Namilumab or infliximab compared with standard of care in hospitalised patients with COVID-19 (CATALYST)

CATALYST investigators; Fisher, Benjamin; Veenith, Tonny; Slade, Daniel; Gaskell, Charlotte; Rowland, Matthew; Whitehouse, Tony; Scriven, James; Parekh, Dhruv; Balasubramaniam, Madhu; Cooke, Graham; Morley, Nick; Gabriel, Zoe; Wise, Matthew; Porter, Joanna; McShane, Helen; Ho, Ling-Pei; Newsome, Philip; Rowe, Anna; Sharpe, Rowena

DOI:
10.1016/S2213-2600(21)00460-4
10.1016/s2213-2600(21)00460-4

Document Version
Peer reviewed version

Citation for published version (Harvard):
CATALYST investigators, Fisher, B, Veenith, T, Slade, D, Gaskell, C, Rowland, M, Whitehouse, T, Scriven, J, Parekh, D, Balasubramaniam, M, Cooke, G, Morley, N, Gabriel, Z, Wise, M, Porter, J, McShane, H, Ho, L-P, Newsome, P, Rowe, A, Sharpe, R, Thickett, D, Bion, J, Gates, S, Richards, D, Kearns, P & Mussai, F 2022, 'Namilumab or infliximab compared with standard of care in hospitalised patients with COVID-19 (CATALYST): a randomised, multicentre, multi-arm, multistage, open-label, adaptive, phase 2, proof-of-concept trial', *The Lancet Respiratory Medicine*, vol. 10, no. 3, pp. 255-266. https://doi.org/10.1016/S2213-2600(21)00460-4, https://doi.org/10.1016/s2213-2600(21)00460-4

Namilumab or infliximab compared to standard of care in hospitalised patients with COVID-19 (CATALYST): a phase 2 randomised multicentre open adaptive multi-arm multi-stage trial

Benjamin A. Fisher, MD(Res)[1.2.3]*, Prof Tonny Veenith, PhD[4,5]*, Daniel Slade, MSc[3]*, Charlotte Gaskell, MSc[3], Matthew Rowland, DPhil[6], Prof Tony Whitehouse, MD[4,5], James Scriven, PhD[7,8], Dhruv Parekh, PhD[4,5,9] Madhu S. Balasubramaniam, MBBS[10], Prof Graham Cooke, DPhil[11], Nick Morley, MBBS[12], Zoe Gabriel, MBBS[13], Matthew P. Wise, DPhil[14], Prof Joanna Porter, PhD[15], Prof Helen McShane, PhD[16], Prof Ling-Pei Ho, DPhil[17,18], Prof Philip N. Newsome, PhD[2,19], Anna Rowe, PhD[2.3], Rowena Sharpe, PhD[3], Prof David R. Thickett, DM[5,9], Prof Julian Bion, MD[4,5], Prof Simon Gates, PhD[3], Prof Duncan Richards, DM[20] and Prof Pamela Kearns, PhD[2,3] on behalf of CATALYST investigators[&]

* These authors have made an equal contribution.

[&] Additional investigators of the CATALYST Trial are listed in Supplemental Appendix.

Corresponding authors contact details:

Dr Benjamin A. Fisher, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, B15 2TT, UK. Email: b.fisher@bham.ac.uk

Prof. Tonny Veenith, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, B15 2TT, UK. Email: t.v.veenith@bham.ac.uk

1. Rheumatology Research Group, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, UK

2. National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK.

3. Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

4. Department of Critical Care Medicine, University Hospitals Birmingham NHS Trust, United Kingdom

5. Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, UK

6. Kadoorie Centre for Critical Care Research, Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK

7. Department of Infectious Diseases, University Hospitals Birmingham NHS Trust, UK

8. Institute of Microbiology and Infection, University of Birmingham, UK.

9. Department of Respiratory Medicine, University Hospitals Birmingham NHS Trust, UK

10. Department of Critical Care Medicine, Royal Bolton Hospital, Bolton, UK

11. Department of Infectious Disease, Imperial College London, London, UK

12. Department of Haematology, Royal Hallamshire Hospital, Sheffield, UK

13. Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

14. Department of Critical Care Medicine, University Hospital of Wales, Cardiff, UK

15. Department of Respiratory Medicine, University College Hospital, London, UK

16. The Jenner Institute, University of Oxford, Oxford, UK

17. MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

18. Oxford Interstitial Lung Disease Service, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

19. Centre for Liver and Gastrointestinal Research, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK

20. Oxford Clinical Trials Research Unit, Botnar Research Centre, University of Oxford, Oxford, UK

Summary

Background

Dysregulated inflammation is associated with poor outcomes in Coronavirus disease 2019 (COVID-19). We assessed the efficacy of namilumab, a granulocyte-macrophage colony-stimulating factor inhibitor and infliximab, a tumour necrosis factor inhibitor in hospitalised patients with COVID-19 in order to prioritise agents for phase 3 trials.

