

## Healthy lifestyle for the prevention of post-COVID-19 multisystem sequelae, hospitalization, and death: a prospective cohort study

Yunhe Wang, BSMed, MSc<sup>1†</sup>; Binbin Su, PhD<sup>2†</sup>; Marta Alcalde-Herraiz, PhD<sup>3</sup>; Nicola L. Barclay, PhD<sup>3</sup>; Yaohua Tian, PhD<sup>4</sup>; Chunxiao Li, PhD<sup>5</sup>; Nicholas J. Wareham, PhD<sup>5</sup>; Roger Paredes, PhD<sup>6,7</sup>; Junqing Xie, PhD<sup>3\*</sup>; Daniel Prieto-Alhambra, PhD<sup>3,8</sup>

1. Nuffield Department of Population Health, University of Oxford, Oxford, UK
2. School of Population Medicine and Public Health, Chinese Academy of Medical Sciences/Peking Union Medical College, Beijing, China
3. Centre for Statistics in Medicine and NIHR Biomedical Research Centre Oxford, NDORMS, University of Oxford, Oxford, UK
4. School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
5. Medical Research Council Epidemiology Unit, University of Cambridge, Cambridge, UK
6. Department of Infectious Diseases & irsiCaixa AIDS Research Institute, Hospital Universitari Germans Trias i Pujol, Catalonia, Spain
7. Center for Global Health and Diseases, Department of Pathology, Case Western Reserve University School of Medicine, Cleveland, OH, US
8. Department of Medical Informatics, Erasmus Medical Center University, Rotterdam, Netherlands

†Contributed equally

\*Corresponding to:

Junqing Xie, [Junqing.xie@ndorms.ox.ac.uk](mailto:Junqing.xie@ndorms.ox.ac.uk).

Centre for Statistics in Medicine and NIHR Biomedical Research Centre Oxford, NDORMS, University of Oxford, Oxford, UK

Word count: 3930 for main text

## Abstract

**Background** Post-COVID complications are emerging as a global public health crisis. Effective prevention strategies are needed to inform patients, clinicians and policy makers, and to reduce their cumulative burden. We aimed to investigate whether a habitual healthy lifestyle predated pandemic is associated with lower risks of multisystem sequelae and other adverse outcomes of COVID-19, and whether the potential protective effects are independent of pre-existing comorbidities.

**Methods** The prospective population-based cohort study enrolled participants with SARS-CoV-2 infection confirmed by a positive polymerase chain reaction test result between March 1, 2020, and March 1, 2022. Participants with no history of the related outcome one year before infection were included and followed up for 210 days. Exposures included ten modifiable healthy lifestyle factors including past or never smoking, moderate alcohol intake ( $\leq 4$  times week), body mass index  $< 30$  kg/m<sup>2</sup>, at least 150 minutes of moderate or 75 minutes of vigorous physical activity per week, less sedentary time ( $< 4$  hours per day), healthy sleep duration (7-9 hours per day), adequate intake of fruit and vegetables ( $\geq 400$  g/day), adequate oily fish intake ( $\geq 1$  portion/week), moderate intake of red meat ( $\leq 4$  portions week) and processed meat ( $\leq 4$  portions week). Outcomes included multisystem COVID-19 sequelae (consisting of 75 diseases/symptoms in 10 organ systems), death, and hospital admission following SARS-CoV-2 infection, confirmed by hospital inpatient and death records. Risk was reported in relative scale (hazard ratio [HR]) and absolute scale (absolute risk reduction [ARR]) during both the acute (the first 30 days) and post-acute (30-210 days) phases of infection using Cox models.

**Findings** A total of 68,896 participants (mean [SD] age, 66.6 [8.4]; 32,098 women [46.6%]) with COVID-19 were included. A favorable lifestyle (6-10 healthy lifestyle factors; 46.4%) was associated with a 36% lower risk of multisystem sequelae of COVID-19 (HR, 0.64; 95% CI, 0.58-0.69; ARR, 7.08%; 95% CI, 5.98-8.09), compared with unfavorable lifestyle (0-4 factors; 12.3%). Risk reductions were observed across all 10 prespecified organ systems including cardiovascular, coagulation, metabolic and endocrine, gastrointestinal, kidney, mental health, musculoskeletal, neurologic, and respiratory disorders, and general symptoms of fatigue and malaise. This beneficial effect was largely attributable to direct effects of healthy lifestyle, with mediation proportion ranging from 44% to 93% across organ systems. A favorable lifestyle was also associated with lower risk of post-COVID death (HR, 0.59; 95% CI, 0.52-0.66; ARR, 1.99%; 95% CI, 1.61-2.32) and hospitalization (HR, 0.78; 95% CI, 0.73-0.84; ARR, 6.14%; 95% CI, 4.48-7.68). These associations were observed after accounting for potential misclassification of lifestyle factors, and during acute and post-acute infection, in those tested positive in the hospital and community setting, and independent of vaccination status or SARS-CoV-2 variant.

**Interpretation** Adherence to a healthy lifestyle predated pandemic was associated with substantially lower risk of complications across organ systems, death, and hospitalization following COVID-19, regardless of phases of infection, vaccination status, test setting, and SARS-CoV-2 variants, and independent of comorbidities. These findings illustrate the benefits of adhering to a healthy lifestyle to reduce the long-term adverse health consequences following SARS-CoV-2 infection.

## Research in context

### Evidence before this study

We searched PubMed and MEDLINE for articles published between March 1, 2020, and December 1, 2023, using the search terms “healthy lifestyle”, “risk factor”, “post-COVID condition”, “long COVID”, “post-acute sequelae”, “prevention”, “management”, and “treatment”, with no language restrictions. Previous evidence on the prevention and management of long COVID has mainly focused on vaccination and pharmaceutical approaches, including antivirals (e.g., molnupiravir and nirmatrelvir) and other drugs (e.g., metformin). Vaccination before infection or use of antivirals in selected high-risk patients during acute infection only partially mediates the risk of COVID-19 sequelae. Evidence for the non-pharmaceutical prevention strategies are lacking. We identified only two publications on the association between healthy lifestyle and post-COVID condition, and one meta-analysis of the risk factors for long COVID symptoms. A cross-sectional study of 1981 women suggested an inverse association between healthy lifestyle factors and self-reported symptoms following infection of non-Omicron variants, which was mainly driven by BMI and sleep duration. Another study suggested an inverse prospective association between healthy lifestyle prior to infection and post-COVID cardiovascular events. High BMI and smoking are risk factors for long COVID mainly in hospitalized patients. We did not find any study that assessed the association between a composite healthy lifestyle and subsequent post-COVID complications or sequelae across organ systems, hospitalization, and death.

### Added value of this study

In a prospective, population-based cohort of 68,896 participants with COVID-19, adherence to a healthy lifestyle prior to infection was associated with a substantially lower risk of multisystem sequelae (by 20%-36%), death (by 26%-41%), and hospital admission (by 13%-22%) following COVID-19. The reduced risk of sequelae was evident across 10 prespecified organ systems, including cardiovascular, coagulation and hematologic, metabolic and endocrine, gastrointestinal, kidney, mental health, musculoskeletal, neurologic, and respiratory disorders, as well as general symptoms of fatigue and malaise. The reduced risk of multisystem sequelae, hospitalization, and death associated with a healthy lifestyle was consistently observed across participants, regardless of their vaccination status, disease severity, and major SARS-CoV-2 variants, and largely independent of relevant comorbidities. Adherence to a healthy lifestyle prior to infection was consistently and directly associated with reduced risk of sequelae and other adverse health outcomes following COVID-19.

### Implications of all the available evidence

The inverse association of healthy lifestyle with multisystem sequelae was even larger than those observed in previous studies of pharmaceutical interventions in non-hospitalized patients. Considering the restricted scope of currently available therapies, such as antivirals (only selected patients at higher risk are qualified during the acute infection) and limited efficacy of vaccination in preventing long COVID, adherence to a healthy lifestyle, in combination with vaccination and, if necessary, potential medications, emerges as practical prevention and care strategies to mitigate the long-term health consequences of SARS-CoV-2 infection. These strategies are of significant clinical and public health importance in reducing the overall burden of post-COVID conditions and improving preparedness for future pandemics.

## Introduction

COVID-19 cases and deaths have decreased globally, yet the long-term health consequences of SARS-CoV-2 infection, termed as *post-COVID-19 conditions* or *long COVID*, are still being managed as a global public health crisis.<sup>1, 2</sup> These conditions or symptoms can involve pulmonary and multiple extrapulmonary organ systems, and may occur or extend beyond the acute infection of varying severity, with significant impact on daily functioning and quality of life.<sup>3</sup> Increased risk and burden of cardiovascular, pulmonary, neuropsychiatric, and metabolic disorders were reported during the 6 to 12 months following SARS-CoV-2 infection,<sup>4</sup> with persistent risk observed for several diseases up to 2 years.<sup>5, 6</sup>

Despite long COVID has been characterized, however, evidence-based strategies for its prevention or treatment are not yet available.<sup>7, 8</sup> Increasing evidence suggests that vaccination before infection<sup>9, 10</sup> and use of antivirals<sup>11, 12</sup> during acute phase only partially mediate the risk of sequelae. Several potential drugs for long COVID are still under investigation without yielding reliable results.<sup>7, 8</sup> Effective prevention and intervention strategies are needed to inform patients, clinicians and policy makers, and to reduce the cumulative burden of post-COVID conditions.

Modifiable lifestyle factors such as physical activity and healthy diet are potential targets for the prevention of major non-communicable diseases,<sup>13-15</sup> and are associated with lower risk of severe COVID-19 and related mortality,<sup>16, 17</sup> possibly through protection against inflammation,<sup>15, 16</sup> autoimmunity,<sup>18, 19</sup> and clotting abnormality.<sup>20, 21</sup> These mechanisms overlap with the hypothesized pathogenesis of long COVID, and other postviral conditions, such as chronic fatigue syndrome.<sup>3</sup> When examined individually, factors such as smoking and obesity, have been reported to be related to increased risk of post-COVID symptoms mainly in hospitalized patients.<sup>22</sup> Nevertheless, association between combinations of multiple lifestyle factors, which are known to interact synergistically,<sup>23</sup> and risk of COVID-19 sequelae across multiple organ systems remains unclear.

This major knowledge gap should be urgently addressed to inform the prevention and care strategies of long COVID. Based on a large-scale, prospective population-based cohort, we evaluated the relationship between composite healthy lifestyle (including 10 modifiable factors) predated pandemic and subsequent risk of COVID-19 sequelae in 10 organ systems, death and hospital admission, while considering the phase of infection (acute or post-acute), severity of infection (tested positive in community/outpatient setting vs inpatient setting), vaccination status (fully vaccinated vs unvaccinated or partially vaccinated), and variants (alpha [B.1.1.7], delta [B.1.617.2], vs omicron [B.1.1.529]), that differ in transmissibility, disease course, and disease severity.<sup>24</sup>

## Methods

### Data Sources and Study Cohorts

UK Biobank is a large-scale population-based prospective cohort study with deep phenotyping and genomic data, as detailed elsewhere.<sup>25</sup> Briefly, between 2006 to 2010, over 500,000 individuals aged 40-69 years were recruited from 22 assessment centers across the United Kingdom at baseline, with collection of socio-demographic, lifestyle and health-related factors, a range of physical measures, and blood samples.<sup>25</sup> Follow-up information is obtained by linking health and medical records, including national primary and secondary care, disease and mortality registries,<sup>26</sup> with validated reliability, accuracy and completeness.<sup>27</sup> To identify cases of SARS-CoV-2 infection, polymerase chain reaction (PCR)-based test results were obtained by linking all participants to the Public Health England's Second Generation Surveillance System, with dates of specimen collection and healthcare settings of testing.<sup>28</sup> Outbreak dynamics were validated to be broadly similar between UK Biobank participants and the general population of England.<sup>28</sup>

In this study, we enrolled participants who were alive by March 1, 2020 and had a positive SARS-CoV-2 PCR test result between March 1, 2020 (date of the first recorded case in the UK Biobank), and March 1, 2022, with the date of first infection considered as index date ( $T_0$ ). The major prevalent variants during the study period included wildtype,

Alpha (B.1.1.7), Delta (B.1.617.2), and Omicron (B.1.1.529 BA.1). The calendar periods of dominant variants in the UK were based on pandemic data from the Office for National Statistics.<sup>24</sup> Participants with missing data on study exposures and covariates of interest at baseline were further excluded. All participants included in this study provided written informed consent at recruitment. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines and received ethical approval from the UKBB ethics advisory committee. Study design, cohort construction, and timeline are provided in **Supplementary Fig. 1**.

