EPIGENOMICS
2026, VOL. 18, NO. 4, 411–419
https://doi.org/10.1080/17501911.2026.2656361

Taylor & Francis
by informa

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Longitudinal study of genome-wide DNA methylation in individuals with and without post-acute symptoms following SARS-CoV-2 infection

Jon Bohlin [a,b], Yunsung Lee [b], Ida Henriette Caspersen [b], Anna Hayman Robertson[a], Christian M. Page[c], Håkon K. Gjessing[b,d], Astanand Jugessur[b,d], Per Magnus [b], Siri Mjaaland [a] and Lill Trogstad [a]

[a]Department of Method Development and Analytics, Norwegian Institute of Public Health, Oslo, Norway; [b]Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway; [c]Department of Physical Health and Ageing, Division of Public Health and Prevention, Norwegian Institute of Public Health, Oslo, Norway; [d]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

ABSTRACT

Background: Symptoms following SARS-CoV-2 infection, referred to as Long-COVID, have been reported since the pandemic. We investigated whether COVID-19 or Long-COVID is associated with persistent genome-wide DNA methylation (DNAm) changes in whole blood using a longitudinal design.
Methods: DNAm was measured using the Illumina EPIC V2 platform (859,651 CpGs) in 297 adult participants (594 samples in total) from two Norwegian population-based cohorts, with samples collected pre-infection (2020) and during the pandemic (2023). Participants were classified as Long-COVID, COVID-19 (no persistent symptoms), or not infected.
Results: No significant DNAm differences were observed between Long-COVID and not infected at either time point ($p = 0.745$(FDR)) or during the pandemic specifically ($p = 0.629$(FDR)). Likewise, no differences were detected between COVID-19 and not infected across both time points ($p = 0.883$(FDR)) or during the pandemic ($p = 0.287$(FDR)). Sex-stratified analyses of the X chromosome revealed no significant DNAm differences for Long-COVID or COVID-19 in males (both $p = 0.999$(FDR)) or females (both $p = 0.999$(FDR)). Epigenetic age acceleration was also evaluated using DunedinPACE (DP) and PhenoAge (PA), but no significant differences were detected for Long-COVID ($p = 0.695$ [DP], $p = 0.528$ [PA]) or COVID-19 ($p = 0.624$ [DP], $p = 0.348$ [PA]).
Conclusion: No persistent epigenetic age- or DNAm based differences due to Long-COVID or SARS-CoV-2 infection were detected in our cohorts.

PLAIN LANGUAGE SUMMARY

Some people continue to experience symptoms for months after a COVID-19 infection. If symptoms persist for more than three months, the condition is often diagnosed as Long-COVID. Scientists are still trying to understand why some people develop long-lasting symptoms while others recover fully. One possible explanation is that COVID-19 might cause lasting changes in how genes are regulated in the body. In this study, we looked at a biological process called DNA methylation that helps control how genes are turned on or off. Changes in this process can sometimes reflect long-term effects of illness or aging. We analyzed blood samples from 297 adults in Norway. Each person gave a sample before the COVID-19 pandemic and another sample during the pandemic resulting in 594 samples in total. We compared people who had Long-COVID, people who had COVID-19 but recovered without long-term symptoms, and people who did not report being infected. We did not find any persistent clear differences in DNA methylation both within- or between these groups. We also looked at measures of biological aging based on DNA methylation patterns but found no differences related to COVID-19 or Long-COVID. Hence, we did not detect long-term changes in DNA methylation in blood that could be attributed to COVID-19 and Long-COVID in our cohorts. Larger studies, employing higher resolution DNA methylation methodology, may be needed to determine whether smaller or more specific changes occur.

ARTICLE HISTORY

Received 5 January 2026
Accepted 2 April 2026

KEYWORDS

Host DNA methylation; epigenetics; EWAS; SARS-CoV-2 infection; Long-COVID; epigenetic age acceleration; X chromosome; longitudinal design

1. Introduction

A substantial proportion of individuals infected with SARS-CoV-2 report persistent post-acute symptoms lasting months or even years after the initial infection [1]. When symptoms continue for more than three months, the condition is referred to as Long-COVID or post-acute sequelae of SARS-CoV-2 infection (PASC) [2]. Multiple factors have been proposed to contribute to the development of Long-COVID, including host genetic predisposition [3,4], preexisting comorbidities [5], psychological stress, immune dysregulation, and age-related

CONTACT Jon Bohlin jon.bohlin@fhi.no Department of Method Development and Analytics, Norwegian Institute of Public Health, Skøyen, Oslo N-0213, Norway

Supplemental data for this article can be accessed online at https://doi.org/10.1080/17501911.2026.2656361

Article highlights

- We examined whether COVID-19 or Long-COVID leads to lasting changes in DNA methylation, an epigenetic mechanism that may be linked to more persistent change in gene activity.
- Using a longitudinal design with matched samples, collected before and after infection, we analyzed genome-wide DNA methylation in 297 adults.
- We found no evidence of genome-wide DNA methylation differences associated with COVID-19 infection or persistent Long-COVID symptoms.
- DNA methylation analyses of the X chromosome, including regions near the ACE2 gene used by SARS-CoV-2 to enter cells, also showed no detectable changes.
- Measures of biological aging based on DNA methylation (DunedinPACE and PhenoAge clocks) were not affected by COVID-19 or Long-COVID status.

susceptibility [6,7]. Alterations in the upper respiratory tract microbiome have also been suggested as a contributing factor in certain Long-COVID phenotypes [8]. The persistence of symptoms in some individuals may reflect underlying long-term physiological changes [9], potentially mediated by epigenetic mechanisms [10]. In particular, the link between DNA methylation (DNAm) and gene expression is well established and may be altered in response to viral infection, leading to sustained transcriptional changes [11]. Furthermore, it has been hypothesized that viral infections may induce demethylation of host genomic regions or of integrated viral elements, potentially contributing to chronic symptoms [12]. Several recent studies have reported genome-wide DNAm alterations in individuals following SARS-CoV-2 infection and in those experiencing Long-COVID [13–16].