Methods

In this randomised, multi-arm, parallel group, open label, adaptive phase 2 proof-of-concept trial (CATALYST) we recruited hospitalised patients ≥16 years with COVID-19 pneumonia and C-reactive protein (CRP) ≥40mg/L in nine UK hospitals. Participants were randomly allocated with equal probability to usual care, or usual care plus a single 150mg intravenous dose of namilumab (150mg) or infliximab (5mg/kg). Randomisation was stratified for ward versus ICU. The primary endpoint was improvement in inflammation in intervention arms compared to control as measured by CRP over time, analysed using Bayesian multi-level models. ISRCTN registry number 40580903.

Findings

Between 15th June 2020 and 18th February 2021 we randomised 146 participants: 54 to usual care, 57 to namilumab and 35 to infliximab. The probabilities that namilumab and infliximab were superior to usual care in reducing CRP over time were 97% and 15% with point estimates for treatment-time interactions of -0.09 (-0.19, 0.00) and 0.06 (-0.05, 0.17) respectively. Consistent effects were seen in ward and ICU patients and aligned with clinical outcomes, such that the probability of discharge (WHO levels 1-3) at day 28 was 47% and 64% for ICU and ward patients on usual care, versus 66% and 77% for patients treated with namilumab. Death occurred in 6 (11%) and 10 (19%) namilumab and usual care patients respectively, and 4 (14%) and 5 (15%) infliximab and usual care patients respectively. 134 adverse events occurred in 30/55 (55%) namilumab patients compared to 145 in 29/54 (54%) usual care patients. 102 events occurred in 20/29 (69%) infliximab patients versus 112 events in 17/34 (50%) usual care patients.

Interpretation

Namilumab, but not infliximab, demonstrated proof-of-concept evidence for reduction in inflammation in hospitalised patients with COVID-19 pneumonia which was consistent with secondary clinical outcomes. Namilumab should be prioritised for further investigation in COVID-19.

Funding

81 Medical Research Council.

Introduction

Severe Coronavirus disease 2019 (COVID-19) is associated with high mortality and disability in survivors. An excessive and dysregulated immune response contributes to these poor outcomes, as evidenced by the ability of corticosteroids and IL-6 receptor blockade to reduce mortality in hospitalised patients requiring oxygen.[1,2]

Inflammatory monocytes/macrophages (IMM) appear central to this dysregulated response,[3] resulting in disruption of pulmonary endothelial barrier integrity, microvascular thrombosis,[4] and lung tissue damage.[5] A genome-wide association study has identified the monocyte/macrophage chemotactic protein CCR2 as being associated with severe COVID-19.[6] Transcriptomic analysis of blood, lung and bronchoalveolar fluid has revealed a predominance of activated IMM within the lung, alongside expression of pro-coagulant genes within alveolar macrophages.[7] Notably, the aberrant expression of proliferation markers in blood monocytes correlates with severe disease,[8] and likely reflects a pathological early release of monocytes from the bone marrow.[9] IMM may be further activated and polarised to an inflammatory phenotype in severe disease by interaction with immune complexes containing hypoglycosylated anti-spike protein antibodies.[10]

IMM or their activity may be targeted therapeutically in a number of different ways. Given that trials with clinical outcomes require large numbers of patients to show effects, we designed a multi-arm proof of concept trial with a biomarker primary outcome to expedite decision-making on potentially effective therapeutic options for COVID-19. The aim was to provide early biological signals of efficacy to efficiently prioritise agents with the highest likelihood of success for study in established phase 3 platform trials.[11] The first two agents studied were namilumab and infliximab.

