Dear Professor Ross,

We are grateful to the editors for their comments, which we have addressed in detail below and in our revised manuscript.

We look forward to hearing from you with your decision.

Yours sincerely,
Daniel Ayoubkhani (on behalf of all co-authors)

Missing values are handled by the missing indicator approach (i.e., using an 'unknown' category) – this is widely considered as being potentially problematic, as it can lead to unpredictable bias outside a randomized controlled setting (Knol et al J Clin Epidemiol 2010; Henry et al J Vasc Surg 2013; Groenwold et al CMAJ 2012); Pedersen et al Clin Epidemiol 2017 and many more). There is a non-ignorable amount of missing data for some covariates. Please provide additional justification for using the missing indicator approach and address the limitations of this approach in your Discussion.

Thank you for this comment. We agree that the missing indicator approach has limitations, which we have emphasised in the discussion section of the revised draft, including the reference to the Groenwold paper you kindly supplied.

However, as we already state in the methods section, it was not practical to perform multiple imputation (or even single regression imputation) due to the size of the study dataset, comprising approximately 50 million individuals. In addition, missingness in variables such as BMI is likely to be informative when working with electronic health records, and it could be argued that these data are to some extent missing not at random (for example, individuals who are neither underweight nor overweight may be less likely to have their BMI measured), which would preclude the use of standard imputation techniques. We have added this justification to the discussion section of the revised draft.

Continuous predictors (age, BMI) handled by arbitrary (coarse) categorisation (losing information) for the main analysis, unclear why this was done, keeping them on their continuous scale would've been more appropriate, and not assuming a functional form and examined using splines. A sensitivity analysis was done using polynomials (for age) though not BMI (why not) – these results are not presented, if they are making claims that the results remain unchanged, then these should be presented in the supplementary material (unless I missed them). At the very least, please further justify these decisions.

Thank you for this comment. We used categorical rather than continuous BMI to facilitate the exact matching process; exact matching would not be possible with continuous matching variables. We have added a sentence to explain this in the methods section of the revised draft.

We started with the underweight (<18 kg/m^2), normal weight (18 to <25 kg/m^2), overweight (25 to <30 kg/m^2), and obese (≥30 kg/m^2) categorisation routinely used in healthcare delivery and clinical research. However, we had to coarsen this categorisation by grouping together individuals with BMI <25 kg/m^2 or unknown in order to ensure a suitable match rate. In sensitivity analysis, we conducted post-matching regression adjustment for age, BMI, and smoking status as these variables were all coarsened during matching. For BMI, we included

the non-coarsened version of the categorical variable in the regression model. The results of the sensitivity analysis can be found in section 2 of the supplementary appendix. We found that our main results are robust to possible residual confounding due to the coarsening of age, smoking status, and BMI during the matching process.

Some justification is needed for examining by <70 versus >70 years of age in the Methods. One would normally keep this on the continuous scale and handle with an interaction and possibly a spline. However, I appreciate interest in this cut-off might have been driven by other considerations.

We agree that keeping age on a continuous scale would usually be the preferred approach. However, in this instance the UK government has consistently communicated that individuals in the ≥70 age group are at higher risk of severe illness from COVID-19, hence the focus on this group in our paper. For example, see the UK government's definition of clinically vulnerable people in its social distance guidelines from May 2020: https://www.gov.uk/government/publications/full-guidance-on-staying-at-home-and-away-from-others/full-guidance-on-staying-at-home-and-away-from-others. We have included this justification in the method section of the revised draft.

The analyses examining white versus non-white, how were the 'unknowns' handled?

Individuals with missing ethnicity information were omitted from all analyses stratified by ethnic group. In the revised draft, we have clarified this in the methods section (statistical techniques) and in the footnotes for Figure 3.

Rather than using the phrase "compared with background levels", please instead say "when compared with expected general population risk".

This has been addressed in our revised draft.

Please write out post-COVID syndrome rather than use the acronym PCS.

This has been addressed in our revised draft.

Please clarify how readmission risk was assessed when compared to the general population. Since these patients were not hospitalized, we presume that they were simply "admitted" (as opposed to being readmitted).

Thank you for this observation, which is correct. In the revised draft, we have clarified this fact in the abstract, in the methods section when defining the outcome measures, and in the footnotes to any tables presenting results relating to readmission. Aside from these clarifications, we have generally retained the succinct terminology "readmission" throughout the paper to avoid cumbersome prose.

Rather than using the phrases "adverse events", we would prefer you consistently use the phrase or "post-discharge multi-organ dysfunction".

This has been addressed in our revised draft.