## Lifestyle factors

Ten prespecified potentially modifiable lifestyle factors were assessed, including smoking, alcohol consumption, body mass index (BMI), physical activity, sedentary time, sleep duration, intake of fruit and vegetable, intake of oily fish, intake of red meat, and intake of processed meat. Selection and categorization of lifestyle factors was based on literature review, previous knowledge, and UK national health service guidelines.<sup>29, 30</sup> Multiple lifestyle factors were measured by validated questionnaire for all participants at baseline recruitment. Part of participants took part in up to two further touchscreen interviews with lifestyle and health-related factors similarly measured. There were generally stable responses to lifestyle factors between baseline assessment and the latest repeat assessment (median time difference from baseline, 8 years) as shown in **Supplementary Fig. 2**. Detailed definitions on measurement and classification of lifestyle factors are provided in **Supplementary Table 1**. Briefly, healthy lifestyle components including past or never smoker, moderate alcohol intake ( $\leq 4$  times week), BMI  $< 30$  kg/m<sup>2</sup>, at least 150 minutes of moderate or 75 minutes of vigorous physical activity per week, less sedentary time ( $< 4$  hours per day), healthy sleep duration (7-9 hours per day), adequate intake of fruit and vegetables ( $\geq 400$  g/day), adequate oily fish intake ( $\geq 1$  portion/week), moderate intake of red meat ( $\leq 4$  portion week) and processed meat ( $\leq 4$  portion week) were defined, in accordance with previous evidence or UK national health service guidelines.<sup>29, 30</sup>

A binary variable was created for each of the 10 factors, with 1 point assigned for those meeting the healthy criteria and 0 otherwise. A composite lifestyle score was then calculated for each participant by summing the total number of healthy lifestyle factors, ranging from 0 to 10. Based on the composite score, participants were classified into three lifestyle categories: unfavorable (0-5), intermediate (6-7), and favorable (8-10). The lifestyle score was also used as a continuous variable of number of healthy lifestyle factors. Similar methods of defining lifestyle score have been used in the same UK Biobank cohort<sup>31</sup> as well as external cohorts.<sup>14, 32</sup> Distributions of lifestyle score and categories are provided in **Supplementary Table 2**.

## Outcomes

The outcomes after COVID-19 were prespecified, including a set of multisystem sequelae, death, and hospital admission following the SARS-CoV-2 infection. The multisystem sequelae were selected and defined based on previous evidence of the long COVID, including 75 systemic diseases or symptoms in 10 organ systems: cardiovascular,<sup>33</sup> coagulation and hematologic,<sup>33</sup> metabolic and endocrine,<sup>34</sup> gastrointestinal,<sup>35</sup> kidney,<sup>36</sup> mental health,<sup>37</sup> musculoskeletal,<sup>38</sup> neurologic,<sup>38</sup> and respiratory disorders,<sup>9, 11, 12</sup> and general symptoms of fatigue and malaise.<sup>3, 4, 39, 40</sup> Detailed definitions of multisystem sequelae are listed in the **Supplementary Table 3**. Outcomes were identified as follows: individual sequela from the hospital inpatient ICD-10 (*International Classification of Diseases 10th Revision*) diagnosis codes, deaths from the records of national death registry, and hospital admission from hospital inpatient data from the Hospital Episode Statistics. Incident outcomes were assessed in participants with no history of the related outcome within one year before the date of the first infection.

As SARS-CoV-2 infection has been associated with both multisystem manifestations during its acute phase and with sequelae during its post-acute phase,<sup>6, 40</sup> we conducted analyses stratified by phase of infection. We reported risk of each outcome during the acute phase ( $T_0$  to  $T_0 + 30$ d), post-acute phase ( $T_0 + 30$ d to  $T_0 + 210$ d), and overall period following infection ( $T_0$  to  $T_0 + 210$ d) to reflect the full spectrum of post-COVID conditions. The end of follow-up for the overall cohort was September 30, 2022, with the maximum follow-up period censored to 210 days.

## Covariates

We prespecified a list of covariates for adjustment or stratification based on literature review and prior knowledge: sociodemographic characteristics including age, sex, education level (mapped to the international standard for classification of education), index of multiple deprivation (IMD, a summary measure of crime, education, employment, health, housing, income, and living environment),<sup>41</sup> and race and ethnicity; and infection related factors including healthcare settings of the testing (community/outpatient vs inpatient setting as proxy of severity of infection), COVID-19 vaccination status, and SARS-CoV-2 variants.

## Statistical analysis

Baseline characteristics of the overall cohort of participants with SARS-CoV-2 infection and by composite healthy lifestyle categories were reported as mean and standard deviation or frequency and percentage, when appropriate. Multivariable cox proportional hazard (PH) model was used to assess the association between composite healthy lifestyle and risk of multisystem sequelae (composite or by organ systems), death, and hospital admission, with adjustment for age, sex, ethnicity, education level, and IMD. PH assumption across lifestyle categories was tested by Schoenfeld residuals with no violations observed for outcomes. Hazard ratio (HR) and absolute risk reduction (ARR, difference in incidence rate between lifestyle groups per 100 persons during the corresponding follow-up period) were estimated from the Cox model. We also assessed the association between individual lifestyle factor instead of composite categories (each component as a categorical variable with or without mutual adjustment for others, or the number of factors as continuous variables) and risk of outcomes.

We conducted causal mediation analysis<sup>42, 43</sup> to quantify the extent to which the habitual healthy lifestyle may affect COVID-19 sequelae through the potential pathway of relevant pre-infection medical conditions (mediator), with proportion of direct and indirect effects estimated by quasi-Bayesian Monte Carlo methods with 1000 simulations for each. Detailed modelling procedures and a directed acyclic graph are provided in eMethods in Supplement.

We examined the association between composite healthy lifestyle and overall risk of multisystem sequelae in prespecified clinical subgroups by demographic and infection-related factors. The demographic factors included age ( $\leq 65$  and  $> 65$  years), sex (male and female), and ethnicity (White and other ethnic groups). As the risk profile of COVID sequelae was related to vaccination and severity of infection, and may change with the evolving pandemic, infection-related factors including vaccine status (no or one-dose partial vaccination and two-dose full vaccination), test setting (inpatient and outpatient or community), dominant variants during the study period (wildtype, Alpha, Delta, and Omicron BA.1) were assessed. Multiplicative interactions between the composite healthy lifestyle and the stratification variables were tested, with *P*-value reported.

We conducted multiple sensitivity analyses to assess the robustness of primary findings. First, to reflect the multisystem and potentially comorbid nature of COVID sequelae, accounting for both the number of sequelae by an individual and the relative health impact of each sequela. Weights based on Global Burden of Disease study data and methodologies for general diseases and long COVID were assigned to each sequela (**Supplementary Table 1**).<sup>44, 45</sup> The weighted score was calculated for each participant by summing the weights of all incident sequelae during the follow-up period. Zero inflated Poisson regression was then used to calculate relative risk (RR), with follow-up time set as the offset of the model and adjustment for covariates. Second, to further account for potential reverse causality and more accurately define incident cases, extending the washout period for outcomes from one year to two years. Third, defining events of post-acute sequelae 90 days after infection (follow-up period  $T_0 + 90d$  to  $T_0 + 210d$ ), instead of 30 days in the main analyses. The adjustment was made as there is no uniform definition for long COVID, which is currently described as conditions occurring 30-90 days after infection in existing guidelines.<sup>46</sup> Fourth, restricting the identification of outcomes to the first three ICD diagnoses, which are the main causes for each hospital admission. Fifth, we conducted quantitative sensitivity analysis to adjust for changes in lifestyle factors over time since the baseline assessment. We used odds ratios to quantify associations and assumed a sensitivity and specificity of 90% for each lifestyle component (eMethods in Supplement).

Statistical significance was determined by a 95% confidence interval (CI) that excluded 1 for ratios and 0 for rate differences. All analyses and data visualizations were conducted using R statistical software (version 4.2.2).

## Results

### Baseline Characteristics

Out of 472,977 eligible UK Biobank participants, 68,896 participants with a positive SARS-CoV-2 test result between March 1, 2020 and March 1, 2022 were included in the current study. The demographic and health characteristics of the eligible participants, and participants with COVID-19 overall and by healthy lifestyle category are provided in **Table 1**. Of the COVID-19 cohort, the mean (SD) age was 66.6 (8.4) years, 53.4% were male and 82.1% were White. For composite healthy lifestyle prior to the infection, 12.3% followed an unfavorable lifestyle, 41.3% followed an intermediate lifestyle, and 46.4% followed a favorable lifestyle. The median [IQR] number of healthy lifestyle factors participants engaged in was 7 [6-8]. For prespecified COVID-19 sequelae, 5.5% and 7.8% had sequelae in at least one organ system during the acute and post-acute phase of infection, respectively.

### Risk of multisystem sequelae

Overall, the risk of multisystem COVID-19 sequelae decreased monotonically across healthy lifestyle categories during both the acute and post-acute phases of infection. Compared with those with an unfavorable lifestyle, participants with an intermediate (HR, 0.80; 95% CI, 0.74-0.87; ARR at 210 days, 3.89%; 95% CI, 2.56-5.12) and favorable (HR, 0.64; 95% CI, 0.58-0.69; ARR at 210 days, 7.08%; 95% CI, 5.98-8.09) lifestyle were at significantly lower risk of multisystem sequelae of COVID-19 (**Figure 1A** and eFigure 3 in **Supplement**), with similar trends observed in both the acute and post-acute phases of COVID-19 (**Figure 1B**). The number of healthy lifestyle factors (range, 0-10) was associated with risk of sequelae in a dose-dependent manner (**Figure 1A**).

The inverse associations with multisystem sequelae were largely attributable to the direct protective effect of healthy lifestyle (proportion of direct effect on any sequela: 71%), with proportion of direct effect ranging from 44% to 93% across organ systems (**Figure 2A**). Pre-infection medical conditions were associated with substantially increased risk of COVID-19 sequelae, particularly history of cardiovascular diseases, diabetes and mental disorders (**Figure 2B**).

For individual components of the healthy lifestyle, each of the 10 studied behavioral and dietary factors was associated with a lower or non-differential risk of sequelae, with smoking, physical activity, obesity and sleep duration contributing most (**Figure 3**).

### Risk of death and hospitalization

Adherence to a healthy lifestyle was associated with lower risk of death and hospitalization following COVID-19 during both the acute and post-acute phases of infection. Compared with those with an unfavorable lifestyle, participants with a favorable lifestyle were at significantly lower risk of death (HR, 0.59; 95% CI, 0.52-0.66; ARR at 210 days, 1.99%; 95% CI, 1.61-2.32) and hospitalization (HR, 0.78; 95% CI, 0.73-0.84; ARR at 210 days, 6.14%; 95% CI, 4.48-7.68) following COVID-19 (**Figures 1A** and eFigure 3 in **Supplement**), with similar trend observed in both the acute and post-acute phases (**Figures 1B**). Each of 10 lifestyle factors was associated with a lower or non-differential risk of death and hospitalization (**Figure 3**).

### Risk of system-specific sequelae

Compared with those following an unfavorable lifestyle, participants with a favorable lifestyle had a significantly lower risk of sequelae in all 10 organ systems examined, including cardiovascular, coagulation and hematologic, metabolic and endocrine, gastrointestinal, kidney, mental health, musculoskeletal, neurologic, and respiratory disorders, as well as general symptoms of fatigue and malaise, with overall HRs ranging from 0.38 to 0.76 (**Figure 1A** and **Figure 1B**). The associations with intermediate lifestyle were consistently in the same protective direction across system-specific

sequelae (**Figure 1A** and **Figure 1B**). Similar trend were observed in both the acute and post-acute phases (**Figure 1B**).

### Risk of outcomes by subgroups

The inverse associations between healthy lifestyle and risk of multisystem sequelae, death, and hospitalization held across the different subgroups of clinical interest, including those by age, sex, and ethnicity, vaccine status, test setting, and variants of infection (**Table 2**). The reduced risk of outcomes was observed in participants who received two doses of vaccine (breakthrough infection) and those who were unvaccinated or partially 1-dose vaccinated (non-breakthrough infection). The reduced risk was evident in participants tested positive in inpatient setting and in those tested positive in community/outpatient settings. The reduced risk was consistently observed across predominant variants of SARS-CoV-2 infection during study period, including wild-type, Alpha, Delta, and Omicron BA.1. Notably, composite healthy lifestyle was associated with decreased risk of outcomes following infection of Omicron variant, which remains currently dominant variant worldwide. No significant interaction was observed in any of the subgroups across outcomes, except for age, ethnicity, and vaccination. The observed association between favorable lifestyle and a reduced risk of outcomes was more evident for people aged <65, in white subjects, and in those fully vaccinated (for mortality only).

### Sensitivity analyses

A similar pattern of associations was observed in multiple sensitivity analyses, including assigning weight to individual sequela and using zero inflated Poisson regression to estimate the association, excluding participants with history of related outcomes in the past two years rather than one year, defining post-acute outcomes 90 days after infection rather than 30 days, and restricting the identification of outcomes to the first three ICD diagnoses (eTable 4 in **Supplement**). Stronger associations were identified after accounting for potential misclassification of lifestyle factors, suggesting that the observed associations may have been underestimated (eTable 4 in **Supplement**).

### Discussion

Based on a large, prospective population-based cohort, this study provides a comprehensive assessment of the health effects of multiple lifestyle factors on a systematic range of disease outcomes following COVID-19. Adherence to a healthy lifestyle prior to infection was associated with significantly lower risk of COVID-19 multisystem sequelae, death, and hospital admission, during both the acute and post-acute phases of SARS-CoV-2 infection. The reduced risk was evident across 10 prespecified organ systems, including cardiovascular, coagulation and hematologic, metabolic and endocrine, gastrointestinal, kidney, mental health, musculoskeletal, neurologic, and respiratory disorders, as well as general symptoms of fatigue and malaise. The reduced risk of multisystem sequelae associated with a healthy lifestyle was consistently observed across participants, regardless of their vaccination status (unvaccinated/partially vaccinated or fully vaccinated), disease severity (testing positive in community/outpatient settings or inpatient settings), and major SARS-CoV-2 variants, including Omicron variants, whose subvariants are currently dominant. Moreover, the benefits of healthy lifestyle on sequelae were largely independent of pathways related to pre-existing disease conditions. Overall, the findings suggest that adherence to a healthy lifestyle prior to infection was consistently and directly associated with reduced risk of adverse health outcomes following COVID-19.