Namilumab is an anti-granulocyte-macrophage colony stimulating factor (GM-CSF) monoclonal antibody with a good safety profile up to phase 2 that has been studied in inflammatory conditions such as rheumatoid arthritis. GM-CSF is a multifunctional cytokine that is a growth factor for granulocytes and monocytes and has an important role in immune responses. In particular, it drives the activation, maturation, survival and trafficking of monocyte-derived macrophages, and their polarisation towards a more inflammatory phenotype. Elevated GM-CSF levels are closely associated with disease severity in COVID-19,[12] with GM-CSF-expressing T cells being clonally expanded in the lungs.[13] Notably, GM-CSF may also enhance the pro-coagulant activities of macrophages,[14] and blood clots are a recognised side effect of recombinant GM-CSF (sargramostim), suggesting that dysregulated GM-CSF expression may predispose to the microvascular thrombosis characteristic of COVID-19.[4].

Infliximab is a widely used anti-tumour necrosis factor (TNF) monoclonal antibody. TNF is an important pro-inflammatory cytokine and its inhibition has shown efficacy in many chronic immune mediated inflammatory diseases (IMIDs). TNF inhibition reduces mortality and severity in several mouse models of viral respiratory infection.[15,16] An IMM subset associated with severe COVID-19 shares transcriptional similarities to macrophages stimulated with both TNF and interferon gamma (IFNγ).[17] Some data suggest that IMID patients who contract COVID-19 whilst treated with TNF inhibitors have better outcomes.[18]

We sought to provide early proof-of-concept signal in a randomised trial to efficiently prioritise these approaches for subsequent testing in larger trials powered for clinical outcomes. Data on harms were also collected as a secondary objective.

Methods

Study design

The CATALYST trial is a randomised, open label, phase 2, multi-arm proof-of-concept trial.[11] A placebo control was not included due to the operational difficulties imposed by the pandemic and the proposed multi-arm design and following advice from patient and public involvement. Participants were recruited from nine hospital sites in the UK (Queen Elizabeth Hospital Birmingham; Heartlands Hospital, Birmingham; John Radcliffe Hospital, Oxford; Royal Bolton Hospital, Bolton; Imperial St Mary's Hospital, London; Royal Hallamshire Hospital, Sheffield; University Hospital of Wales, Cardiff; Good Hope Hospital, Birmingham and University College Hospital, London). The trial was approved by the East Midlands-Nottingham 2 Research Ethics Committee (20/EM/0115) and given national Urgent Public Health Status.

Participants

Eligible patients were 16 years or over, with a clinical picture strongly suggestive of SARS-CoV-2 pneumonia (confirmed by chest X-ray or CT scan, with or without a positive reverse transcription-polymerase chain reaction (RT-PCR) assay), and with a C-Reactive Protein (CRP) ≥40 mg/L. The requirement for raised CRP replaced an inclusion criterion for low oxygenation status (oxygen saturation ≤94% while breathing ambient air or a ratio of the partial pressure of oxygen to the fraction of inspired oxygen ≤300 mmHg) early in the course of recruitment following a change in primary outcome (see below; all protocol changes are summarised in supplementary Table 1). Exclusion criteria are detailed in Supplementary Information.

Written informed consent was obtained from all patients with capacity. If the patient lacked capacity, from severity of illness for example, informed consent was obtained from the patient's personal legal representative or, if unavailable, a professional legal representative according to the requirements of the UK Health Research Authority. Patients with representative consent were re-consented as soon as possible after regaining capacity.

Randomisation

Randomisation was performed by an automated minimisation procedure that attempted to allocate participants in a balanced manner between treatment arms available at the site allowing for the sole stratification variable (ward or ICU) and with a 20% random component (further details in Supplementary Appendix). At one site (Bolton) infliximab was unavailable as an intervention. Although clinical staff were aware of treatment allocation, aggregate outcomes were not provided to them, the trial management committee or the trial steering committee.

Procedures

Participants assigned to namilumab received a single intravenous (IV) dose of 150mg given over 1 hour on day 1. Those receiving infliximab had a single IV dose of 5 mg/kg over 2 hours on day 1. Participants were followed for 28 days. Blood tests were taken on days 1, 3, 5, 7, 9 and 14 until truncated by discharge or death. Physiological measures were collected until day 14, discharge, or death, and included the ratio of the oxygen saturation to fractional inspired oxygen concentration (SpO_2/FiO_2; SF ratio). The World Health Organisation (WHO) Clinical Progression Improvement Scale was assessed daily for 28 days on a 1-10 scale (online supplementary Table 2) where 1 is asymptomatic, 4 is hospitalised without oxygen, 6 is hospitalised with non-invasive ventilation or high-flow nasal oxygen, 7 is hospitalised with mechanical ventilation and 10 is death; data for level 0 (no viral load detected) was not collected.[20] If a patient was discharged earlier than day 28 this outcome was collected by means of a diary and scheduled telephone calls.

