving-advancements-multiple-sclerosis>
 4. Big Multiple Sclerosis Data Network – The home of MS Real World Evidence! [Internet]. [cited 2021 Nov 9]. Available from: <https://bigmsdata.org/>
 5. Valdez AR, Hancock EE, Adebayo S, Kiernicki DJ, Proskauer D, Attewell JR, et al. Estimating prevalence, demographics, and costs of ME/CFS using large scale medical claims data and machine learning. *Front Pediatr* [Internet]. 2019 Jan [cited 2021 Nov 9];6(412). Available from: <https://pubmed.ncbi.nlm.nih.gov/30671425/>
 6. Jason LA, Mirin AA. Updating the National Academy of Medicine ME/CFS prevalence and economic impact figures to account for population growth and inflation. *Fatigue Biomed Heal Behav* [Internet]. 2021 [cited 2021 Nov 23];9(1):9–13. Available from: <https://sci-hub.st/10.1080/21641846.2021.1878716>
 7. Rivera MC, Mastronardi C, Silva-Aldana CT, Arcos-Burgos M, Lidbury BA. Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Comprehensive Review. *Diagnostics (Basel, Switzerland)* [Internet]. 2019 Aug 7 [cited 2021 Nov 9];9(3). Available from: <https://pubmed.ncbi.nlm.nih.gov/31394725/>
 8. Collin SM, Bakken IJ, Nazareth I, Crawley E, White PD. Trends in the incidence of chronic fatigue syndrome and fibromyalgia in the UK, 2001–2013: a Clinical Practice Research Datalink study. *J R Soc Med* [Internet]. 2017 Jun 30;110(6):231–44. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28358988>
 9. Ramsay AM, Myalgic Encephalomyelitis Association. Myalgic encephalomyelitis and postviral fatigue states : the saga of Royal Free Disease. Gower Medical Pub. for the Myalgic Encephalomyelitis Association; 1988.
 10. Siegel SD, Antoni MH, Fletcher MA, Maher K, Segota MC, Klimas N. Impaired natural immunity, cognitive dysfunction, and physical symptoms in patients with chronic fatigue syndrome: preliminary evidence for a subgroup? *J Psychosom Res* [Internet]. 2006 Jun [cited 2021 Oct 26];60(6):559–66. Available from: <https://pubmed.ncbi.nlm.nih.gov/16731230/>
 11. Bansal AS, Bradley AS, Bishop KN, Kiani-Alikhan S, Ford B. Chronic fatigue syndrome, the immune system and viral infection. *Brain Behav Immun* [Internet]. 2012 Jan [cited 2021 Oct 26];26(1):24–31. Available from: <https://pubmed.ncbi.nlm.nih.gov/21756995/>
 12. Cliff JM, King EC, Lee J-S, Sepúlveda N, Wolf A-S, Kingdon C, et al. Cellular Immune Function in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *Front Immunol* | www.frontiersin.org [Internet]. 2019 [cited 2019 Jun

- 6];1:796. Available from: www.frontiersin.org
13. Sweetman E, Noble A, Edgar C, Mackay A, Helliwell A, Vallings R, et al. Current Research Provides Insight into the Biological Basis and Diagnostic Potential for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *Diagnostics* [Internet]. 2019 Sep [cited 2021 Oct 26];9(3). Available from: [/pmc/articles/PMC6787691/](https://pubmed.ncbi.nlm.nih.gov/3487691/)
14. Helliwell AM, Sweetman EC, Stockwell PA, Edgar CD, Chatterjee A, Tate WP. Hypermethylation and decreased expression of TMEM240 are potential early-onset biomarkers for colorectal cancer detection, poor prognosis, and early recurrence prediction. *Clin Epigenet* [Internet]. 2020 [cited 2021 Oct 26];12(167):167. Available from: <https://doi.org/10.1186/s13148-020-00960-z>
15. Cairns R, Hotopf M. A systematic review describing the prognosis of chronic fatigue syndrome. *Occup Med (Lond)* [Internet]. 2005 Jan [cited 2021 Nov 9];55(1):20–31. Available from: <https://pubmed.ncbi.nlm.nih.gov/15699087/>
16. Strayer DR, Young D, Mitchell WM. Effect of disease duration in a randomized Phase III trial of rintatolimod, an immune modulator for Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *PLoS One* [Internet]. 2020 Oct 29 [cited 2021 Oct 26];15(10). Available from: [/pmc/articles/PMC7595369/](https://pubmed.ncbi.nlm.nih.gov/3395369/)