We found that a favorable lifestyle, in comparison with an unhealthy one, was associated with a 36% lower risk of multisystem sequelae, a 41% lower risk of death, and a 22% lower risk of hospitalization, which corresponded to an absolute risk reduction of 7.08, 1.99, and 6.14 fewer cases per 100 people at 210 days after infection. This association was even larger than those observed in previous studies of pharmaceutical interventions in non-hospitalized patients, which reported a 14% risk reduction in post-acute sequelae at 180 days for vaccination before infection, and 14% and 26% risk reductions at 180 days for the use of molnupiravir and nirmatrelvir during acute phase of infection, respectively.<sup>9, 11, 12</sup> It is important to note that participants with breakthrough infection were still at risk of sequelae compared with those without infection.<sup>9</sup> In addition, only selected patients at risk of progression to severe COVID-19



are qualified for antivirals during the acute infection,<sup>11, 12</sup> and their benefit-risk profile in wider population with milder infection, or when used during the post-acute stage, remains unclear. These previous findings highlighted the restricted scope of currently available therapies and limited efficacy of vaccination in preventing long COVID.<sup>9, 11, 12, 46</sup> Our results are consistent with a cross-sectional study of 1981 women suggesting an inverse association between composite healthy lifestyle (mainly driven by BMI and sleep duration) and self-reported symptoms following infection of non-Omicron variants.<sup>32</sup> However, outcomes purely based on self-report symptoms are less clinically relevant and the inclusion of only women may limit the generalization of findings to other populations and settings.

The mechanisms underlying the benefit of adhering to healthy lifestyle for the alleviation of sequelae are likely multifaceted. Previous research has established causal links between several individual lifestyle factors, such as smoking, obesity, and physical inactivity, and increased susceptibility and severity in relation to COVID-19.<sup>47</sup> Smoking and high BMI were also risk factors for long COVID symptoms mainly in hospitalized patients.<sup>22</sup> Indeed, we observed that these factors and additionally sleep duration and sedentary behavior were significant contributors to the reduced risk of sequelae. However, these pathways are unlikely to fully explain the protective effects conferred by healthy lifestyle. In our study, all participants had a confirmed infection, and the protective associations persisted even among those who were hospitalized. In addition, it has been suggested that individuals with an unhealthy lifestyle are more likely to have prevalent chronic conditions, such as cardiovascular diseases and diabetes—which are strong risk factors for severe COVID-19—and are therefore more vulnerable to post-acute complications. Through mediation analysis, our study supported this hypothesis and, for the first time, further demonstrated that the healthy lifestyle's direct protection accounts for the majority of the overall associations with COVID-19 sequelae.

Although our findings align with previous evidence on the broader benefits of healthy lifestyle on chronic disease prevention and life expectancy,<sup>13-15</sup> this potentially beneficial effect should not be interpreted as changing behaviors around the time of acute infection or during post-acute infection. The current practical guide recommends that patients without long COVID should gradually and safely return to pre-infection physical activity when appropriate, although direct evidence is lacking.<sup>48</sup> Previous study also characterized long COVID as a multifactorial condition determined by pathogen, host response, external pandemic-associated factors, and supported a multidisciplinary treatment including both pharmacological and rehabilitation approaches, but also social and welfare support to promote healthy lifestyle habits.<sup>49, 50</sup> Future research is warranted to assess the effect of composite lifestyle interventions in prevention of long COVID or alleviating associated symptoms among patients with long COVID. Adherence to a healthy lifestyle, in combination with vaccination and, if necessary, potential medications, may be a viable and practical approach to further reduce the long-term health consequences of SARS-CoV-2 infection. These strategies hold significant public health and scientific importance in mitigating the overall burden of post-COVID complications and enhancing preparedness for future pandemics.

## Limitation

This study has several limitations. First, the UK Biobank participants are likely to be older and healthier than the general population of the UK and are mostly of European ancestry, which may limit the generalizability of study findings. Despite the relative risk were largely not influenced, the absolute risk should be interpreted with caution. Second, residual confounding and reverse causality cannot be ruled out in this observational study. Third, we assumed the baseline lifestyle unchanged over years until the time of infection, which may be subject to exposure misclassification and underestimated any genuine associations. However, reassuringly, there was high consistency of lifestyle measures between cohort baseline and repetitive interviews after a median of 8 years, and consistent associations were observed after accounting for potential changes in lifestyle factors. Forth, as sequelae outcomes were based on inpatient records, milder long COVID symptoms were less likely to be detected. Fifth, given the potential non-linear effects of lifestyle factors, such as alcohol consumption, caution is warranted when interpreting associations between binarized lifestyle factors and outcomes. Finally, despite we included a range of sequelae across organ systems, it is difficult to link these outcomes directly to the infection, especially given the lack of consensus standard for diagnosis of long COVID. This limitation is applied to all related studies in the field.<sup>33-40, 51</sup>

Nevertheless, the sequelae prespecified were most relevant to long COVID based on prior evidence, with increased risk and burden consistently reported beyond the acute phase in the currently largest electronic dataset,<sup>33-39</sup> the same UK Biobank,<sup>51</sup> and other nationwide cohorts.<sup>40, 51</sup>

## Conclusion

Adherence to a healthy lifestyle predated pandemic was associated with substantially lower risk of sequelae across organ systems, death, and hospitalization following COVID-19, regardless of phases of infection, vaccination status, test setting, and SARS-CoV-2 variants, and independent of relevant comorbidities. These findings suggest the benefit of population adhering to a healthy lifestyle to reduce the potential long-term adverse health consequences of COVID-19.

## Acknowledgements

This study was based on data from UK Biobank. All participants provided written informed consent at the UK Biobank cohort recruitment. This study received ethical approval from UK Biobank Ethics Advisory Committee (EAC) and was performed under the application of 65397. Mr Wang is funded through the Clarendon Fund Scholarship. Dr Xie is funded through Jardine-Oxford Graduate Scholarship and a titular Clarendon Fund Scholarship. The research was partially supported by the Oxford National Institute for Health and Care Research (NIHR) Biomedical Research Centre. Dr Prieto-Alhambra is funded through an NIHR Senior Research Fellowship (grant SRF-2018-11-ST2-004). The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

## Author contributions

Drs Wang and Xie had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analyses. *Concept and design:* Wang and Xie. *Acquisition, analysis, or interpretation of data:* Wang, Xie, and Prieto-Alhambra. *Drafting of the manuscript:* Wang. *Critical revision of the manuscript for important intellectual content:* All authors. *Statistical analysis:* Xie. *Obtained funding:* Prieto-Alhambra. *Administrative, technical, or material support:* Prieto-Alhambra. *Supervision:* Nicholas J. Wareham and Prieto-Alhambra.

## Competing interests

Dr Prieto-Alhambra's department has received grant/s from Amgen, Chiesi-Taylor, Lilly, Janssen, Novartis, and UCB Biopharma. His research group has received consultancy fees from Astra Zeneca and UCB Biopharma. Amgen, Astellas, Janssen, Synapse Management Partners and UCB Biopharma have funded or supported training programmes organised by Dr Prieto-Alhambra's department. Dr Paredes has participated in advisory boards for Gilead, MSD, ViiV Healthcare, Theratechnologies and Lilly. His institution has received research support from Gilead, MSD, and ViiV Healthcare. The remaining authors declare no competing interests.

## Table Legends

**Table 1.** Baseline characteristics of eligible participants, and COVID-19 cohort overall and by lifestyle category

**Table 2.** Association of composite healthy lifestyle with multisystem sequelae, death, and hospitalization following SARS-CoV-2 infection in key clinical subgroups

## Figure Legends

**Figure 1.** Association of healthy lifestyle with multisystem sequelae of COVID-19, death, and hospital admissions, during overall, acute, and post-acute phases of SARS-CoV-2 infection

**Figure 2.** Direct and indirect effects of healthy lifestyle, and association of pre-infection medical conditions with multisystem sequelae of COVID-19

**Figure 3.** Association of individual components of composite healthy lifestyle with multisystem sequelae, death, and hospitalization

**Table 1. Baseline characteristics of eligible participants, and COVID-19 cohort overall and by lifestyle category**

Characteristics	All eligible participants <sup>a</sup>	Participants with COVID-19			
		Overall	By composite lifestyle category <sup>b</sup>		
			Unfavorable	Intermediate	Favorable
No. (%)	472977	68896	8476 (12.3)	28457 (41.3)	31963 (46.4)
Age, mean (SD)	68.1 (8.1)	66.6 (8.4)	66.6 (8.3)	66.8 (8.4)	66.5 (8.4)
Sex, No. (%)					
Male	261436 (55.3)	36798 (53.4)	3206 (37.8)	14207 (49.9)	19385 (60.6)
Female	211540 (44.7)	32098 (46.6)	5270 (62.2)	14250 (50.1)	12578 (39.4)
Ethnicity, No. (%)					
White	382010 (80.8)	56566 (82.1)	6842 (80.7)	23247 (81.7)	26477 (82.8)
Other ethnic groups	90967 (19.2)	12330 (17.9)	1634 (19.3)	5210 (18.3)	5486 (17.2)
Education categories, <sup>c</sup> No. (%)					
I	85686 (18.1)	11398 (16.5)	1946 (23.0)	5140 (18.1)	4312 (13.5)
II	125549 (26.5)	20644 (30.0)	2826 (33.3)	8670 (30.5)	9148 (28.6)
III	52738 (11.2)	7748 (11.2)	869 (10.3)	3112 (10.9)	3767 (11.8)
IV	24148 (5.1)	3205 (4.7)	364 (4.3)	1340 (4.7)	1501 (4.7)
V	184856 (39.1)	25901 (37.6)	2471 (29.2)	10195 (35.8)	13235 (41.4)
Index of multiple deprivation, <sup>d</sup> mean (SD)	17.5 (12.9)	18.6 (13.6)	21.9 (15.3)	19.4 (13.9)	17.1 (12.5)
Individual lifestyle factors, No. (%)					
Smoking					
Never or past	425810 (90.0)	62066 (90.1)	6043 (71.3)	25126 (88.3)	30897 (96.7)
Current	47167 (10.0)	6830 (9.9)	2433 (28.7)	3331 (11.7)	1066 (3.3)
Alcohol consumption					
≤ 4 times week	377958 (79.9)	56248 (81.6)	5526 (65.2)	21944 (77.1)	28778 (90.0)
Daily or almost daily	95019 (20.1)	12648 (18.4)	2950 (34.8)	6513 (22.9)	3185 (10.0)
BMI					
<30 kg/m <sup>2</sup>	359704 (76.1)	51001 (74.0)	3592 (42.4)	19011 (66.8)	28398 (88.8)
≥30 kg/m <sup>2</sup>	113273 (23.9)	17895 (26.0)	4884 (57.6)	9446 (33.2)	3565 (11.2)
Physical activity					
≥150 min/week of MIPA or ≥75 min/week of VIPA	398880 (84.3)	57930 (84.1)	4913 (58.0)	22857 (80.3)	30160 (94.4)
<150 min/week of MIPA and <75 min/week of VIPA	74097 (15.7)	10966 (15.9)	3563 (42.0)	5600 (19.7)	1803 (5.6)
TV viewing/sedentary time					
< 4 h/day	339946 (71.9)	50500 (73.3)	3388 (40.0)	18819 (66.1)	28293 (88.5)
≥ 4 h/day	133031 (28.1)	18396 (26.7)	5088 (60.0)	9638 (33.9)	3670 (11.5)
Sleep duration					
7-9 h/day	349546 (73.9)	50777 (73.7)	3718 (43.9)	18982 (66.7)	28077 (87.8)
<7 or >9 h/day	123431 (26.1)	18119 (26.3)	4758 (56.1)	9475 (33.3)	3886 (12.2)
Fruit and vegetable intake					
≥ 400 g/day	377154 (79.7)	53635 (77.8)	3699 (43.6)	20696 (72.7)	29240 (91.5)
<400 g/day	95823 (20.3)	15261 (22.2)	4777 (56.4)	7761 (27.3)	2723 (8.5)
Oily fish intake					
≥1 portion/week	265642 (56.2)	35996 (52.2)	1727 (20.4)	11504 (40.4)	22765 (71.2)
<1 portion/week	207335 (43.8)	32900 (47.8)	6749 (79.6)	16953 (59.6)	9198 (28.8)
Red meat intake					
≤3 portion/week	216206 (45.7)	34003 (49.4)	3375 (39.8)	12650 (44.5)	17978 (56.2)
>3 portion/week	256771 (54.3)	34893 (50.6)	5101 (60.2)	15807 (55.5)	13985 (43.8)

Processed meat intake					
≤1 portion/week	328721 (69.5)	46295 (67.2)	2536 (29.9)	16352 (57.5)	27407 (85.7)
>1 portion/week	144256 (30.5)	22601 (32.8)	5940 (70.1)	12105 (42.5)	4556 (14.3)

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); MIPA, moderate intensity physical activity; VIPA, vigorous intensity physical activity.

<sup>a</sup>All eligible participants included were those who were alive when the study began (March 1, 2020) in UK Biobank.