Outcomes

The primary objective of the trial was to investigate whether candidate treatments could reduce inflammation compared to usual care alone, in order to prioritise drugs to be evaluated in phase 3 trials. The primary outcome measure was CRP, collected over time until day 14. Published data indicate that CRP levels and trajectory are strongly associated with clinical outcomes including respiratory failure and death as well as with lung changes observed on CT.[11] With the objective of

having a rapid, biologically-driven efficacy signal using continuous, readily available data and a small sample size, we had initially chosen the oxygen saturation to fraction of inspired oxygen ratio (SF ratio) as the primary outcome. However, subsequent modelling of data from a large cohort of patients hospitalised in the first wave, indicated that the SF ratio might not be a viable outcome measure of sickness. This led to an early change in primary outcome to CRP, before any analysis of trial data, as previously described.[11]

Secondary outcome measures included the WHO Clinical Progression Scale as a principal clinical efficacy measure as well as hospital survival status and hospital free days, all assessed up to day 28. Hospital free days were defined as the number of days between date of hospital discharge to day 28, with patients who died or who were alive in hospital on day 28 being incorporated as 0 hospital free days. Physiological outcomes measured up to day 14 or discharge, if earlier, included the SF ratio.

Safety data were survival status and adverse events defined by the Common Terminology Criteria for Adverse Events (CTCAE), version 4.03 which fulfilled one of the following criteria: grade ≥3, secondary infection or allergic reaction. Data on harms were collected until day 28, utilising telephone follow-up if participant discharged earlier, and were submitted on case report forms by site investigators. Attribution for SAEs was made by site investigators and reviewed by arm leads or chief investigator. Given the known safety profile of infliximab, infection and allergic reaction were anticipated events. Low neutrophil count was an anticipated adverse event with namilumab.

Statistical analysis

The data were analysed according to a pre-specified Statistical Analysis Plan. Each intervention arm was compared against the control group independently, including only control patients for whom that intervention was a randomisation option i.e. usual care patients randomised after the infliximab arm closed or at the single site where infliximab was not a randomisation option were not included in the infliximab comparison. For the primary endpoint of CRP, we used Bayesian multi-level regression models[21] that allowed for nesting of the repeated measures data within patient, and non-linear responses, implemented using brms.[22] Default priors as chosen by 'brms' were utilised in all models, updated at any analysis point; these are chosen to be very weakly informative, the default covariance structure was implemented. The full details on how these are decided upon are provided in the package documentation [23].

Posterior probabilities for the treatment/time interaction covariates were used to conduct decision making at interim analyses, specifically the probability that the covariate was <0 indicating a positive

treatment effect in the direction of the intervention as per the model formulation The fitted models incorporated population-level effects for both the intercept and time, random effects for the intercept and time for patient, and fixed effects for age, location (ward/ICU), a main treatment effect, a treatment-time interaction, a treatment-location interaction and a higher order time term.

For the WHO scale, we used Bayesian longitudinal ordinal regression models, implemented using brms,[22] including in the model formulation fixed effects for location, age, a main treatment effect and a treatment-time interaction and random effects for both the intercepts and time for patient. For consistency with other trials, we also calculated the time to a two-point improvement for this outcome. Kaplan-Meier curves were produced for time to improvement and the Greenwood method was utilised in calculating confidence intervals. Results for other outcome measures were not modelled; the results are summarised graphically or tabulated. The full outline for the statistical analysis of all secondary endpoints for the study is provided in the statistical analysis plan in the supplementary appendix.

We present for the aforementioned models conditional probability plots, which show the mean predicted values of the natural logarithm of CRP, and, for the WHO scale, the predicted probability of being in each of the WHO outcome categories, conditioned on model parameter values. This enables an easy to interpret visualisation of effect of treatment on these outcomes through time.

Where relevant we include estimates of uncertainty for any point estimates at the stated confidence/probability level typically 95%.

Interim analyses were planned every 20 participants per arm up to 60 participants, and CRP data was considered by the data monitoring committee (DMC) in the context of the emerging safety data to make a recommendation as outlined in the supplementary appendix. No form of bias adjustment was applied.