With access to ongoing longitudinal cohorts containing COVID-19/Long-COVID specific information, we aimed to investigate whether long-lasting symptoms from Long-COVID, as well as COVID-19 infections, are associated with persistent DNAm changes in the host. Whole blood samples and detailed COVID-related health information were collected from 297 adult participants in two Norwegian population-based studies: the Norwegian Mother, Father and Child Cohort Study (MoBa) [17] and the Norwegian Influenza Study (NorFlu) cohort [18]. In this population, we previously observed that Long-COVID symptoms, after SARS-CoV-2 Omicron variant infection, included fatigue, poor memory, brain fog, dizziness, heart palpitations, shortness of breath, reduced lung function, and altered smell or taste, based on questionnaires administered 3–5 months after a primary infection in January-February 2022 [19]. Using the same design, we now investigate DNAm changes in whole-blood after COVID-19, also considering persistence of any of the eight above-mentioned Long-COVID symptoms. Blood samples were obtained at two time points for each individual: prior to SARS-CoV-2 infection and vaccination (April-December 2020) and during the later stages of the pandemic (April 2023). Hence, our longitudinal study consisted of 594 samples in total.

The availability of COVID-19 infection status and self-reported Long-COVID symptoms allowed us to directly assess within-individual DNAm changes pre- and post-infection. This longitudinal design enabled comparisons both within participants (before vs. after infection), between groups (Long-COVID, COVID-19, and not infected controls) as well as within participants and between groups. By including individuals who had never reported a COVID-19 infection (i.e., "not infected"), we were able to assess differences between symptomatic and asymptomatic groups, both for each individual (i.e., before and after infection) and between groups. We also conducted targeted analyses of CpG sites linked to the *ACE2* gene, located on the X chromosome, which encodes the cell surface receptor used by the SARS-CoV-2 virus for host-cell entry [20]. Finally, we examined potential health-span implications of Long-COVID and COVID-19 using the DunedinPACE [21] and PhenoAge [22] epigenetic health-span clocks.

2. Materials and methods

2.1. Study population and design

We used two population-based cohorts, MoBa and NorFlu sampled from the Norwegian population, with regular questionnaire data collection and data linkages to national health registries. MoBa recruited pregnant women and their partners from all over Norway during 1999 to 2008 [17,23]. Norflu was initiated during the A H1N1(pdm09) influenza pandemic in 2009/2010 [24] where pregnant women were invited to participate (https://www.fhi.no/en/id/studies/norflu/). Since March/April 2020, adult participants in MoBa (mothers and fathers) and NorFlu (mothers) have been invited to answer electronic questionnaires concerning Long-COVID symptoms and COVID-19 testing. A subsample of participants in MoBa and NorFlu were additionally asked to donate biological samples at multiple time points prior to and during the pandemic. The cohorts were linked to the National Surveillance System for Communicable Diseases (MSIS), the Norwegian Immunisation Registry (SYSVAK), and the Medical Birth Registry of Norway via each participant's unique national identification number.

Eligible participants in this study included adult members in the two cohorts with available whole blood samples (the NorFlu cohort also included peripheral blood mononuclear cells) collected prior to SARS-CoV-2 infection, in 2020 (MoBa) and 2020/21 or 2014/15 (NorFlu). For all exposure groups, NorFlu participants and a randomly selected subset of MoBa participants were invited to provide a second blood sample in April 2023.

2.2. Exposure

We defined three exposure groups based on the participant's SARS-CoV-2 infection and Long-COVID status: 1) Long-COVID; 2) only COVID-19, no reported Long-COVID symptoms; 3) no reported COVID-19 infection (designated as "not infected"). Participants in the COVID-19 strata (group 2) had tested positive for SARS-CoV-2 infection in January or February 2022. This period largely overlapped with the first Omicron surge in Norway. A positive test result in this period was defined as having a record of laboratory confirmed SARS-CoV-2 infection (MSIS), or any self-reported positive test for SARS-CoV-2 from questionnaires, either a self-administrated antigen test or a laboratory test (PCR or antigenic). The Long-COVID strata (group 1) included participants with COVID-19 according to the criteria given above who after 3–5 months post infection (June 2022) still reported one of the following symptoms:

fatigue, poor memory, brain fog, dizziness, heart palpitations, shortness of breath, reduced lung function, and altered smell and/or taste. These symptoms were selected because of their association with prior SARS-CoV-2 Omicron infection [19]. The last, uninfected strata (group 3) included participants with no reported SARS-CoV-2 infection by January 2023. These participants had no record of a positive SARS-CoV-2 test in MSIS by January 2023, no self-reported infection in January or February 2022, and responded "No" to the question "Have you ever been infected with SARS-CoV-2?" in January 2023. Supplementary Figure S1 shows the timeline of vaccination and infection (both MSIS-registered and self-reported) in the Norwegian population based on data from the MoBa cohort (91,360 participants).

2.3. Covariates

We included information about sex from the participants' cohort member status as mother or father. Age at blood sampling was calculated based on birth year, obtained from the Medical Birth Registry of Norway. Information about body mass index (BMI, calculated from height and weight) and smoking status around the time of blood sampling were obtained from questionnaires administered to cohort participants before (2013, for NorFlu only) and during the pandemic (2020–2022).

2.4. Biological materials

Whole blood samples were collected from all 297 participants in 10 ml EDTA tubes at two timepoints. DNA extraction was performed using the FlexiGene DNA Kit (250) (QIAGEN, Hilden, Germany) for most samples. For NorFlu samples collected in 2014/15 a ArchivePure DNA Blood 1000 ml basekit was used (5 PRIME, Inc.). All samples were subsequently stored at −20°C. For each participant, approximately 20 µl of TE-diluted DNA was aliquoted into skirted 96-well plates, with a minimum concentration of 60 ng/µl and an absorbance ratio (260/280) between 1.8 and 2.0. All DNA samples, from both cohorts, were collectively shipped on dry ice to Life & Brain GmbH (Bonn, Germany) for methylation measurement using the Illumina Infinium EPIC v2 BeadChip (Illumina, San Diego, CA, USA).