<sup>b</sup>Participants were classified into three lifestyle categories: favorable (0-4 unhealthy lifestyle factors), intermediate (5-6), and unfavorable (7-10), based on the composite score of 10 modifiable lifestyle factors including smoking, alcohol consumption, BMI, physical activity, TV watching/sedentary time, sleep duration, intake of fruit and vegetables, oily fish intake, red meat intake, and processed meat intake.

<sup>c</sup>Education levels were mapped to the international standard for classification of education. Category I includes self-reported "None of the above" and "Prefer not to answer", II includes "CSEs or equivalent" and "O levels/GCSEs or equivalent", III includes "A levels/AS levels or equivalent", IV includes "Other professional qualifications e.g.: nursing, teaching", and V includes "NVQ or HND or HNC or equivalent" and "College or University degree."

<sup>d</sup>A summary measure of crime, education, employment, health, housing, income, and living environment, with high score indicating higher levels of deprivation.

**Table 2. Association of composite healthy lifestyle with multisystem sequelae, death, and hospitalization following SARS-CoV-2 infection in key clinical subgroups**

	Multisystem sequelae		Death		Hospitalization	
	HR (95% CI)	P for interaction	HR (95% CI)	P for interaction	HR (95% CI)	P for interaction
<b>Age, y</b>						
<65	0.74 (0.64-0.85)	0.08	0.58 (0.39-0.85)	0.06	0.86 (0.76-0.96)	0.47
	0.51 (0.44-0.59)	<.001	0.39 (0.25-0.61)	0.01	0.69 (0.61-0.78)	0.03
≥65	0.84 (0.76-0.93)		0.77 (0.68-0.88)		0.88 (0.81-0.97)	
	0.70 (0.64-0.78)		0.61 (0.54-0.70)		0.83 (0.76-0.91)	
<b>Sex</b>						
Female	0.78 (0.69-0.89)	0.30	0.63 (0.51-0.78)	0.04	0.82 (0.73-0.92)	0.42
	0.58 (0.51-0.67)	0.02	0.52 (0.42-0.64)	0.09	0.73 (0.65-0.81)	0.14
Male	0.80 (0.72-0.89)		0.81 (0.70-0.93)		0.90 (0.82-0.99)	
	0.69 (0.61-0.77)		0.62 (0.53-0.73)		0.83 (0.75-0.91)	
<b>Ethnicity</b>						
White	0.75 (0.69-0.82)	0.01	0.74 (0.65-0.85)	0.67	0.84 (0.78-0.90)	0.04
	0.59 (0.54-0.64)	<.001	0.58 (0.50-0.67)	0.36	0.73 (0.68-0.79)	<.001
Other ethnic groups	1.10 (0.86-1.42)		0.78 (0.61-1.01)		1.01 (0.81-1.25)	
	0.93 (0.72-1.21)		0.65 (0.50-0.85)		1.02 (0.82-1.27)	
<b>Vaccine status<sup>b</sup></b>						
No or partial vaccination	0.88 (0.78-0.99)	0.13	0.85 (0.74-0.98)	0.02	0.89 (0.80-0.99)	0.63
	0.72 (0.63-0.81)	0.21	0.69 (0.60-0.80)	0.04	0.78 (0.70-0.87)	0.19
Full vaccination	0.75 (0.67-0.84)		0.62 (0.49-0.79)		0.86 (0.78-0.95)	
	0.60 (0.53-0.67)		0.53 (0.41-0.68)		0.79 (0.72-0.87)	
<b>Test setting</b>						
Inpatient	0.80 (0.73-0.88)	0.91	0.80 (0.67-0.96)	0.87	0.89 (0.82-0.97)	0.07
	0.64 (0.58-0.70)	0.42	0.64 (0.53-0.77)	0.33	0.80 (0.74-0.87)	0.13
Community/outpatient	0.83 (0.70-0.99)		0.78 (0.66-0.92)		0.81 (0.68-0.95)	
	0.75 (0.63-0.90)		0.69 (0.58-0.83)		0.81 (0.68-0.96)	
<b>Variants<sup>a</sup></b>						
Wildtype, alpha, and delta	0.76 (0.69-0.84)	0.07	0.77 (0.68-0.88)	0.56	0.80 (0.73-0.87)	0.01
	0.63 (0.57-0.70)	0.27	0.60 (0.53-0.69)	0.44	0.76 (0.69-0.83)	0.06
Omicron BA.1	0.94 (0.80-1.09)		0.70 (0.51-0.95)		1.05 (0.92-1.20)	
	0.72 (0.61-0.84)		0.69 (0.51-0.95)		0.89 (0.78-1.01)	

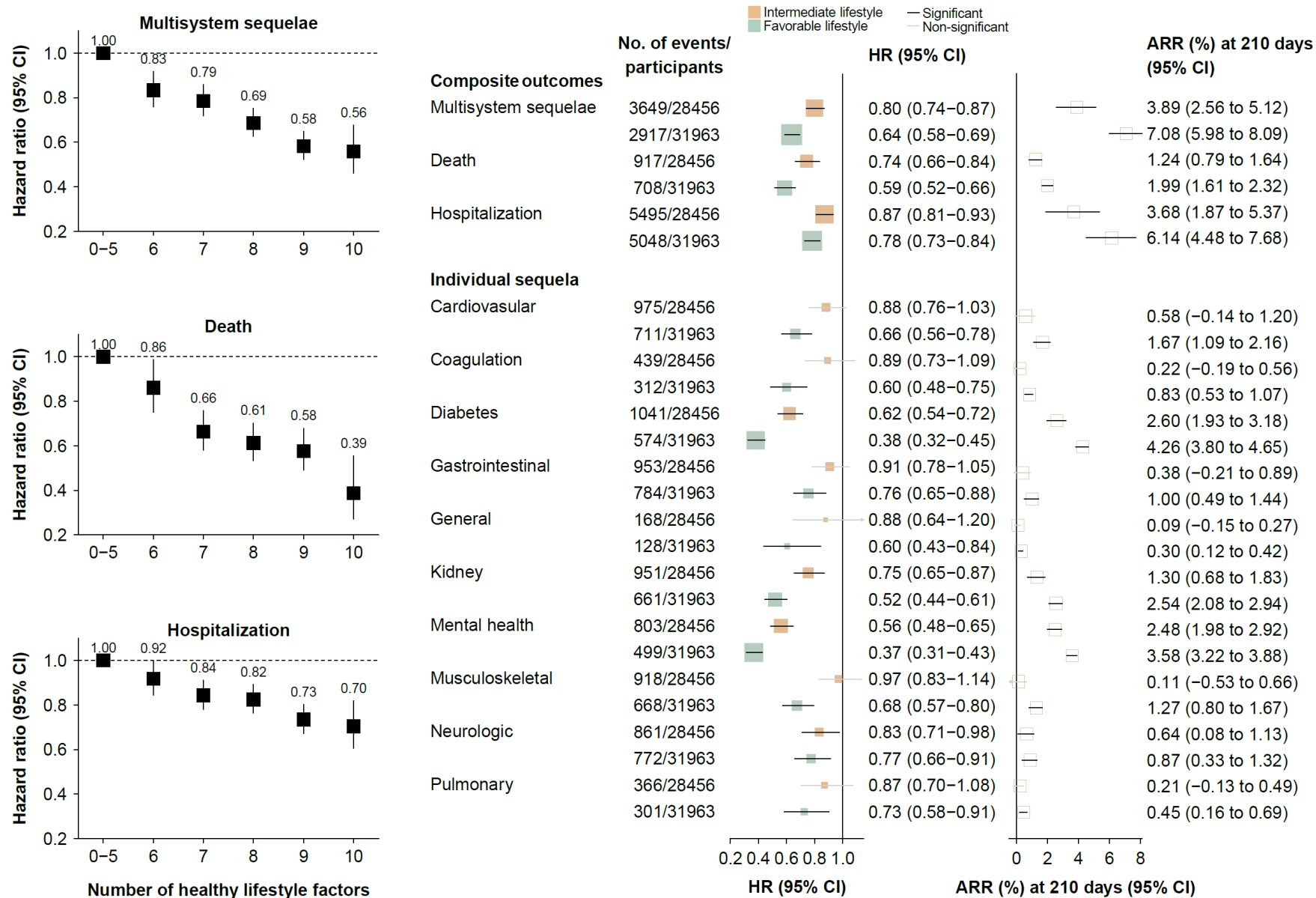
The first row of each category of the subgroup refers to intermediate vs unfavorable lifestyle; and the second row refer to favorable vs unfavorable lifestyle.

<sup>a</sup>The calendar periods of dominant variants in UK were based on pandemic data from the Office for National Statistics: wildtype (March 1, 2020–December 7, 2020), alpha [B.1.1.7] (December 8, 2020– May 17, 2021), delta [B.1.617.2] (May 18, 2021–December 13, 2021), Omicron [B.1.1.529 BA.1] (December 14, 2021–March 1, 2022).

<sup>b</sup>Subgroup analysis by vaccine status was conducted in participants from England, as vaccination data for Scotland and Wales were not available.

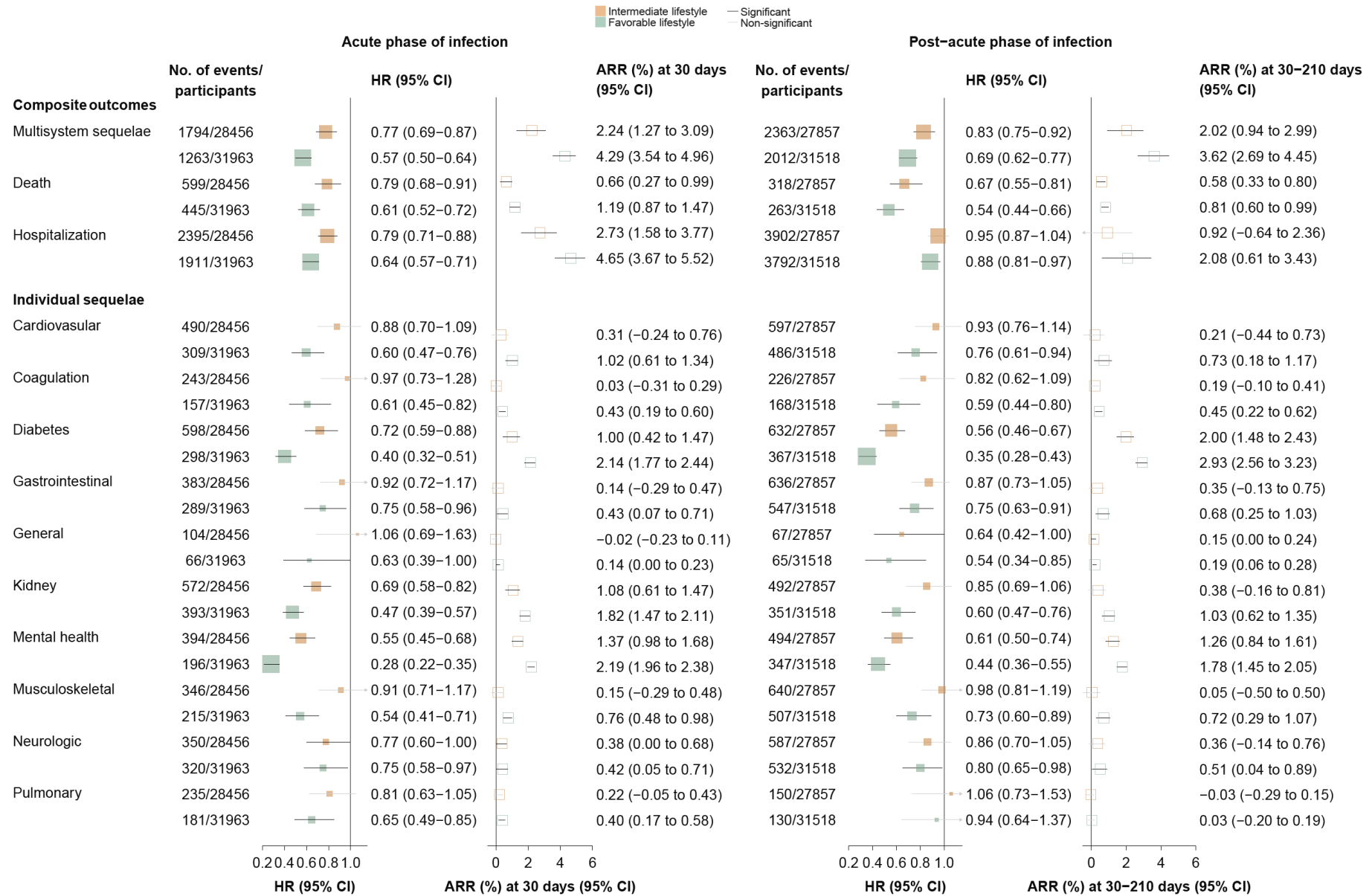
**Figure 1. Association of healthy lifestyle with multisystem sequelae of COVID-19, death, and hospital admissions, during overall, acute, and post-acute phases of SARS-CoV-2 infection**