Success was defined as a 90% probability of an intervention arm being better than usual care in reducing CRP as per the posterior probability for the treatment-time interaction covariate outlined above, whereas less than 50% probability defined futility. The operating characteristics, based on a simpler analysis of area under the curve for sequential CRP data, have been previously published,[11] and are presented in the supplementary appendix. These indicated a mean total sample size of between 43 and 70 patients per comparison would be required dependent on the assumptions

Pre-planned subgroup analyses were conducted to ascertain the effect of treatment on the primary outcome measure in participants recruited from ward and ICU, and with non-severe and severe

disease at baseline, with severe defined as requiring non-invasive or invasive ventilation. The effect of age was also studied.[12] Post-hoc analyses were conducted to exclude participants without a positive SARS-CoV2 PCR, and to assess the impact of baseline remdesivir use, smoking status and frailty.

The primary outcome was analysed on a modified intention to treat population, which included all participants who received trial treatment and had a baseline and at least one post treatment CRP measurement.

The modified intention to treat population for secondary outcomes included all patients who received any trial treatment and with available data for the respective outcome. The safety population included all patients in the usual care arm and all patients who had received a trial intervention in the active arms. Data on all reported harms, as well as for those meeting the pre-specified criteria, were summarised descriptively.

An independent data monitoring committee (DMC) reviewed unblinded data at interim analyses to advise the Trial Steering Committee on whether the trial data (and results from other relevant research), justified the continuing recruitment of further patients. The DMC operated in accordance with a trial-specific charter based on the template created by the Damocles Group. Statistical analysis was conducted in Stata 16 and R Version 4.0.3. The ISRCTN registry number is 40580903.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between 15th June 2020 and 18th February 2021 we assessed 299 patients for eligibility and randomised 146 participants to usual care (n=54), namilumab (n=57) and infliximab (n=35) (Figure 1). Data from the COVID-19 genomics UK consortium (COG-UK) shows that the main circulating strains in the UK within this time period were the original B lineage, the B.1.177 lineage and the B.1.1.7 lineage (alpha variant). Following a DMC review on 21st January 2021 that made recommendations on both arms based on primary outcome analysis, the Trial Steering Committee advised stopping the infliximab arm for futility (probability of benefit 21%) but to continue to recruit to usual care and namilumab, which met criteria for success (probability of benefit 99%), in order to collect further secondary outcome clinical data. A subsequent DMC meeting on 23rd February 2021, advised closing the remaining arms as the trial was close to maximal recruitment for these arms and recent changes to

standard of care with routine use of tocilizumab would affect conduct of the trial. In total 9 patients withdrew post-randomisation but before treatment and were not included in the analysis: 5 participants at their own or a relative's request (1 namilumab and 4 infliximab), 1 patient in the usual care group at the request of the treating physician, 1 patient in the infliximab group was reassessed as not having COVID-19, 1 patient in the namilumab group due to initial non-disclosure of information that met an exclusion criterion, and 1 patient in the infliximab group who was withdrawn before treatment when the DMC recommendation to stop the arm was made known.

Table 1 shows the baseline characteristics for participants. Denominators are provided below to indicate available data. Overall, groups were evenly matched although fewer patients in the infliximab group had remdesivir at enrolment. Most participants had a positive PCR assay for SARS-CoV2. Overall, 53/54 (98%) patients in the usual care group, 55/57 (97%) of the namilumab group and 33/35 (94%) of the infliximab group received oxygen at baseline. For the usual care and namilumab comparison, 16/54 (30%) and 21/57 (37%) received high-flow nasal oxygen or CPAP, and 11/54 (20%) and 11/57 (19%) were intubated and mechanically ventilated. Almost all patients received dexamethasone as part of usual care at enrolment, and around half received remdesivir. Subsequent to enrolment, all patients bar one in the namilumab group received dexamethasone, 36/53 (68%) and 37/55 (67%) patients in the usual care and namilumab arms received remdesivir, and 3/53 (6%) and 5/55 (9%) in the usual care and namilumab comparison received tocilizumab respectively. For the infliximab comparison, all patients received dexamethasone, 26/33 (79%) and 16/30 (53%) received remdesivir before or following randomisation, and 2/33 (6%) and /30 (3%) in usual care and infliximab comparison received tocilizumab respectively).

