Reviewer: 1
This paper has been revised to take account of many of the statistical issues previously raised.

Follow-up has been extended and thus there are fewer patients with no set outcome data which reduces bias with respect to the critical illness logistic regression analysis. More details are provided about those who test negative. The analyses now include study week as a covariate and adjust for hospital. A new survival analysis (mortality) is also presented, and sensitivity analyses have also been carried out. The analysis involving critical illness in relation to all patients has been omitted, and the decision tree classification analysis has been removed from the paper.

There are just a few points that remain:

1. I agree that the very heterogenous group of tested subjects does not necessarily result in bias when comparing hospitalised vs non-hospitalised Covid-19 patients, although I am still concerned the decision to admit may be influenced by the 'reason/site for testing' which would, in turn, affect the hospitalised vs non-hospitalised comparison. It is useful to have the site data (Appendix Table S2) which shows that 59% of positive tests were carried out in the ED, and 96% of those hospitalised were tested in the ED. But it would be useful to know whether (in a sensitivity-type analysis) restricting the study sample to those tested in the ED gave similar results.

> **Response:** We have added this analysis as **Appendix Table S6**; of note, only 505 patients were tested in the ED and not hospitalized, so the comparison sample size is substantially smaller. Nonetheless, results are qualitatively similar to Table 1 (and especially similar to Appendix S5, also restricted to ED patients) with the main differences being attenuated effect of age, as expected since older patients are more likely to become sicker and present to the ED for first evaluation.
>
> **New text, page 11:** A sensitivity analysis adding patients seen in the ED for suspected COVID-19 but not tested produced similar results (**Appendix Table S5**), as did a different sensitivity analysis restricting non-hospitalized patients to those tested in the ED (**Appendix Table S6**).
>
> **New table:** Appendix Table S6

2. Tables 2-4. Because some of the significant multivariable-adjusted factors need to be interpreted carefully (eg a 'change' from a higher to a lower risk for hyperlipidema with statistical adjustment in Table 2) it would seem useful to add 'univariate' OR/HR results alongside the multivariable results here. This would highlight the sometimes complex relationship between a risk factor and hospitalisation/critical illness etc. Alternatively, specific factor relationships should be discussed more extensively in the text and the univariate results included in Supplementary Tables.

> **Response:** We have added a column of univariate ORs to Tables 1 and 3 as requested. In several cases, unadjusted ORs are much larger than adjusted, as would be expected given

both co-existing occurrence of diseases and increasing prevalence of disease with advancing age.

3. Out of interest, in the multivariable analyses, were interactions between any of the factors considered?

Response: Yes, as briefly noted in the text, we tested overall multicollinearity using the determinant of correlation matrix with no significant results. In fact, out of six different tests for overall multicollinearity, only one (Farrar Chi-square) was significant. We also tested a handful of individual specific interaction terms that seemed most likely to be problematic. For reference, overall collinearity results are as follows:

Admitted Model

Test	MC Results	Detection
Determinant |X'X|:	0.03	0
Farrar Chi-Square:	18589.60	1
Red Indicator:	0.12	0
Sum of Lambda Inverse:	33.54	0
Theil's Method:	-2.75	0
Condition Number:	8.79	0

Critical Illness Model

Test	MC Results	Detection
Determinant |X'X|:	0.02	0
Farrar Chi-Square:	21856.06	1
Red Indicator:	0.12	0
Sum of Lambda Inverse:	38.05	0
Theil's Method:	-16.13	0
Condition Number:	11.53	0

Existing text, page 8: We also tested for overall multicolllinearity among all variables simultaneously using the determinant of correlation matrix implemented in R's mctest library and found no significant results.

4. Table 1. I would suggest omitting the temperature and oxygen saturation results here as they are open to misinterpretation.

Response: We have omitted these results

5. Figures. The time axis should include the units, ie. ‘days’.

Response: We have corrected Figures 2-4