**A Overall phase**



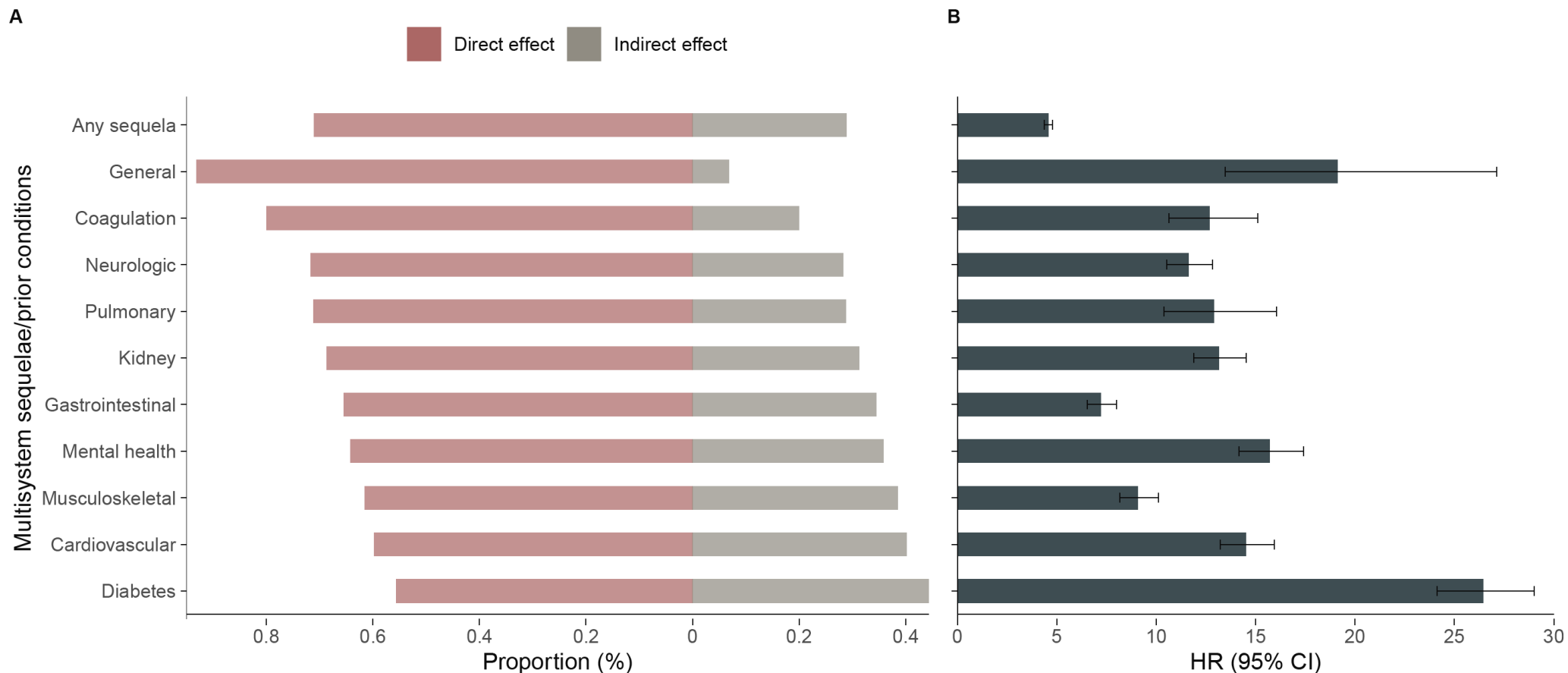


## B Acute and post-acute phases



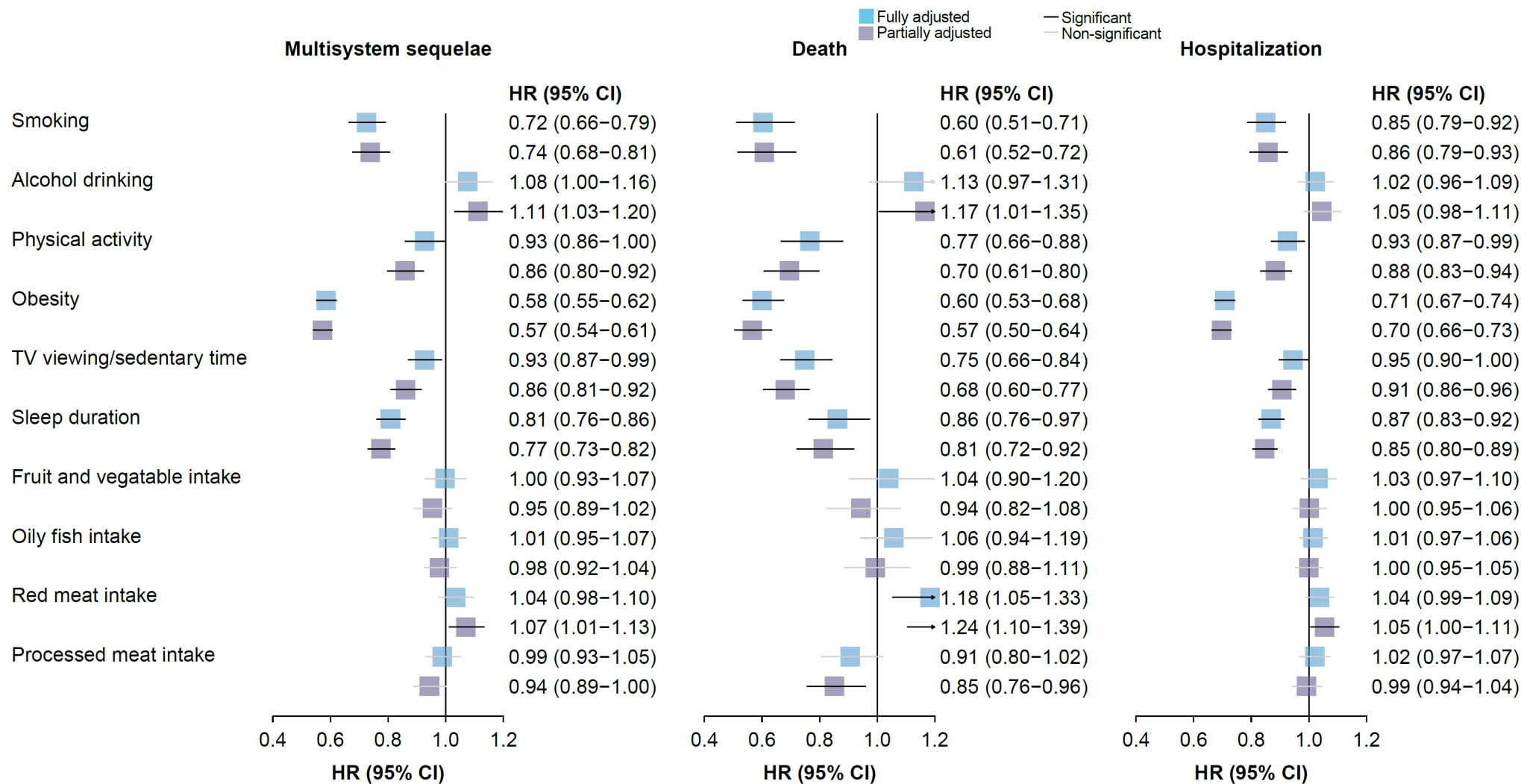
A, Healthy lifestyle (composite or number) and risk of multisystem sequelae (composite or by organ systems), death, and hospitalization during the overall phase (0-210 days) of SARS-CoV-2 infection. B, Composite healthy lifestyle and risk of multisystem sequelae, death, and hospitalization during the acute phase (first 30 days) and post-acute (30-210 days) phases of SARS-CoV-2 infection. Adjusted RRs and 95% CI are presented for composite multisystem sequelae, and adjusted HRs and 95% CIs for other outcomes. Absolute risk reduction per 100 persons at 210 days, 30 days, and 30-210 days and 95% CI were calculated. Solid square represents HRs with the area inversely proportional to the variance of the log HR. Hollow square represents ARR. The horizontal lines indicate 95% CIs, with black line representing statistically significant results and grey line representing non-significant results. Intermediate lifestyle category are in orange, favorable lifestyle category in green.

**Figure 2. Direct and indirect effects of healthy lifestyle, and association of pre-infection medical conditions with multisystem sequelae of COVID-19**



A, Proportion of the direct and indirect effect of healthy lifestyle on multisystem sequelae (Intermediate/favorable vs unfavorable lifestyle). Direct association were accounted for pre-infection medical conditions (mediator), identified as any relevant event recorded between baseline measurement and infection date. B, Association of corresponding pre-infection medical conditions with risk of sequelae following SARS-CoV-2 infection. Outcomes were ascertained 0-210 days after SARS-CoV-2 infection.

**Figure 3. Association of individual components of composite healthy lifestyle with multisystem sequelae, death, and hospitalization**



Blue square represents risk estimates from models fully adjusted for age, sex, education level, ethnicity, IMD, and mutually for all lifestyle factors. Purple square represents risk estimates from models partially adjusted for age, sex, education level, ethnicity, and IMD. The horizontal lines indicate 95% CIs, with black line representing statistically significant results and grey line representing non-significant results.

## Supplemental Online Content

### Healthy lifestyle for the prevention of post-COVID-19 multisystem sequelae, hospitalization, and death: a prospective cohort study

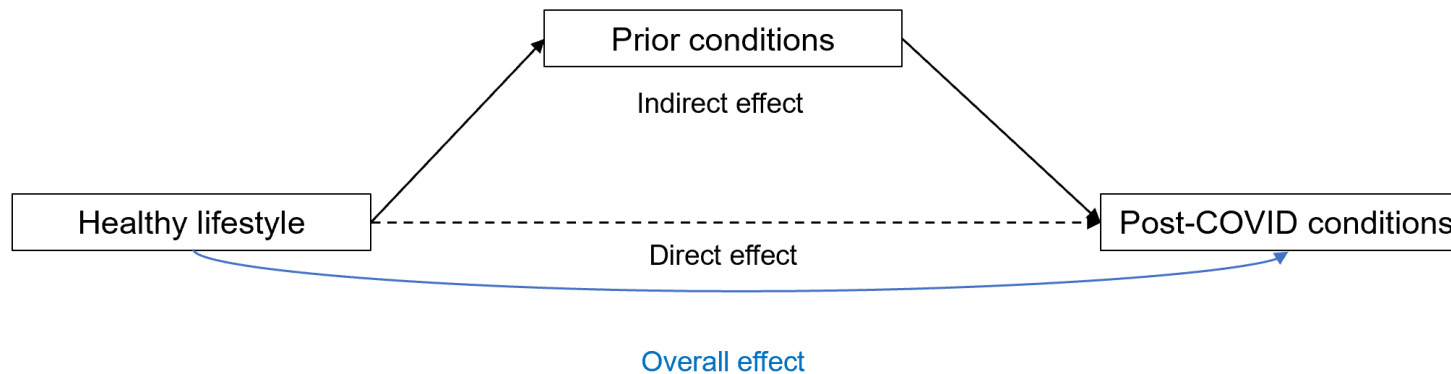
<b>eMethods</b>
<b>Supplementary Tables</b>
eTable 1. Detailed definitions on measurement and classification of lifestyle factors
eTable 2. Distributions of lifestyle score and categories
eTable 3. Definitions and weights of multisystem sequelae
eTable 4. Sensitivity analyses of the risk of composite multisystem sequelae, death, and hospitalization
<b>Supplementary Figures</b>
eFigure 1. Study design, cohort construction, and timeline.
eFigure 2. Change in lifestyle factors between baseline and the latest repeat assessment for those who undertook both visits
eFigure 3. Cumulative incidence curves of composite multisystem sequelae, death, and hospitalization following SARS-CoV-2 infection

## eMethods

### Causal mediation analysis

We used causal mediation analysis to specifically evaluate the extent to which a habitual healthy lifestyle may affect an COVID-19 sequelae outcome through a potential pathway of relevant medical conditions (mediator) prior to the infection. The mediator of interest was defined as relevant events that occurred between the baseline lifestyle assessment and the date of COVID-19 diagnosis. The statistical analysis was executed in a four-step process: (1) we first fitted a logistic regression model with the mediator as dependent variables and all covariates from the primary analysis as independent variables. (2) another logistic regression model was then fitted with multisystem sequelae as the dependent variable and the corresponding mediator along with other covariates as independent variables. (3) the direct effects (DE), indirect effects and their 95% confidence intervals (IE) were estimated using quasi-Bayesian Monte Carlo methods with 1,000 simulations for each. (4) finally, the mediation proportion was computed as  $IE/(IE+DE)*100$ .

Directed acyclic graph (DAG) for identifying confounding variables:



### Quantitative sensitivity analysis

Quantitative sensitivity analysis was used to adjust for changes in lifestyle factors over time since the baseline assessment. We used a standard algebraic approach that recalculates the expected cell frequencies for a correctly classified 2 x 2 contingency table based on the observed frequencies in a misclassified table. We then derived the bias-adjusted odds ratio using the formula  $A*D/B*C$ , assuming non-differential sensitivity and specificity both at 0.9. The equations used to calculate the expected true data, considering exposure misclassification, are as follows:

	Observed		Misclassification adjusted data	
	<i>E1</i>	<i>E0</i>	<i>E1</i>	<i>E0</i>
<i>D1</i>	<i>a</i>	<i>b</i>	$[a - D1 \text{ Total} (1 - SPD1)] / (SED1 - 1 + SPD1)$	<i>D1 Total - A</i>
<i>D0</i>	<i>c</i>	<i>d</i>	$[c - D0 \text{ Total} (1 - SPD0)] / (SED0 - 1 + SPD0)$	<i>D0 Total - C</i>
Total	<i>a + c</i>	<i>b + d</i>	<i>A + C</i>	<i>B + D</i>

Here, SED1 and SED0 denote the sensitivity in cases and non-cases, respectively, while SPD1 and SPD0 represent specificity in cases and non-cases, respectively.