2.5. Preparation of DNA methylation array data

Since the NorFlu cohort consisted exclusively of women, quality control (QC) was performed separately for the MoBa and NorFlu cohorts using the RnBeads (v2.0) R package [25]. Of the 923,798 probes on the EPIC v2 array, the 'greedycut' algorithm excluded 19,142 probes in MoBa, resulting in 904,656 retained sites (881,456 autosomal, 23,062 X-chromosomal, and 138 Y-chromosomal). For NorFlu, 6,842 probes were excluded, along with 161 Y-chromosomal probes. This yielded 916,795 retained probes (893,116 autosomal and 23,679 X-chromosomal).

Sample-level QC in MoBa began with 531 samples from 269 participants (262 with two measurements, one for each time point, and 7 with one). One sample was excluded due to being the only sample on its plate; another was removed by the greedycut algorithm, resulting in 260 participants with paired measurements and 9 with single measurements.

Data preprocessing included 'noob' background correction using SeSAMe [26], specifically adapted to the EPIC V2 platform, followed by BMIQ normalization to adjust for Illumina Type I/II CpG probe bias [27].

2.6. Estimation of cell-type composition

Cell-type proportions were estimated separately for the MoBa and NorFlu cohorts using the IDOL method [28] in combination with the FlowSorted.Blood.EPIC reference panel [29]. Estimated proportions included $CD4^+$ helper T cells (CD4T), $CD8^+$ cytotoxic T cells (CD8T), $CD19^+$ B cells (Bcell), $CD14^+$ monocytes (Mono), $CD56^+$ natural killer cells (NK), and neutrophils (Neu).

2.7. Statistical analyses

Linear mixed-effects models were used to assess DNAm differences between, as well as within (for each individual), the Long-COVID, COVID-19, and not infected groups. Models included cross-classified individual- and plate-level random intercepts to account for within-subject correlation across two time points, pre- vs. during pandemic, including potential MoBa and NorFlu cohort DNAm differences (not detected). Both main effects and interaction terms between infection status (Long-COVID, COVID-19 or not infected) and time point (pre- vs. during pandemic) were evaluated.

Covariates included in all models were sampling date, sex, BMI, age, smoking status (never/ever), and estimated cell-type proportions. Analyses were also conducted separately for the X chromosome due to the location of the *ACE2* gene, with sex-stratified models used to account for DNAm differences due to X-inactivation in females [30].

Health-span analyses were performed using the DunedinPACE [21] and PhenoAge [22] epigenetic clocks as outcome variables, with models adjusted using the same covariates as described above. Significance was evaluated using false discovery rate (FDR) correction [31], with FDR <0.01 as the primary threshold.

To verify the integrity of DNAm data, additional models assessed known associations with smoking status and BMI. These included DNAm patterns at established smoking-associated CpG sites (see Supplementary Methods) [32], as well as associations with both epigenetic clocks, which were consistent with prior literature.

Visual diagnostics included volcano and Q–Q plots for all models. All analyses were conducted in R, using the nlme package [33] for linear mixed-effects models.

Because age was missing for several individuals due to strict privacy regulations in Norway, DNAm-based age was estimated using MinLinMo [34], a 16-CpG clock with a reported correlation of $r = 0.94$ with chronological age. This estimator does not incorporate health-span components, making it suitable for unbiased age estimation in the current context.

Further methodological details, including full statistical model specifications, are provided in the Supplementary Methods.

3. Results

It should be noted that only the results from what we consider to be the primary regression models are presented here. Additional model specifications, as well as further methodological details and extended results, are provided in the Supplementary Methods (see also Supplementary Figure S1) for readers seeking a more in-depth view.

We first assessed genome-wide DNAm across 859,651 CpG sites using mixed-effects models, adjusting for estimated cell-type proportions (CD4+ T cells, CD8+ T cells, NK cells, B cells, monocytes, and granulocytes), as well as age, sampling date, sex, smoking status (never/ever), and BMI (see Table 1). The longitudinal design incorporated two time points per participant: one prior to SARS-CoV-2 infection and widespread vaccination (April-December 2020) and one during the later stages of the pandemic (April 2023). To account for repeated measures and technical variation we applied mixed-effects models with nested random intercepts for individual (two time points: pre- vs. during pandemic) and array plate. While multiple alternative model specifications were explored (see Supplementary Methods for full details), the models presented below reflect the approach we judged to be the most biologically and statistically appropriate for the research question.

We tested for persistent DNAm differences between Long-COVID and COVID-19 groups, compared with not infected participants, at both time points. Interaction terms between group status (Long-COVID or COVID-19) and time point (pre- vs. during pandemic) were included to assess within-individual and between group average DNAm changes over time.

No significant genome-wide differences were detected either within or between groups: Long-COVID vs. not infected ($P_{FDR} = 0.745$), or COVID-19 vs. not infected ($P_{FDR} = 0.883$). These null findings suggest that infection status, as confirmed by positive PCR or antigen self-test, did not significantly influence genome-wide DNAm profiles (see Figures 1–3). Our study could also indicate that vaccination against COVID did not

Table 1. Study population statistics.

Group	1 (Long-COVID)	2 (COVID only)	3 (Not infected)
Time point 1: During COVID, *n*	58	110	129
Average age, years	48.7	49.2	50.9
Brain fog, *n*	19	0	2
Fatigue, *n*	19	0	3
Number of vaccine doses, *n* (%)			
0–1	9 (15.5%)	12 (10.9%)	11 (8.5%)
2	9 (15.5%)	19 (17.3%)	16 (12.4%)
3+	40 (69.0%)	79 (71.8%)	102 (79.1%)
Time point 0: Pre COVID infection, *n*	58	110	129
Average age, years	46.4	46.5	48.6
Pooled			
Sample size, *n*	116	220	258
Females, %	72.40%	65.50%	59.70%
Smoking (yes), %	7%	3%	7%
Average BMI, kg/m^2	25.00	24.50	25.40