17. Fluge Ø, Rekeland IG, Lien K, Thürmer H, Borchgrevink PC, Schäfer C, et al. B-Lymphocyte Depletion in Patients With Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: A Randomized, Double-Blind, Placebo-Controlled Trial. *Ann Intern Med* [Internet]. 2019 [cited 2021 Oct 26];170(9):585–93. Available from: <https://pubmed.ncbi.nlm.nih.gov/30934066/>
18. Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness [Internet]. Vol. 2016. Institute of Medicine of the National Academies; 2015. Available from: <https://www.nap.edu/html/19012/MECFSc cliniciansguide.pdf>
19. Zhou Z, Kang H, Li S, Zhao X. Understanding the neurotropic characteristics of SARS-CoV-2: from neurological manifestations of COVID-19 to potential neurotropic mechanisms. *J Neurol* [Internet]. 2020 Aug [cited 2021 Nov 9];267(8):2179–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/32458193/>
20. Román GC, Spencer PS, Reis J, Buguet A, Faris MEA, Katrak SM, et al. The neurology of COVID-19 revisited: A proposal from the Environmental Neurology Specialty Group of the World Federation of Neurology to implement international neurological registries. *J Neurol Sci* [Internet]. 2020 Jul 15 [cited 2021 Nov 9];414(116884). Available from: <https://pubmed.ncbi.nlm.nih.gov/32464367/>
21. Moldofsky H, Patcai J. Chronic widespread musculoskeletal pain, fatigue, depression and disordered sleep in chronic post-SARS syndrome; a case-controlled study. *BMC Neurol*. 2011 Mar;11:37.
22. Collins FS. NIH launches new initiative to study “Long COVID” [Internet]. NIH. 2021 [cited 2021 Sep 10]. Available from: <https://www.nih.gov/about-nih/who-we-are/nih-director/statements/nih-launches-new-initiative-study-long-covid>
23. Ladds E, Rushforth A, Wieringa S, Taylor S, Rayner C, Husain L, et al. Persistent symptoms after Covid-19: qualitative study of 114 “long Covid” patients and draft quality principles for services. *BMC Health Serv Res* [Internet]. 2020 Dec 20 [cited 2021 Nov 9];20(1144). Available from:

- <https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-020-06001-y>
24. Goërtz YMJ, Van Herck M, Delbressine JM, Vaes AW, Meys R, Machado FVC, et al. Persistent symptoms 3 months after a SARS-CoV-2 infection: the post-COVID-19 syndrome? *ERJ Open Res.* 2020 Oct;6(4):00542–2020.
 25. Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing Long COVID in an International Cohort: 7 Months of Symptoms and Their Impact. *medRxiv [Internet]*. 2020 Dec 24 [cited 2021 Nov 9]; Available from: <https://www.medrxiv.org/content/10.1101/2020.12.24.20248802v2>
 26. Trueland J. Long COVID – we've been here before [Internet]. *BMA.* 2021 [cited 2021 Nov 9]. Available from: <https://www.bma.org.uk/news-and-opinion/long-covid-we-ve-been-here-before>
 27. Topol EJ, Verghese A, Fauci AS. Fauci to Medscape: "We're All in It Together and We're Gonna Get Through It" [Internet]. *Medscape.* 2020 [cited 2021 Oct 11]. Available from: https://www.medscape.com/viewarticle/933619#vp_2
 28. Holeman I, Kane D. Human-centered design for global health equity. *Inf Technol Dev [Internet]*. 2019 Jul 2 [cited 2021 Oct 27];26(3):477–505. Available from: <https://www.tandfonline.com/doi/abs/10.1080/02681102.2019.1667289>
 29. NINDS Common Data Elements. Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Standards - NINDS Common Data Elements [Internet]. 2018 [cited 2018 Nov 14]. Available from: https://www.commondataelements.ninds.nih.gov/MECFs.aspx#tab=Data_Standards
 30. Lupton D. Social media, activism, trucker caps: the fascinating story behind long COVID [Internet]. *The Conversation.* 2021 [cited 2021 Oct 27]. Available from: <https://theconversation.com/social-media-activism-trucker-caps-the-fascinating-story-behind-long-covid-168465>
 31. Lacerda EM, Bowman EW, Cliff JM, Kingdon CC, King EC, Lee JS, et al. The UK ME/CFS Biobank for biomedical research on Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) and Multiple Sclerosis. *Open J Bioresour [Internet]*. 2017;4. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28649428>
 32. Ware J, Sherbourne C. User's manual for the SF-35v2 health survey second edition. Lincoln, RI: QualityMetric Incorporated; 2007.