**eTable 1. Detailed definitions on measurement and classification of lifestyle factors**

Modifiable lifestyle factor	Source and guidelines	UK Biobank questionnaire/definitions	Binary category	
			Healthy (1 point)	Unhealthy (0 point)
<b>Smoking status</b>	UK national health guidelines recommend smoking cessation for better health: such as <a href="#">National Health Service (NHS) Live Well guideline</a> : Quit smoking and <a href="#">guidance QS82 from NICE</a> : Reducing and preventing tobacco use	"Do you smoke tobacco now?" and "In the past, how often have you smoked tobacco?"	Past or never smoker	Current smoker
<b>Alcohol consumption</b>	<a href="#">NHS Live Well guideline</a> : alcohol advice recommends to drink no more than 14 units of alcohol a week, spread across 3 days or more and there's no completely safe level of drinking. <a href="#">The US 2020-2025 Dietary Guidelines</a> for Americans suggest that adults of legal drinking age can choose not to drink or to drink in moderation by limiting intake to 2 drinks or less in a day for men and 1 drink or less in a day for women when alcohol is consumed.	"About how often do you drink alcohol?"	≤ 4 times week	Daily or almost daily
<b>BMI</b>	<a href="#">NHS Live Well guideline</a> and a recent <a href="#">meta-analysis on risk factors for long COVID</a> suggest that obesity is a serious health concern that increases the risk of many health outcomes, including post-COVID conditions.	Height was measured in whole centimetres with a Seca 202 device. Weight was measured to the nearest 0.1 kg. BMI was calculated as weight in kilograms divided by the square of height in meters.	BMI <30 kg/m <sup>2</sup>	BMI ≥ 30 kg/m <sup>2</sup>
<b>Physical activity</b>	<a href="#">WHO guidelines on physical activity</a> and <a href="#">NHS Live Well guideline</a> suggest moderate- and vigorous-intensity physical activity improve health. All adults should undertake 150-300 min of moderate-intensity, or 75-150 min of vigorous-intensity physical activity, or some equivalent combination of moderate-intensity and vigorous-intensity aerobic physical activity, per week.	IPAQ short form 2 <sup>a</sup> – total time walking or moderate and vigorous-intensity PA in previous week	≥150 min/week moderate or ≥ 75 min/week vigorous PA	< 75 min/week vigorous PA
<b>TV viewing/sedentary time</b>	<a href="#">WHO guidelines on sedentary behaviour</a> and <a href="#">NHS Live Well guideline</a> : reducing sedentary behaviours is recommended across all age groups and abilities	"In a typical day, how many hours do you spend watching TV?" <sup>b</sup>	< 4 h/day	≥ 4 h/day
<b>Sleep duration</b>	In accordance with <a href="#">NHS Live Well guideline</a> , <a href="#">American Academy of Sleep Medicine (AASM) and Sleep Research Society (SRS) recommendation</a> , sleep duration of 7 to 9 hours was classified into healthy lifestyle category.	"About how many hours sleep do you get in every 24 hours?"	7-9 h/day	<7 or >9h/day

<b>Fruit and vegetable intake</b>	Consuming a healthy diet throughout the life-course helps to prevent malnutrition in all its forms as well as a range of noncommunicable diseases and conditions. <a href="#">NHS Eat Well guideline</a> , <a href="#">WHO guideline on healthy diet</a> , and <a href="#">World Cancer Research Fund International</a>	"About how many of .... would you eat per day?" Separate questions for pieces of fresh and dried fruit, tablespoons of salad or cooked/raw vegetables. Combined and converted to g/day (1 portion = 80 g)	≥ 400 g/day	<400 g/day
<b>Oily fish intake</b>	suggest that at least 400 g, or five portions, of fruit and vegetables per day; at least 2 portions of fish a week, including 1 of oily fish; limit consumption to red meat and processed meat (though threshold varied by guidelines: NHS suggests keep intake of red and processed meat to 70g or less a day; WCRF suggest limit consumption to no more than about three portions per week (equivalent to about 350–500g (about 12–18oz) cooked weight.) and consume very little, if any, processed meat).	"How often do you eat oily fish? (e.g. sardines, salmon, mackerel, herring)"	≥1 portion/week	<1 portion/week
<b>Red meat intake</b>		"How often do you eat...?" Separate questions for Beef / lamb or mutton / pork (excluding processed meats such as ham or bacon). Red meat included due to clear link between red meat and mortality.	≤3 portion/week	>3 portion/week
<b>Processed meat intake</b>		"How often do you eat processed meats (such as bacon, ham, sausages, meat pies, kebabs, burgers, chicken nuggets)?"	≤1 portion/week	>1 portion/week

<sup>a</sup>UK Biobank physical activity (PA) data were analyzed in accordance with the International Physical Activity Questionnaire (IPAQ) scoring protocol, with total physical activity computed as the sum of walking, moderate, and vigorous activity, measured as metabolic equivalents (MET-hours/week).

<sup>b</sup>UK Biobank participants were asked about work-related sitting time and were not asked about discretionary/leisure time sedentary behavior. Therefore, TV viewing time, which is strongly associated with health outcomes, was used instead of sedentary time. UK Biobank participants were asked separately about time spent using a computer, excluding computer time at work. This information was used to check for implausible values for TV viewing time.

**eTable 2. Distributions of lifestyle score and categories**

**A, All eligible participants**

Lifestyle category <sup>a</sup>	Lifestyle score	No. of participants (N=472,977)		Percentage, %	
Unfavorable	0	6	55400	0.00%	11.7%
	1	97		0.02%	
	2	777		0.16%	
	3	3949		0.83%	
	4	13914		2.94%	
	5	36657		7.75%	
Intermediate	6	75141	192407	15.89%	40.7%
	7	117266		24.79%	
Favorable	8	127905	225170	27.04%	47.6%
	9	80470		17.01%	
	10	16795		3.55%	

**B, COVID-19 cohort**

Lifestyle category <sup>a</sup>	Lifestyle score	No. of participants (N=68,896)		Percentage, %	
Unfavorable	0	1	8476	0.00%	12.3%
	1	18		0.03%	
	2	122		0.18%	
	3	639		0.93%	
	4	2142		3.11%	
	5	5554		8.06%	
Intermediate	6	11258	28457	16.34%	41.3%
	7	17199		24.96%	
Favorable	8	18182	31963	26.39%	46.4%
	9	11273		16.36%	
	10	2508		3.64%	

<sup>a</sup>A binary variable was created for each of the 10 lifestyle factors, with 1 point assigned for healthy category and 0 otherwise. A composite lifestyle score was then calculated for each participant by summing the total number of unhealthy lifestyle factors, with a range of 0 to 10. Based on the composite score, participants were classified into three lifestyle categories: unfavorable (0-5), intermediate (6-7), and favorable (8-10).



**eTable 3. Definitions and weights of multisystem sequelae**

Organ system	Sequela	Cause of death or disability	ICD-10 code	Weight
Cardiovascular	Acute coronary disease	Ischemic heart disease	I24	1.009
Cardiovascular	Myocardial infarction	Ischemic heart disease	I21,I22	1.009
Cardiovascular	Atrial fibrillation	Atrial fibrillation and flutter	I480,I481,I482	0.124
Cardiovascular	Atrial flutter	Atrial fibrillation and flutter	I483,I484	0.124
Cardiovascular	Angina	Chest pain	I20	0.138
Cardiovascular	Bradycardia	Atrial fibrillation and flutter	R001	0.124
Cardiovascular	Cardiac arrest	Ischemic heart disease	I46	1.009
Cardiovascular	Non-ischemic cardiomyopathy	Other cardiomyopathy	I42,I43,B332	0.740
Cardiovascular	Cariogenic shock	Ischemic heart disease	R570	1.009
Cardiovascular	Heart failure	Ischemic heart disease	I50	1.009
Cardiovascular	Ischemic cardiomyopathy	Other cardiomyopathy	I255	0.740
Cardiovascular	Myocarditis	Myocarditis	I514	0.584
Cardiovascular	Pericarditis	Myocarditis	I30,I311,I312,I313,I318,I319,I32	0.584
Cardiovascular	Tachycardia	Atrial fibrillation and flutter	R000	0.124
Cardiovascular	Ventricular arrhythmias	Atrial fibrillation and flutter	I490,I470,I471,I472	0.124
Coagulation	Anemia	Anemia	D60,D61,D63,D64,D62	0.002
Coagulation	Deep vein thrombosis	Stroke	I824,I825	0.540
Coagulation	Coagulation defect	Stroke	D689	0.540
Coagulation	Pulmonary embolism	Stroke	I26	0.540
Coagulation	Venous thrombotic embolism	Stroke	I820,I821,I822,I823,I826,I827,I82A,I82B,I82C,I828,I829	0.540
Diabetes	Diabetes	Diabetes mellitus type 2	E10,E11	0.114
Gastrointestinal	Abdominal pain	Abdominal pain	R100,R101,R103,R108,R109	0.124
Gastrointestinal	Acute gastritis	Gastritis and duodenitis	K290	0.047
Gastrointestinal	Acute pancreatitis	Pancreatitis	K85	0.503
Gastrointestinal	Cholangitis	Gallbladder and biliary diseases	K830	0.031
Gastrointestinal	Constipation	Constipation	K590	0.022
Gastrointestinal	Diarrhea	Moderate diarrheal diseases	K591,R197	0.188
Gastrointestinal	Gastroesophageal reflux disease	Gastroesophageal reflux disease	K21	0.008
Gastrointestinal	Irritable bowel syndrome	Inflammatory bowel disease	K55	0.282
Gastrointestinal	Liver diseases	Total burden related to Non-alcoholic fatty liver disease	K70,K71,K73,K74,K7,K721,Z944,B18,I850,I859,I982	0.005
Gastrointestinal	Peptic ulcer disease	Peptic ulcer disease	K25,K26,K27,K28	0.214
Gastrointestinal	Vomiting	Vomitting	R111	0.223
General	Fatigue and malaise	Fatigue and malaise*	R53	0.215
Kidney	Acute kidney injury	Acute kidney disease	N17,N19	0.605

Kidney	Chronic Kidney Disease	Chronic kidney disease	N18	0.057
Mental health	Psychotic disorders	Schizophrenia	F20,F21,F22,F23,F24,F25,F26,F27,F28,F29	0.049
Mental health	Mania/Bipolar affective disorder	Depressive disorders	F30,F31	0.136
Mental health	Depressive episode	Depressive disorders	F32	0.136
Mental health	Phobic anxiety disorder	Anxiety disorders	F40	0.093
Mental health	Panic disorder	Anxiety disorders	F410	0.093
Mental health	Generalized anxiety disorder	Anxiety disorders	F411	0.093
Mental health	Posttraumatic stress disorder	Anxiety disorders	F431	0.093
Mental health	Substance use disorders	Drug use disorders	F10,F11,F12,F13,F14,F15,F16,F17,F18,F19	0.621
Mental health	Sleep disorders	Fatigue*	F51,G47	0.215
Musculoskeletal	Joint pain	Musculoskeletal disorders	M255	0.110
Musculoskeletal	Myalgias	Musculoskeletal disorders	M791	0.110
Musculoskeletal	Myopathy	Musculoskeletal disorders	G72	0.110
Musculoskeletal	Osteoarthritis	Musculoskeletal disorders	M15,M16,M17,M18,M19	0.110
Neurologic	Abnormal involuntary movements	Parkinson's disease	R25	0.796
Neurologic	Alzheimers disease	Alzheimer's disease and other dementias	G30	0.413
Neurologic	Memory problems	Alzheimer's disease and other dementias	R413	0.413
Neurologic	Central Venous thrombosis	Stroke	I636	0.540
Neurologic	Dizziness	Fatigue*	R42	0.215
Neurologic	Dysautonomia	Neuropathy	G900,G903,G904,G905,G908,G909	0.133
Neurologic	Dystonia	Parkinson's disease	G24	0.796
Neurologic	Bells palsy	Neuropathy	G510	0.133
Neurologic	Headache disorders	Tension-type headache	G44	0.002
Neurologic	Loss of hearing	Age-related and other hearing loss	H90,H91,H93	0.031
Neurologic	Migraine	Migraine	G43	0.037
Neurologic	Neurocognitive decline	Cognitivie symptoms*	G454,F440,F04, F01,F02,G311,G318,G319	0.128
Neurologic	Paresthesia	Neuropathy	R202	0.133
Neurologic	Parkinsons-like disease	Parkinson's disease	G20	0.796
Neurologic	Polyneuropathy	Neuropathy	G611,G618,G619,G62	0.133
Neurologic	Epilepsy and seizures	Idiopathic epilepsy	G40	0.303
Neurologic	Loss of smell	Other sense organ diseases	R430,R431	0.023
Neurologic	Somnolence	Fatigue*	R400	0.215
Neurologic	Ischemic stroke	Stroke	G46,I63	0.540
Neurologic	Loss of taste	Other sense organ diseases	R439	0.023
Neurologic	Transient cerebral ischemic attack	Transient cerebral ischemic attack	G45	0.250
Neurologic	Tremor	Parkinson's disease	R251	0.796

Neurologic	Vision abnormalities	Vision problems	H53,H54,H34,H356	0.036
Pulmonary	Cough	Mild and moderate respiratory disease*	R05	0.138
Pulmonary	Hypoxemia	Severe respiratory disease*	R090	0.410
Pulmonary	Interstitial lung disease	Interstitial lung disease	J841, J848, J849	0.112
Pulmonary	Shortness of breath	Mild and moderate respiratory disease*	R060	0.138

Health burden coefficients were obtained from [T. Vos et al., Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 396, 1204-1222 \(2020\)](#) and other GBD sources: [\\*Health burden weights were obtained from S. W. Hanson et al., Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. Jama 328, 1604-1615 \(2022\)](#). In sensitivity analysis, the weighted score was calculated for each participant by summing the weights of all incident sequelae during the follow-up period.