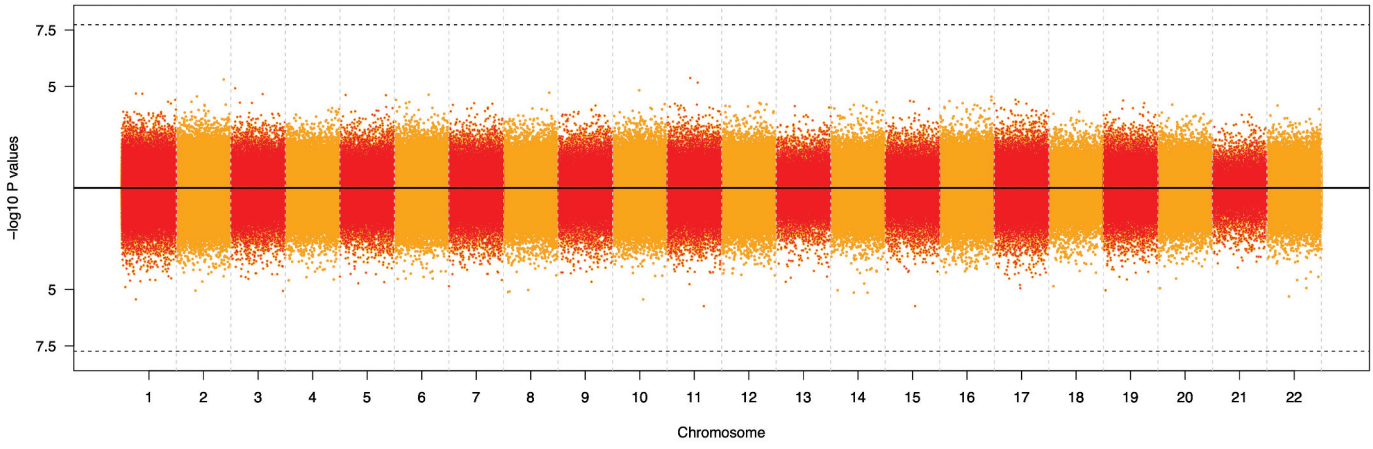

Figure 1. Miami plot of main longitudinal regression models for Long-COVID/COVID. The top part of the plot, above the thick black line in the middle, shows the -$\log_{10}$ transformed p values (vertical axis) from the longitudinal Long-COVID regression models assessing DNAm differences with regards to non-COVID infected at 859,651 CpGs. The bottom part of the plot shows the corresponding longitudinal regression models for COVID infected, which also includes Long-COVID cases, versus non-infected. The horizontal axis represents genomic position with regards to each chromosome, excluding the X chromosome. The horizontal dashed black line at the top and bottom of the Miami plot is the Bonferroni limit for $p < 0.05$ conditioned on 859,651 tests.

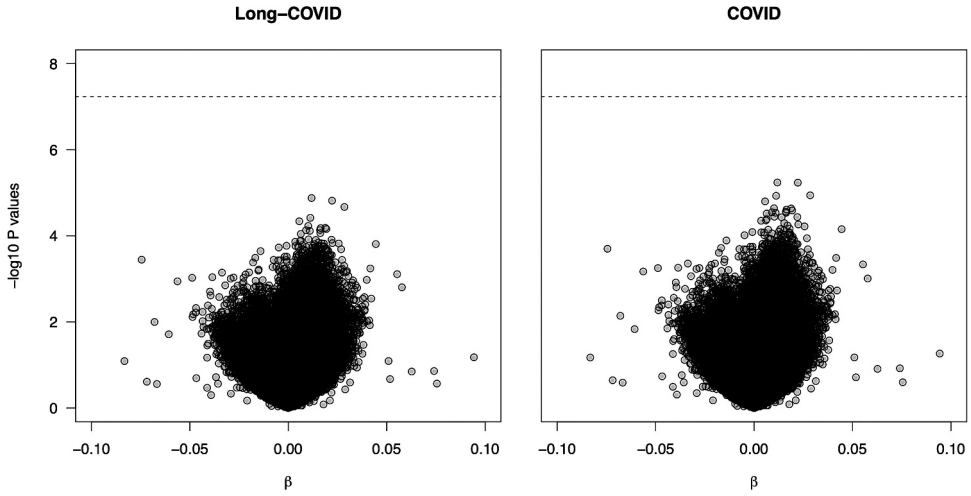

Figure 2. Volcano plots for longitudinal Long-COVID and COVID regression models. The volcano plots show -$\log_{10}$ transformed p values (vertical axis) plotted against the estimated slopes (β, horizontal axis) for the respective main- longitudinal Long-COVID (left) and COVID (right) regression models. The dashed line at the top designates the Bonferroni limit for $p < 0.05$ conditioned on 859,651 tests.

