Bijlage figuren en tabellen ME/CFS Lines

Table 1: Problems of current ME/CFS research and proposed solutions in ME/CFS Lines

Limitations of current ME/CFS research	Our solution
Lack of researchers focusing on ME/CFS ME/CFS is a diagnosis that has not been taken very seriously in the past. This has resulted in a lack of funding and a resulting lack of researchers focusing on this problem.	We will educate a **new generation of early stage researchers** from various fields. We will do so by building a **multidisciplinary consortium** of experienced researchers with relevant expertise for the development of new hypotheses and approaches to study ME/CFS, and early stage researchers leading specific projects.
Lack of data because of limited inclusion of ME/CFS in large scale biobanks Low public awareness and the limited number of researchers working in the field has acted as a barrier to include ME/CFS information in large scale biobanks, hampering etiological research. Current studies consequently have been performed with small sample numbers resulting in insufficient power. This is particularly problematic given the presumed heterogeneity of ME/CFS. In addition, previous studies typically compared ME/CFS patients to controls differing on many other aspects than the disease, such as lifestyle or geographical area.	Lifelines, a national resource for health research, is the **only large population cohort and biobank** that **includes diagnostic criteria for ME/CFS** since 2014. We will leverage this exceptionally large cohort of 167,000 individuals for ME/CFS research and build a national longitudinal in-depth cohort study and biobank to enable mechanistic studies into ME/CFS. Even in the current patient set, we have **2,500 individuals meeting the diagnostic criteria of ME/CFS** and this number will increase over the project runtime of 8 years. To the best of our knowledge this cohort size exceeds state of the art studies by an order of magnitude, and thus will allow to **unravel disease heterogeneity**. We will develop a protocol/criteria to implement in other cohorts, via consortia that we have connections to (CHARGE-consortium) to accelerate research into ME/CFS. In addition, Lifelines includes a large group of controls from the same geographical area with extensive data on health and lifestyle, which can be used for matching or adjustments in analyses to reduce environmental impact.
Reverse causality Virtually all previous studies on the aetiology of ME/CFS have compared patients with healthy controls, and were not able to distinguish causal factors from associations. Many proposed biomarkers could also be explained by differences in sleep or activity patterns associated with ME/CFS and may not necessarily play a causal role.	As our Lifelines cohort has included ME/CFS diagnostic criteria since 2014, we have so far identified **400 new onset ME/CFS cases**. For these individuals **blood samples are available before and after diagnosis**. This cohort is worldwide unique enabling to **predict new onsets** from samples collected before disease outbreak. Thereby, we can avoid reverse causality and also compare differences pre/post on an individual level.
Lack of gold standard diagnosis In the absence of objectively assessable diagnostic features, ME/CFS is diagnosed based on the presence of specific self-reported symptoms, while molecular markers are missing. Current diagnostic criteria are based on expert-opinion instead of scientific arguments and there is no consensus on which criteria identify patients best. Current ME/CFS studies are thus based on patient groups that may remarkably differ in symptom type and severity. This heterogeneity largely contributes to the inconsistent results of studies aiming to find biomarkers, or studies seeking efficient treatment options.	We will include **multiple diagnostic criteria** sets by adding the DePaul Symptom Questionnaire 2.0 for all participants fulfilling the CDC criteria in the 2014-2018 data wave (N=2,500), and a random sample of other participants (N=2,500). By profiling ME/CFS patients with an array of multi-omics methods (genetics, proteomics, microbiome, antibody repertoires, metabolomics), we will perform a holistic, integrated search for **molecular ME/CFS biomarkers**, representing the largest scale effort to date. Ultimately, the data of the molecular markers will be compared to the state of the art criteria from questionnaires to investigate subtypes of ME/CFS and understand disease heterogeneity.
Lack of robustness and replication Several factors potentially involved in ME/CFS (such as the involvement of the immune system, microbiome, or metabolism) have been assessed in various small cohorts of ME/CFS patients. But in nearly all cases replication of these findings in other cohorts (or even larger patient numbers) is lacking	We have various strategies to enable replication studies: (1) Our cohort of **2,500 individuals meeting ME/CFS criteria and 400 new onset cases**, has been collected over a time span of currently 8 years over a wide geographic area. Thereby we can test for temporal effects as well as biases from local collection, providing **unprecedented replication;** (2) **Harmonisation** of diagnostic criteria and instruments with other cohorts (such as funded under this call or the UK ME/CFS Biobank); (3) **Replication of findings in other layers of data** (e.g., replicate genetic findings at protein level)
Lack of coordination and integration of results Every research group uses its own sample and own methods, and produces its own results without taking into account mechanisms suggested by researchers from other disciplines.	We will build a research resource with several layers of **multi-omics data in a uniformly recruited group of patients** to allow translation of results from one discipline and perspective to another level. Therefore, the consortium will include a coordinating senior researcher to integrate the findings into an integrated view on the aetiology of ME/CFS.