The selection and definition of prespecified multisystem sequelae of COVID-19 were based on literature review and previous knowledge:

- High-dimensional characterization of post-acute sequelae of COVID-19. Nature
- Risk of clinical sequelae after the acute phase of SARS-CoV-2 infection: retrospective cohort study. BMJ
- Long-term cardiovascular outcomes of COVID-19. Nat Med
- Risks and burdens of incident diabetes in long COVID: a cohort study. Lancet Diabetes Endocrinol
- Risks and burdens of incident dyslipidaemia in long COVID: a cohort study. Lancet Diabetes Endocrinol
- Long-term neurologic outcomes of COVID-19. Nat Med
- Risks of mental health outcomes in people with COVID-19. BMJ
- Kidney outcomes in long COVID. J. Am. Soc. Nephrol.
- Long-term gastrointestinal outcomes of COVID-19. Nat Commun
- Postacute sequelae of COVID-19 at 2 years. Nat Med

**eTable 4. Sensitivity analyses of the risk of composite multisystem sequelae, death, and hospitalization**

**A,**

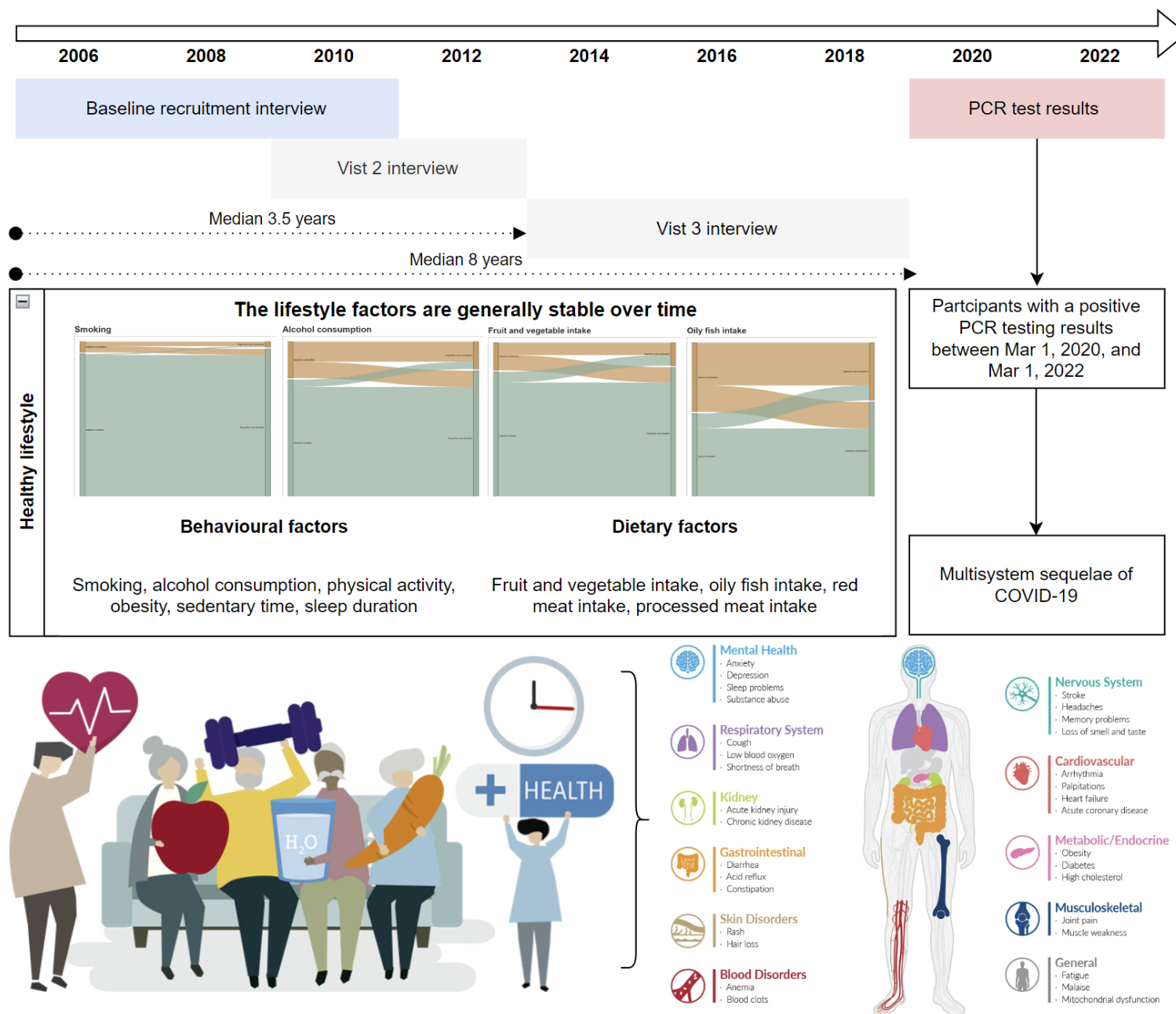
	<b>Multisystem sequelae (HR and 95% CI)</b>		<b>Death (HR and 95% CI)</b>		<b>Hospitalization (HR and 95% CI)</b>	
	Intermediate	Favorable	Intermediate	Favorable	Intermediate	Favorable
<b>Assigning weight to each individual sequela and using zero inflated Poisson regression</b>	0.78 (0.70-0.88)	0.56 (0.50-0.64)	NA	NA	NA	NA
<b>Extending the washout period from one year to two years</b>	0.80 (0.73-0.88)	0.65 (0.59-0.72)	0.74 (0.66-0.84)	0.59 (0.52-0.66)	0.90 (0.83-0.98)	0.83 (0.76-0.90)
<b>Defining events of post-acute sequelae 90 days after infection</b>	0.83 (0.73-0.95)	0.71 (0.62-0.81)	0.66 (0.49-0.88)	0.57 (0.42-0.77)	0.95 (0.86-1.03)	0.88 (0.81-0.97)
<b>Using only the first three ICD disease diagnoses to identify outcomes</b>	0.73 (0.62-0.86)	0.64 (0.54-0.76)	0.81 (0.70-0.94)	0.59 (0.50-0.69)	0.75 (0.67-0.85)	0.59 (0.52-0.67)

**B, account for potential misclassification of lifestyle factors<sup>a</sup>**

	<b>Multisystem sequelae (OR and 95% CI)</b>		<b>Death (OR and 95% CI)</b>		<b>Hospitalization (OR and 95% CI)</b>	
	Intermediate	Favorable	Intermediate	Favorable	Intermediate	Favorable
<b>Without adjustment</b>	0.75 (0.69-0.81)	0.54 (0.50-0.59)	0.74 (0.69-0.79)	0.50 (0.47-0.54)	0.74 (0.67-0.81)	0.53 (0.48-0.58)
<b>Adjusted for potential misclassification</b>	0.70 (0.63-0.77)	0.47 (0.42-0.52)	0.69 (0.63-0.74)	0.43 (0.39-0.47)	0.68 (0.61-0.77)	0.45 (0.40-0.51)

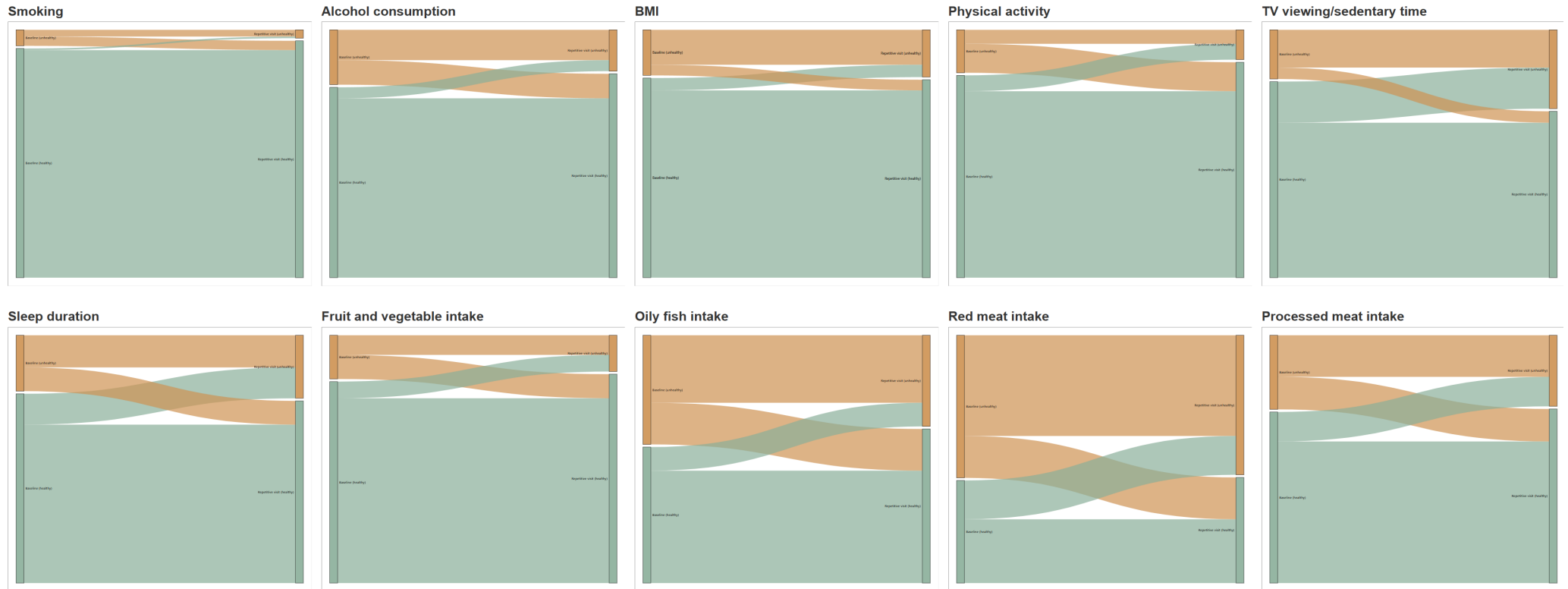
A, qualitative sensitivity analyses. B, quantitative sensitivity analysis accounting for potential misclassification of lifestyle factors over time. <sup>a</sup>Odds ratios were used to quantify associations and assumed a sensitivity and specificity of 90% for each lifestyle component.

eFigure 1. Study design, cohort construction, and timeline

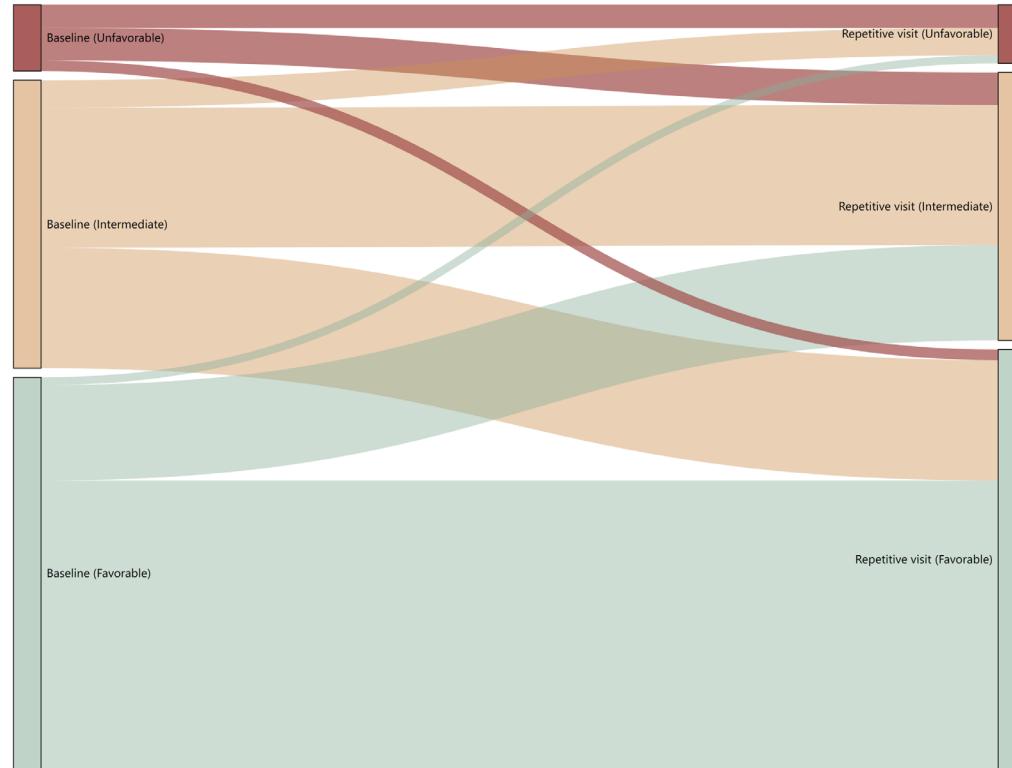


eFigure 2. Change in lifestyle factors between baseline and the latest repeat assessment for those who undertook both visits