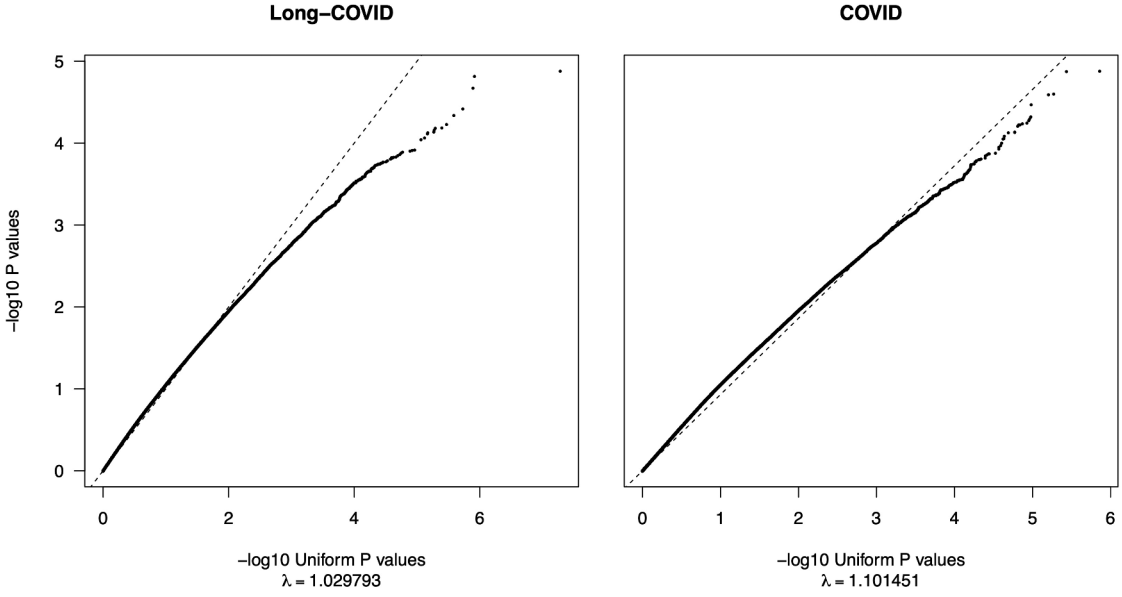

Figure 3. QQ-plots for longitudinal Long-COVID and COVID regression models. The QQ-plots show -$\log_{10}$ transformed p values (vertical axis) for the main longitudinal Long-COVID (left) and COVID (right) regression models plotted against correspondingly transformed values between 0 and 1 drawn from a uniform distribution (horizontal axis). A slope (λ) approaching 1 designates absence of bias.

have any genome-wide effect on DNAm as well as most participants across all three groups were vaccinated at the time of second sampling [35].

To refine the Long-COVID phenotype, we conducted targeted analyses focusing on individuals reporting fatigue and brain fog symptoms [36], reducing the Long-COVID group to 19 participants for each time point (see Table 1). Despite the more specific phenotype definition, there were no significant DNAm differences for Long-COVID with fatigue ($P_{FDR} = 0.554$) or brain fog ($P_{FDR} = 0.999$) when compared to the not infected group. These models included the same covariate adjustments described above (see Methods section for details).

We further examined DNAm changes on the X chromosome, focusing on CpG sites linked to the *ACE2* gene, and stratified analyses by sex. Again, no significant DNAm differences were found between Long-COVID and not infected individuals for either males or females ($P_{FDR} = 0.999$ for both). Similarly, no differences were observed for the COVID-19 group ($P_{FDR} = 0.883$ for males; $P_{FDR} = 0.999$ for females) (see Figure 4).

Lastly, we tested whether infection status was associated with changes in biological aging, using the DunedinPACE (DP) and PhenoAge (PA) epigenetic clocks. Longitudinal models with the same covariates were used, including group-by-time interaction terms. Again, no significant associations were detected for Long-COVID ($p = 0.695$ [DP], $p = 0.528$ [PA]) or COVID-19 (including

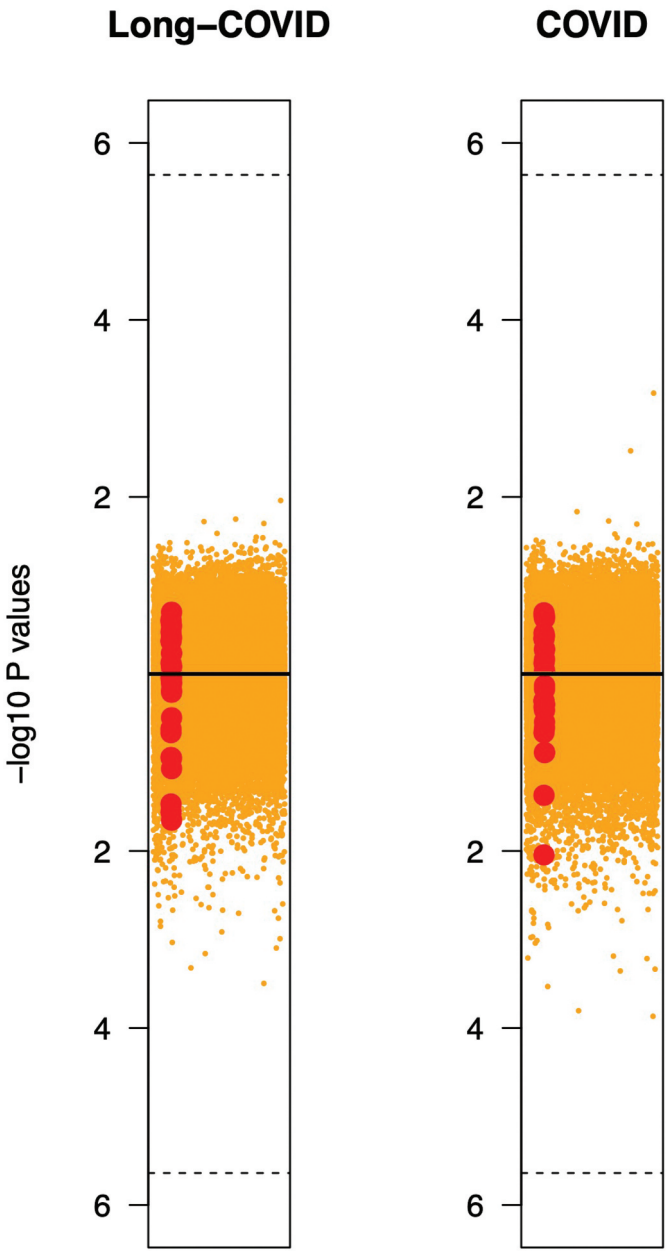

Figure 4. Miami plots of main longitudinal regression models for Long-COVID /COVID on X chromosomes. The Miami plots show the -$\log_{10}$ transformed p values (vertical axis) from the Long-COVID (left) and COVID (right) regression models assessing DNAm differences with regards to non-COVID infected at 21,761 CpGs for males (top) and females (bottom). The horizontal axis represents genomic position with regards to the X chromosome, and the red dots highlight the position of the CpG sites linked to the *ACE2* gene that the SARS-CoV-2 virus employs for host-cell entry. The horizontal dashed black line at the top and bottom of both plots is the Bonferroni limit for $p < 0.05$ conditioned on 21,761 tests.

Long-COVID cases) ($p = 0.624$ [DP], $p = 0.348$ [PA]) when compared to not infected participants (see Supplementary Methods for details).