Table 2: ME/CFS patient cohort and biobank

	Method	Sample	Power considerations
Diagnostic criteria	DePaul Symptom Questionnaire (DSQ-2)	2,500 participants fulfilling CDC criteria and 2,500 controls from wave 3, repeated in wave 4	Not applicable. ME/CFS Lines be the first population-based patient cohort in which the four widely used diagnostic criteria sets and their overlap will be assessed.
Genetics	Illumina global screening array (GSA) Beadchip-24 v1.0	All participants fulfilling CDC criteria in wave 2 (2,500 cases) or wave 3 (currently 400 extra); genetic data already available in N=1,400 cases and 48,000 controls	Genetic data will be collected in all participants fulfilling the CDC criteria; genetic analyses will be performed on several ME/CFS cohorts combined.
Proteomics	Olink Explore 3072 panels	800 cases and 800 matched controls in wave 3, 100 severe cases outside Lifelines	Olinks' power calculation indicates a minimum sample size of 595 cases and 595 controls to detect an effect size of 0.3 (α=0.05; power=0.8).
Microbiome	Metagenomics sequencing (MGS)	800 cases in wave 3 and 100 severe cases outside Lifelines, control data already available in N=8,000.	The current evidence of association between ME/CFS is based on the studies with a sample size for ME/CFS patient group smaller or equal than 100 patients and similar number of controls. Given the expected design of 900 cases vs 8,000 controls, we are able to detect from 4.1 to 4.8 times smaller effect sizes (in terms of log odds ratios) of ME/CFS on quantitative microbial markers (Wald statistics, considering the population prevalence of ME/CFS).
Antibody repertoire	PhIP-Seq with a library of 344,000 microbial and viral antigens	800 cases at wave 3, including 400 new onset cases that will also be assessed pre-disease at wave 2, 400 matched controls, and 100 severe cases outside Lifelines	The necessary cohort sizes will depend on the strength of the signal and abundance of shared diagnostic antibody responses. From pilot studies on 40 ME/CFS cases and 40 healthy controls, we could already detect statistically significant differences, hence increasing the cohort size 20x should additionally increase power.
Metabolome	Metabolon Inc.	800 cases and 800 matched controls in wave 3, 100 severe cases outside Lifelines	A sample size of at least 600 people is required to be representative of the entire sample population (Dunn et al, Metabolomics 11, 9–26 (2015).

Table 3: Expertise consortium members

Member	Google scholar link	Expertise
Prof. Rosmalen (UMCG)	https://scholar.google.nl/citations?user=mJJu06wAAAAJ&hl	Epidemiology, Immunology
Prof. Bakker (UMCG)	https://scholar.google.com/citations?hl=nl&user=9PIuH7gAAAAJ	Medical Systems Biology, Metabolomics
Dr. Boer (ErasmusMC)	https://scholar.google.com/citations?user=IT1MURgAAAAJ&hl=en	Population/Functional genomics, proteomics
Prof. Gans (UMCG)	https://scholar.google.com/citations?hl=nl&user=q_dugKcAAAAJ	Internal Medicine, medical education
Prof. Kendler (Virg Commonw Univ)	https://scholar.google.com/citations?hl=nl&user=PKXVXTgAAAAJ	Genetic Epidemiology, Nosology
Dr. Kurilshikov (UMCG)	https://scholar.google.com/citations?hl=nl&user=F47NF4wAAAAJ	Microbiome
Prof. van Meurs (ErasmusMC)	https://scholar.google.com/citations?hl=nl&user=ssvwassAAAAJ	Population/Functional genomics, proteomics
Dr. van Ockenburg (UMCG)	https://scholar.google.com/citations?hl=nl&user=tpEizMMAAAAJ	Endocrinology, Metabolic disorders
Prof. Reinders (Delft Univ)	https://scholar.google.com/citations?hl=nl&user=h52_bg0AAAAJ	Bioinformatics, pattern recognition
Dr. Vogl (Cleveland Clinic)	https://scholar.google.at/citations?user=jua0UDgAAAAJ&hl=eng	Systems Immunology, Data science
Prof. Zhernakova (UMCG)	https://scholar.google.at/citations?user=jua0UDgAAAAJ&hl=eng	Microbiome, Data science
Dr. Zijdewind (UMCG)	https://scholar.google.com/citations?hl=nl&user=uV728XcAAAAJ	Neurosciences, physiology

Figure 1: The general design of ME/CFS Lines