A, Change in individual lifestyle factor (unhealthy and healthy)

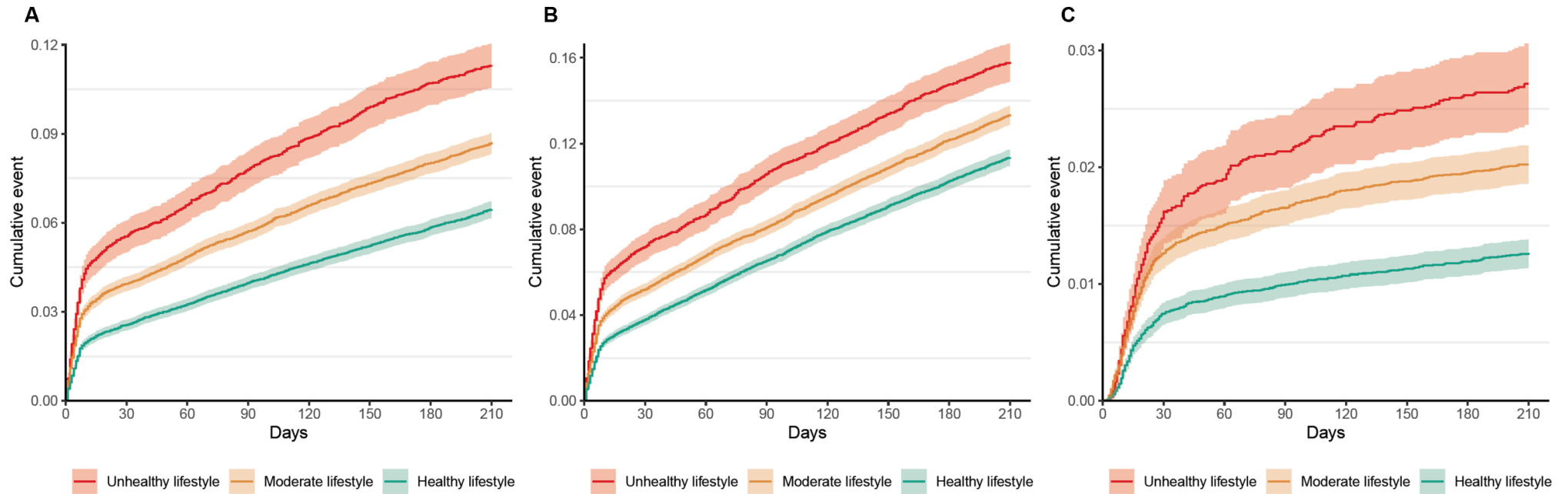


**B, Change in composite lifestyle category (unfavorable, intermediate, and favorable)**



A, Change in 10 individual lifestyle factor between baseline and the latest repeat assessment. B, Change in composite lifestyle category between baseline and the latest repeat assessment. Left column of each plot indicates baseline and right column indicates repeat assessment. For each individual lifestyle factor as binary variable, unhealthy category is in orange, favorable category in green. For each the composite lifestyle as categorical variable, unfavorable category is in red, intermediate category in yellow, and favorable category in green. 34.9% of participants with an unfavorable lifestyle, 48.6% with an intermediate lifestyle, and 73.7% with a favorable lifestyle at baseline remained in the same corresponding lifestyle category at the latest repeat assessment ~8 years later (overall proportion of stable categories, 60.6%). The sample size for each factor varied by missing conditions (N=~60,000).

eFigure 3. Cumulative incidence curves of composite multisystem sequelae, death, and hospitalization following SARS-CoV-2 infection



A, Composite multisystem sequelae. B, Death. C, Hospitalization. Outcomes were ascertained 0-210 days after SARS-CoV-2 infection. Event rates presented for the unhealthy lifestyle category (red), the intermediate lifestyle category (orange), and the favorable lifestyle category (green). The shadow of curves represents 95% confidence interval.



## Reference

1. Parotto M, Gyöngyösi M, Howe K, et al. Post-acute sequelae of COVID-19: understanding and addressing the burden of multisystem manifestations. *Lancet Respir Med*. Aug 2023;11(8):739-754. doi:10.1016/s2213-2600(23)00239-4
2. Munblit D, Nicholson T, Akrami A, et al. A core outcome set for post-COVID-19 condition in adults for use in clinical practice and research: an international Delphi consensus study. *Lancet Respir Med*. Jul 2022;10(7):715-724. doi:10.1016/s2213-2600(22)00169-2
3. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nature Reviews Microbiology*. 2023;21(3):133-146.
4. Al-Aly Z, Agarwal A, Alwan N, Luyckx VA. Long COVID: long-term health outcomes and implications for policy and research. *Nature Reviews Nephrology*. 2023;19(1):1-2.
5. Bowe B, Xie Y, Al-Aly Z. Postacute sequelae of COVID-19 at 2 years. *Nature Medicine*. 2023:1-11.
6. Ballouz T, Menges D, Anagnostopoulos A, et al. Recovery and symptom trajectories up to two years after SARS-CoV-2 infection: population based, longitudinal cohort study. *bmj*. 2023;381
7. Ledford H. Long-COVID treatments: why the world is still waiting. *Nature*. 2022;608(7922):258-260.
8. Shaffer L. Lots of long COVID treatment leads, but few are proven. *Proceedings of the National Academy of Sciences*. 2022;119(36):e2213524119.
9. Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nature medicine*. 2022;28(7):1461-1467.
10. Català Martí M-BN, Kolde Raivo, et al. The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: staggered cohort analyses of data from the UK, Spain, and Estonia. *Lancet Respir Med*. 2023;doi:In press
11. Xie Y, Choi T, Al-Aly Z. Association of treatment with nirmatrelvir and the risk of post-COVID-19 condition. *JAMA Internal Medicine*. 2023;183(6):554-564.
12. Xie Y, Choi T, Al-Aly Z. Molnupiravir and risk of post-acute sequelae of covid-19: cohort study. *bmj*. 2023;381
13. Patnode CD, Redmond N, Iacocca MO, Henninger M. Behavioral counseling interventions to promote a healthy diet and physical activity for cardiovascular disease prevention in adults without known cardiovascular disease risk factors: updated evidence report and systematic review for the us preventive services task force. *Jama*. 2022;328(4):375-388.
14. Nyberg ST, Singh-Manoux A, Pentti J, et al. Association of healthy lifestyle with years lived without major chronic diseases. *JAMA internal medicine*. 2020;180(5):760-768.
15. Firth J, Solmi M, Wootton RE, et al. A meta-review of "lifestyle psychiatry": the role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. *World psychiatry*. 2020;19(3):360-380.
16. Hamer M, Kivimäki M, Gale CR, Batty GD. Lifestyle risk factors, inflammatory mechanisms, and COVID-19 hospitalization: A community-based cohort study of 387,109 adults in UK. *Brain, behavior, and immunity*. 2020;87:184-187.
17. Ahmadi MN, Huang B-H, Inan-Eroglu E, Hamer M, Stamatakis E. Lifestyle risk factors and infectious disease mortality, including COVID-19, among middle aged and older adults: Evidence from a community-based cohort study in the United Kingdom. *Brain, behavior, and immunity*. 2021;96:18-27.
18. Perricone C, Versini M, Ben-Ami D, et al. Smoke and autoimmunity: The fire behind the disease. *Autoimmunity reviews*. 2016;15(4):354-374.
19. Sharif K, Watad A, Bragazzi NL, Lichtbroun M, Amital H, Shoenfeld Y. Physical activity and autoimmune diseases: Get moving and manage the disease. *Autoimmunity reviews*. 2018;17(1):53-72.
20. Ghouse J, Tragante V, Ahlberg G, et al. Genome-wide meta-analysis identifies 93 risk loci and enables risk prediction equivalent to monogenic forms of venous thromboembolism. *Nature Genetics*. 2023;55(3):399-409.
21. Gregson J, Kaptoge S, Bolton T, et al. Cardiovascular risk factors associated with venous thromboembolism. *JAMA cardiology*. 2019;4(2):163-173.
22. Tsampasian V, Elghazaly H, Chattopadhyay R, et al. Risk factors associated with Post- COVID-19 condition: a

systematic review and meta-analysis. *JAMA Internal Medicine*. 2023;

23. Loef M, Walach H. The combined effects of healthy lifestyle behaviors on all cause mortality: a systematic review and meta-analysis. *Preventive medicine*. 2012;55(3):163-170.
24. Office for National Statistics. Coronavirus (COVID-19) latest insights: Infections. Accessed Sept 9, 2023. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronavirus-covid19latestinsights/infections>
25. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-209.
26. Denaxas S, Gonzalez-Izquierdo A, Direk K, et al. UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER. *Journal of the American Medical Informatics Association*. 2019;26(12):1545-1559.
27. Kuan V, Denaxas S, Gonzalez-Izquierdo A, et al. A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. *The Lancet Digital Health*. 2019;1(2):e63-e77.
28. Armstrong J, Rudkin JK, Allen N, et al. Dynamic linkage of COVID-19 test results between Public Health England's second generation surveillance system and UK Biobank. *Microbial genomics*. 2020;6(7)
29. NHS. NHS advice about healthy living, including eating a balanced diet, healthy weight, exercise, quitting smoking and drinking less alcohol. Accessed Sept 1, 2013. <https://www.nhs.uk/live-well/>
30. Office for Health Improvement & Disparities. Healthy eating: applying All Our Health. Accessed Sept 1, 2023. <https://www.gov.uk/government/publications/healthy-eating-applying-all-our-health/healthy-eating-applying-all-our-health>
31. Foster HM, Celis-Morales CA, Nicholl BI, et al. The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. *The Lancet Public Health*. 2018;3(12):e576-e585.
32. Wang S, Li Y, Yue Y, et al. Adherence to healthy lifestyle prior to infection and risk of post-COVID-19 condition. *JAMA Internal Medicine*. 2023;183(3):232-241.
33. Xie Y, Xu E, Bowe B, Al-Aly Z. Long-term cardiovascular outcomes of COVID-19. *Nature medicine*. 2022;28(3):583-590.
34. Xie Y, Al-Aly Z. Risks and burdens of incident diabetes in long COVID: a cohort study. *The lancet Diabetes & endocrinology*. 2022;10(5):311-321.
35. Xu E, Xie Y, Al-Aly Z. Long-term gastrointestinal outcomes of COVID-19. *Nature communications*. 2023;14(1):983.
36. Bowe B, Xie Y, Xu E, Al-Aly Z. Kidney outcomes in long COVID. *Journal of the American Society of Nephrology*. 2021;32(11):2851-2862.
37. Xie Y, Xu E, Al-Aly Z. Risks of mental health outcomes in people with covid-19: cohort study. *bmj*. 2022;376
38. Xu E, Xie Y, Al-Aly Z. Long-term neurologic outcomes of COVID-19. *Nature medicine*. 2022;28(11):2406-2415.
39. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature*. 2021;594(7862):259-264.
40. Daugherty SE, Guo Y, Heath K, et al. Risk of clinical sequelae after the acute phase of SARS-CoV-2 infection: retrospective cohort study. *Bmj*. 2021;373
41. Smith T, Noble M, Noble S, Wright G, McLennan D, Plunkett E. The English indices of deprivation 2015. *London: Department for Communities and Local Government*. 2015:1-94.
42. Cashin AG, McAuley JH, VanderWeele TJ, Lee H. Understanding how health interventions or exposures produce their effects using mediation analysis. *BMJ (Clinical Research ed)*. 2023;382:e071757. doi:10.1136/bmj-2022-071757
43. Lee H, Cashin AG, Lamb SE, et al. A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement. *JAMA*. 2021;326(11):1045-1056. doi:10.1001/jama.2021.14075
44. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020;396(10258):1204-1222.
45. Hanson SW, Abbafati C, Aerts JG, et al. Estimated global proportions of individuals with persistent fatigue, cognitive,

- and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *Jama*. 2022;328(16):1604-1615.
46. Greenhalgh T, Sivan M, Delaney B, Evans R, Milne R. Long covid—an update for primary care. *bmj*. 2022;378
47. Luo S, Liang Y, Wong THT, Schooling CM, Au Yeung SL. Identifying factors contributing to increased susceptibility to COVID-19 risk: a systematic review of Mendelian randomization studies. *Int J Epidemiol*. 2022;51(4):1088-1105. doi:10.1093/ije/dyac076
48. Salman D, Vishnubala D, Le Feuvre P, et al. Returning to physical activity after covid-19. *bmj*. 2021;372
49. Florencio LL, Fernández-de-Las-Peñas C. Long COVID: systemic inflammation and obesity as therapeutic targets. *The Lancet Respiratory Medicine*. 2022;10(8):726-727. doi:10.1016/S2213-2600(22)00159-X
50. Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: a prospective observational study. *The Lancet Respiratory Medicine*. 2022;10(8):761-775. doi:10.1016/S2213-2600(22)00127-8
51. Lam ICH, Wong CKH, Zhang R, et al. Long-term post-acute sequelae of COVID-19 infection: a retrospective, multi-database cohort study in Hong Kong and the UK. *EClinicalMedicine*. 2023;60:102000. doi:10.1016/j.eclinm.2023.102000