4. Discussion

A key motivation for examining whether COVID-19 and Long-COVID are associated with sustained genome-wide DNA methylation (DNAm) changes in white blood cells is the known role of the immune system in adapting to pathogens [10,37]. As immunological memory can persist beyond the acute phase of infection, it is plausible that epigenetic mechanisms, including DNAm alterations, may underlie some of these long-term effects [10,37]. Given that DNAm is both persistent and reversible, reflecting transient or sustained immune responses, epigenetic mechanisms could potentially contribute to post-acute symptoms, such as those observed in Long-COVID [38].

To our knowledge, this is one of the few epigenetic studies to analyze persistent genome-wide DNAm in longitudinally matched samples across COVID-19 and Long-COVID phenotypes, including targeted assessment of the contributions of genes on the X chromosome. Several previous studies have suggested associations between SARS-CoV-2 infection and altered DNAm profiles, including in individuals with Long-COVID. However, these studies have often been limited by small sample sizes, broad Long-COVID definitions, and cross-sectional designs [39]. In contrast, our study benefits from a longitudinal design with matched pre- and post-infection samples, a relatively large sample size, and the ability to classify individuals using robust definitions of both COVID-19 and Long-COVID. Additionally, we conducted sex-stratified analyses on the X chromosome, targeting CpG sites linked to the *ACE2* gene, encoding the SARS-CoV-2 entry receptor [20], a feature rarely included in previous DNAm studies on COVID-19.

Despite a comprehensive set of analyses, including more than 40 mixed-effects models, we found no significant genome-wide DNAm differences between individuals with Long-COVID, COVID-19, and not infected controls. This included both repeated measures and between-group comparisons across one or two time points. We also tested whether infection status influenced biological aging using two epigenetic clocks: DunedinPACE and PhenoAge. No associations were observed between either clock and infection status, suggesting that neither COVID-19 nor Long-COVID has a detectable sustained effect on epigenetic aging in our cohorts, at least as captured by these established metrics.

Importantly, both epigenetic clocks were significantly associated with known predictors of biological aging, such as BMI and smoking status, indicating that the DNAm data were sufficiently sensitive and of adequate quality [21,22]. Additional smoking-related EWAS models also showed expected DNAm patterns [32] (see Supplementary Methods), reinforcing the notion that the absence of findings related to COVID-19 is unlikely the consequence of technical limitations.

Although we did not observe DNAm alterations associated with Long-COVID or COVID-19, this does not preclude the possibility that such effects may exist. The Illumina EPIC V2 array interrogates only a small fraction of the estimated 28 million CpG sites in the human genome [40], and while many of these are in regulatory regions, it is estimated that up to 80% of CpGs are constitutively methylated and may not have any regulatory relevance [41]. Moreover, the methylome is highly correlated within individuals [42], meaning that subtle or localized changes could be missed in genome-

wide analyses. Since biological samples were collected approximately one year after the last time point it is not impossible that DNAm changes can have reverted, in particular for the COVID-19 infected group. Most SARS-CoV-2 infections in Norway occurred after the emergence of the Omicron variant, when public health restrictions were relaxed and routine PCR testing was largely discontinued. As a result, infection status for some participants relied on self-testing and self-report, which may have led to some misclassification of infection status in the COVID-19 and "not infected" groups. In addition, there was a broad, predominantly mRNA, vaccine coverage in Norway (see Table 1 and Supplementary Figure S1) which could potentially also influence results. Although no DNAm differences were detected between time point 0, when no one were vaccinated, and time point 1, vaccination reducing effects from Long-COVID/COVID 19 infection cannot be ruled out. Finally, it is also important to note that our cohorts are ethnically homogeneous, consisting largely of individuals of Northern European ancestry, as epigenetic responses to infection could potentially differ across populations with different genetic and environmental backgrounds.

Further complicating interpretation is the heterogeneity of Long-COVID itself [39]. The condition encompasses a broad range of symptoms, some of which may have distinct etiologies such as fatigue, poor memory, brain fog, dizziness, heart palpitations, shortness of breath, reduced lung function and altered taste/smell, making it difficult to capture a unified biological signal [19,43]. While we attempted to refine the Long-COVID definition by focusing on participants reporting fatigue and brain fog [36], no significant DNAm differences were identified even in this more narrowly defined subgroup. This highlights the need for further work to disaggregate Long-COVID into more biologically coherent subtypes, ideally with larger samples for each phenotype.

5. Conclusions

In this study, we comprehensively investigated whether Long-COVID or COVID-19 infection is associated with persistent genome-wide DNA methylation (DNAm) changes in whole blood from 297 cohort participants collected and matched at two time points (totaling 594 samples): before and after SARS-CoV-2 infection and vaccination. Across more than 850,000 CpG sites, including targeted sites near the *ACE2* receptor gene on the X chromosome, we found no evidence of DNAm alterations associated with infection status in our cohorts. Similarly, no associations were detected between infection and measures of biological aging, as estimated by the DunedinPACE and PhenoAge epigenetic clocks.

Although our sample size was larger than those in prior studies, it may still have been insufficient to detect subtle DNAm changes, particularly given the likely heterogeneity of Long-COVID phenotypes. Refining the case definition to individuals reporting specific symptoms such as brain fog or fatigue did not yield significant findings. Additionally, the Illumina EPIC V2 array, while offering broad coverage, interrogates only a small subset of the estimated 28 million CpG sites in the human genome. As such, we cannot exclude the possibility that infection-related DNAm changes exist outside the regions covered by the array.

Taken together, our findings do not support the presence of persistent, detectable DNAm alterations associated with Long-COVID or COVID-19 infection in our cohorts. Further research using larger and more diverse populations, as well as higher-resolution methylation profiling, as opposed to the array-based EPIC V2 platform employed here, may be required to detect more subtle and phenotype-specific effects.

Acknowledgments

The Norwegian Mother, Father, and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. We are grateful to all the participants in Norway who take part in the ongoing cohorts at NIPH.

Although the manuscript is entirely written by the first author, with additional edits from the coauthors, ChatGPT version 4o was occasionally used for help with English- translation, grammar and improving certain formulations.