 33. Ferrucci L, Koh C, Bandinelli S, Guralnik JM. Disability, Functional Status, and Activities of Daily Living. *Encyclopedia of Gerontology.* 2007. p. 476–436.
 34. Shahid A, Wilkinson K, Marcu S, Shapiro CM. STOP, THAT and One Hundred Other Sleep Scales. In *Springer Science+Business Media, LLC* ; 2012.
 35. Smets EMA, Garssen B, Bonke B, De Haes JCJM. The multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res.* 1995;39(3):315–25.
 36. Sunnquist M, Lazarus S, Jason LA. The Development of a Short Form of the DePaul Symptom Questionnaire. *Rehabil Psychol [Internet]*. 2019 Jul 18 [cited 2021 Nov 9];64(4). Available from: [/pmc/articles/PMC6803042/](https://pmc/articles/PMC6803042/)
 37. Fagotto E. Engaging Patients for Research that Matters: IBD Partners [Internet].

- Ash Center Harvard Kennedy School. 2019. Available from: <https://ash.harvard.edu/publications/engaging-patients-research-matters-ibd-partners>
38. Fleisher L, Bass SB, Shwarz M, Washington A, Nicholson A, Alhajji M, et al. Using theory and user-centered design in digital health: The development of the mychoice communication tool to prepare patients and improve informed decision making regarding clinical trial participation. *Psychooncology* [Internet]. 2019 Oct 26 [cited 2021 Oct 27];29(1):114–22. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/pon.5254>
 39. Azeez-Narain K. Can Human-Centered Design Help Create Better Digital Health Solutions? [Internet]. *Health Tech.* [cited 2021 Oct 27]. Available from: <https://www.healthtechmagazines.com/can-human-centered-design-help-create-better-digital-health-solutions/>
 40. Solve ME/CFS Initiative. About the Ramsay Grant Program [Internet]. [cited 2021 Sep 20]. Available from: <https://solvecfs.org/about-the-ramsay-grant-program/>
 41. Smolensky L, Amondikar N, Crawford K, Neu S, Kopil CM, Daeschler M, et al. Fox Insight collects online, longitudinal patient-reported outcomes and genetic data on Parkinson's disease. *Sci data* [Internet]. 2020 Feb 24 [cited 2021 Oct 27];7(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/32094335/>
 42. Yancey AK, Ortega AN, Kumanyika SK. Effective recruitment and retention of minority research participants. *Annu Rev Public Health* [Internet]. 2006 [cited 2021 Nov 9];27:1–28. Available from: <https://pubmed.ncbi.nlm.nih.gov/16533107/>
 43. Jang M, Vorderstrasse A. Socioeconomic Status and Racial or Ethnic Differences in Participation: Web-Based Survey. *JMIR Res Protoc* [Internet]. 2019 Apr [cited 2021 Nov 19];8(4):e11865. Available from: <https://pmc/articles/PMC6479282/>
 44. Grady C, Wiener L, Abdoler E, Trauernicht E, Zadeh S, Diekema DS, et al. Assent in research: the voices of adolescents. *J Adolesc Heal* [Internet]. 2014 May 1 [cited 2021 Oct 27];54(5):515. Available from: <https://pmc/articles/PMC4047975/>
 45. Brown DL, Cowdery JE, Jones TS, Langford A, Gammage C, Jacobs TL. Adolescent knowledge and attitudes related to clinical trials. *Clin trials* [Internet]. 2015 Jun [cited 2021 Oct 27];12(3):212–4. Available from: <https://pubmed.ncbi.nlm.nih.gov/25673637/>
 46. American College of Obstetricians and Gynecologists. Guidelines for Adolescent Health Research - Committee Opinion No 665 [Internet]. *Obstet Gynecol.* 2016 [cited 2021 Nov 9]. p. 183–6. Available from: <https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2016/06/guidelines-for-adolescent-health-research>
 47. Jason LA, Benton MC, Valentine L, Johnson A, Torres-Harding S. The Economic impact of ME/CFS: Individual and societal costs. *Dyn Med* [Internet]. 2008;7(6). Available from: <http://www.dynamic-med.com/content/pdf/1476-5918-7-6.pdf>
 48. Jason LA, Katz BZ, Sunnquist M, Torres C, Cotler J, Bhatia S. The Prevalence of Pediatric Myalgic Encephalomyelitis/Chronic Fatigue Syndrome in a Community-Based Sample. *Child Youth Care Forum* [Internet]. 2020 Aug [cited 2021 Oct 27];49(4):579. Available from: <https://pmc/articles/PMC8186295/>