Author contributions

CRediT: **Jon Bohlin:** Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing; **Yunsung Lee:** Data curation, Formal analysis, Methodology, Writing – review & editing; **Ida Henriette Caspersen:** Data curation, Formal analysis, Investigation, Writing – review & editing; **Anna Hayman Robertson:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – review & editing; **Christian M. Page:** Formal analysis, Methodology, Writing – review & editing; **Håkon K. Gjessing:** Formal analysis, Methodology, Writing – review & editing; **Astanand Jugessur:** Data curation, Writing – review & editing; **Per Magnus:** Conceptualization, Data curation, Funding acquisition, Project administration, Writing – review & editing; **Siri Mjaaland:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Writing – review & editing; **Lill Trogstad:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Writing – review & editing.

Disclosure statement

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Funding

This work was funded by the Norwegian Research Council's Centers of Excellence Funding Scheme, grant no. [262700] and by the Norwegian Institute of Public Health (NIPH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Ethical declaration

The study has been approved by The Regional Committee for Medical and Health Research Ethics, Southeast Norway (no. 127708 for MoBa and no. 18403 for NorFlu). Written informed consent was obtained to participate. In the NorFlu cohort, mothers consented on behalf of their children. The establishment of MoBa and initial data collection was based on a license from the Norwegian Data Protection Agency and approval from The Regional Committees for Medical and Health Research Ethics. MoBa is now based on regulations related to the Norwegian Health Registry Act. The 2023 blood sample collection and DNAm analysis plan was additionally approved in umbrella project with REK no 229359.

Data availability statement

Access to the genetic datasets can be obtained by applying to the Norwegian Institute of Public Health (NIPH; http://www.fhi.no/en/). Restrictions may apply regarding the availability of these data, which were originally used under specific approvals for the current study and are therefore not publicly available. Access can only be given after approval by REK under the provision that the applications are consistent with the consent provided. For MoBa, an application form can be found on the NIPH website at https://www.fhi.no/en/studies/moba/. Queries on access to genetic data from NorFlu or other specific questions regarding access to data in this study can be directed to Dr. Lill Trogstad (lill.trogstad@fhi.no).

ORCID

Jon Bohlin http://orcid.org/0000-0002-0992-1311
Yunsung Lee http://orcid.org/0000-0002-2512-1082
Ida Henriette Caspersen http://orcid.org/0000-0003-2591-8435
Per Magnus http://orcid.org/0000-0002-6427-4735
Siri Mjaaland http://orcid.org/0000-0002-3870-5861
Lill Trogstad http://orcid.org/0000-0002-9557-5725

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Bartsch SM, Chin KL, Strych U, et al. The current and future burden of Long COVID in the United States (U.S.). J Infect Dis. 2025;231(6):1581–1590. doi: 10.1093/infdis/jiaf030
2. Goldowitz I, Worku T, Brown L, editors. A Long COVID definition: a chronic, systemic disease state with profound consequences. National Academies Press (US); 2024. Available from: http://www.ncbi.nlm.nih.gov/books/NBK605676/
• **Describes the definition of Long-COVID.**
3. Pairo-Castineira E, Rawlik K, Bretherick AD, et al. GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. Nature. 2023;617(7962):764–768. doi: 10.1038/s41586-023-06034-3
4. Lammi V, Nakanishi T, Jones SE, et al. Genome-wide association study of long COVID. Nat Genet. 2025;57(6):1402–1417. doi: 10.1038/s41588-025-02100-w
• **A GWAS study reporting significant findings of SNPs associated with Long-COVID.**
5. Durstenfeld MS, Peluso MJ, Peyser ND, et al. Factors associated with long COVID symptoms in an online cohort study. Open Forum Infect Dis. 2023;10(2):ofad047. doi: 10.1093/ofid/ofad047
6. Hermans LE, Wasserman S, Xu L, et al. Long COVID prevalence and risk factors in adults residing in middle- and high-income countries: secondary analysis of the multinational anti-coronavirus therapies (ACT) trials. BMJ Glob Health. 2025;10(4):e017126. doi: 10.1136/bmjgh-2024-017126
7. Davis HE, McCorkell L, Vogel JM, et al. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol. 2023;21(3):133–146. doi: 10.1038/s41579-022-00846-2
8. Xie L, Luo G, Yang Z, et al. The clinical outcome of COVID-19 is strongly associated with microbiome dynamics in the upper respiratory tract. J Infect. 2024;88(3):106118. doi: 10.1016/j.jinf.2024.01.017
9. Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. Nat Med. 2022;28(7):1461–1467. doi: 10.1038/s41591-022-01840-0
10. Cheong J-G, Ravishankar A, Sharma S, et al. Epigenetic memory of coronavirus infection in innate immune cells and their progenitors. Cell. 2023;186(18):3882–3902.e24. doi: 10.1016/j.cell.2023.07.019
11. Milavetz BI, Balakrishnan L. Viral epigenetics. Methods Mol Biol Clifton NJ. 2015;1238:569–596. doi: 10.1007/978-1-4939-1804-1_30
12. Bernal KDE, Whitehurst CB. Incidence of Epstein-Barr virus reactivation is elevated in COVID-19 patients. Virus Res. 2023;334:199157. doi: 10.1016/j.virusres.2023.199157
13. Lee Y, Riskedal E, Kalleberg KT, et al. EWAS of post-COVID-19 patients shows methylation differences in the immune-response associated gene, IFI44L, three months after COVID-19 infection. Sci Rep. 2022;12(1):11478. doi: 10.1038/s41598-022-15467-1
14. Nikesjö F, Sayyab S, Karlsson L, et al. Defining post-acute COVID-19 syndrome (PACS) by an epigenetic biosignature in peripheral blood mononuclear cells. Clin Epigenet. 2022;14(1):172. doi: 10.1186/s13148-022-01398-1
15. Calzari L, Dragani DF, Zanotti L, et al. Epigenetic patterns, accelerated biological aging, and enhanced epigenetic drift detected 6 months following COVID-19 infection: insights from a genome-wide DNA methylation study. Clin Epigenet. 2024;16(1):112. doi: 10.1186/s13148-024-01724-9
16. Balnis J, Madrid A, Drake LA, et al. Blood DNA methylation in post-acute sequelae of COVID-19 (PASC): a prospective cohort study. eBioMedicine. 2024;106:105251. doi: 10.1016/j.ebiom.2024.105251
17. Magnus P, Birke C, Vejrup K, et al. Cohort profile update: the norwegian mother and child cohort study (MoBa). Int J Epidemiol. 2016;45(2):382–388. doi: 10.1093/ije/dyw029
18. Caspersen IH, Trogstad L, Galanti MR, et al. Current tobacco use and SARS-CoV-2 infection in two Norwegian population-based cohorts. BMC Public Health. 2023;23:846. doi: 10.1186/s12889-023-15822-5
• **Describes the NorFlu cohort in greater detail.**
19. Caspersen IH, Skodvin SN, Blix K, et al. Post-COVID symptoms after SARS-CoV-2 omicron infection and the effect of booster vaccination: a population-based cohort study. Vaccine. 2025;47:126664. doi: 10.1016/j.vaccine.2024.126664
•• **A study describing the timeline of COVID-19 infections, vaccinations and details regarding Long-COVID in the Norwegian population from the cohorts presented here.**
20. Zhou P, Yang X-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020;579(7798):270–273. doi: 10.1038/s41586-020-2012-7
21. Belsky DW, Caspi A, Corcoran DL, et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. elife. 2022;11:e73420. doi: 10.7554/eLife.73420
22. Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573–591. doi: 10.18632/aging.101414
23. Brandlistuen RE, Kristjansson D, Alsaker E, et al. Cohort profile update: the Norwegian mother, father and child cohort (MoBa). Int J Epidemiol. 2025;54(5):dyaf139. doi: 10.1093/ije/dyaf139
• **A detailed description of the MoBa cohort.**
24. Laake I, Tunheim G, Robertson AH, et al. Risk of pregnancy complications and adverse birth outcomes after maternal A(H1N1) pdm09 influenza: a Norwegian population-based cohort study. BMC Infect Dis. 2018;18(1):525. doi: 10.1186/s12879-018-3435-8
25. Müller F, Scherer M, Assenov Y, et al. RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome Biol. 2019;20(1):55. doi: 10.1186/s13059-019-1664-9
26. Zhou W, Triche TJ, Laird PW, et al. SeSAMe: reducing artifactual detection of DNA methylation by infinium beadchips in genomic deletions. Nucleic Acids Res. 2018;46(20):e123. doi: 10.1093/nar/gky691

27. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinforma Oxf Engl. 2013;29(2):189–196. doi: 10.1093/bioinformatics/bts680
28. Koestler DC, Jones MJ, Usset J, et al. Improving cell mixture deconvolution by identifying optimal DNA methylation libraries (IDOL). BMC Bioinf. 2016;17(1):120. doi: 10.1186/s12859-016-0943-7
29. Salas LA, Koestler DC. FlowSorted.Blood.EPIC: illumina epic data on immunomagnetic sorted peripheral adult blood cells, R package version 2.14.0. 2025. Available from: https://bioconductor.org/packages/FlowSorted.Blood.EPIC
30. Cotton AM, Price EM, Jones MJ, et al. Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. Hum Mol Genet. 2015;24(6):1528–1539. doi: 10.1093/hmg/ddu564
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol. 1995;57(1):289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
32. Breitling LP, Yang R, Korn B, et al. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. Am J Hum Genet. 2011;88(4):450–457. doi: 10.1016/j.ajhg.2011.03.003
33. Pinheiro J, Bates D. R Core Team. nlme: linear and nonlinear mixed effects models. 2025. Available from: https://cran.r-project.org/web/packages/nlme/index.html
34. Bohlin J, Håberg SE, Magnus P, et al. MinLinMo: a minimalist approach to variable selection and linear model prediction. BMC Bioinf. 2024;25(1):380. doi: 10.1186/s12859-024-06000-4
35. Tunheim G, Fossum E, Robertson AH, et al. Characterization of the SARS-CoV-2 antibody landscape in Norway in the late summer of 2022: high seroprevalence in all age groups with patterns of primary Omicron infection in children and hybrid immunity in adults. BMC Infect Dis. 2024;24(1):841. doi: 10.1186/s12879-024-09670-w
• **The study shows that there is a high seroprevalence of SARS-CoV-2 antibodies in the Norwegian population following the Omicron surge.**
36. Editorial. Post-COVID-19 condition, brain fog, and fatigue-what do we know now? EBioMedicine. 2024 Jul;105:105252. doi: 10.1016/j.ebiom.2024.105252
37. Gómez-Díaz E, Jordà M, Peinado MA, et al. Epigenetics of host-pathogen interactions: the road ahead and the road behind. PLOS Pathog. 2012;8(11):e1003007. doi: 10.1371/journal.ppat.1003007
38. Fischer N. Infection-induced epigenetic changes and their impact on the pathogenesis of diseases. Semin Immunopathol. 2020;42(2):127–130. doi: 10.1007/s00281-020-00793-1
39. Shekhar Patil M, Richter E, Fanning L, et al. Epigenetic changes in patients with post-acute COVID-19 symptoms (PACS) and long-COVID: a systematic review. Expert Rev Mol Med. 2024;26:e29. doi: 10.1017/erm.2024.32
•• **This review summarizes current studies of Long-COVID and DNA methylation listing also challenges with current such studies.**
40. Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the Illumina MethylationEPIC beadchip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17(1):208. doi: 10.1186/s13059-016-1066-1
41. Tiedemann RL, Liang G, Jones PA. The human epigenome. In: Michels KB, editor. Epigenetic epidemiology. Cham: Springer International Publishing; 2022. p. 3–25.
42. Nustad HE, Steinsland I, Ollikainen M. Modeling dependency structures in 450k DNA methylation data. Bioinforma Oxf Engl. 2022;38(4):885–891. doi: 10.1093/bioinformatics/btab774
43. Caspersen IH, Magnus P, Trogstad L. Excess risk and clusters of symptoms after COVID-19 in a large Norwegian cohort. Eur J Epidemiol. 2022;37(5):539–548. doi: 10.1007/s10654-022-00847-8