The following patients were evaluable for the primary outcome: 45 and 52 for the usual care alone versus namilumab comparison respectively, and 29 and 28 for the usual care versus infliximab comparison respectively. At the whole population level, and consistent with our previous findings and published data, CRP over time was related to the outcomes of discharge, death and continued hospitalisation at day 28 (supplementary Figure 1). Analysis of the primary outcome showed a 97% probability that namilumab plus usual care was superior to usual care alone in reducing CRP over time with a point estimate for the treatment-time interaction of -0.09 (-0.19, 0.00). (Figure 2). Model fitted values were in good agreement with raw data. This effect was consistent in ward and ICU groups based on location at randomisation as visualised in the conditional effects plots (Figure 2), and also in 'severe' and 'non-severe' patients at baseline (Supplementary Figure 2), where severe was defined as use of non-invasive or invasive ventilation. The effect of namilumab on CRP was independent of age (Supplementary Figure 3). The probability of infliximab being superior to usual care alone was 15%

with a point estimate for the treatment-time interaction of 0.06 (-0.05, 0.17). This lack of effect was consistent across ward and ICU groups (Figure 2) and severe and non-severe disease (supplementary Figure 2). Post-hoc sensitivity analyses were conducted to assess the impact of baseline remdesivir use, smoking status and frailty, and the inference for both drugs remained unchanged (data not shown). Likewise, excluding patients without a positive SARS-CoV2 PCR did not change the inference (data not shown). Effects of namilumab and infliximab on CRP were also consistent with an area-under-the curve analysis (data not shown). Supplementary Tables 3 and 4 show point estimates for ln(CRP) predicted values with associated credible intervals at baseline, day 7 and day 14 for both ward and ICU patients.

Amongst secondary endpoints, the principal efficacy outcome was the 1-10 point WHO clinical progression scale. For the modified intention-to-treat comparisons between usual care and namilumab, data were available for 53 and 55 patients respectively. Figure 3 shows the proportion of patients at each WHO scale level over 28 days as well as the conditional modelled probabilities of being at each level over time for ward and ICU. In the namilumab arm for patients recruited from both ward and ICU, the probability of having lower scores is consistently increased over time in comparison with usual care. For example, the arms were similar at baseline but by day 28, the probability of discharge (WHO levels 1-3 combined) was 47% and 64% for ICU and ward patients on usual care, versus 66% and 77% for patients treated with namilumab (supplementary Table 5). At day 14, the probability of an ICU patient still needing non-invasive ventilation, invasive ventilation or to have died (WHO ≥6) was 54% in the usual care arm vs. 36% in the namilumab arm. Time to two point improvement was also seen to be shorter in the namilumab arm (Table 2 and supplementary Figure 4). Comparable improvements on WHO scale were not observed with infliximab (Supplementary Figure 5 and supplementary Table 6). The median hospital free days for usual care and namilumab were 17 (IQR 0, 23) and 20 (IQR 3, 23) respectively, and for usual care and infliximab, 17 (0, 23) and 17 (3, 23). Data were also collected on respiratory rate, body temperature and destination of discharge, however results were non-informative and data is not shown. Similarly, data was collected on lymphocyte and neutrophil counts, neutrophil: lymphocyte ratios, and ferritin, d-dimers and lactate dehydrogenase (LDH). These outcomes will be presented alongside exploratory biological outcomes in a future publication.

By day 28, there were fewer deaths and more discharges in the namilumab group with 43 (78%) participants discharged, 6 (11%) still in hospital and 6 (11%) dead, compared to 33 (62%), 11 (20%), and 10 (19%) for usual care alone (Table 3). Interestingly, despite the challenges we described in

modelling the SF ratio, trends to improvement in oxygenation status were observed with namilumab (supplementary Figure 6).

For the namilumab and usual care comparison, a total of 279 adverse events were reported in 59 of the 109 patients in the safety population (54%; 134 events in n=30 and 145 events in n=29 for namilumab and usual care respectively). Of these, 131 (90%) and 103 events (77%) events were grade 3 or above for usual care and namilumab respectively. Infections were more common in the namilumab group (20 events) compared with usual care (10 events). Supplementary Table 7 shows adverse events that were grade ≥3, secondary infection or allergic reaction, for which more than one event occurred. There were 10 serious adverse events in each of the usual care and namilumab groups respectively. All except one of the namilumab SAEs were considered unrelated, the related case being a re-admission with bacterial pneumonia 26 days after receiving namilumab and on a background of a prolonged admission for social reasons and known COPD.

For the infliximab and usual care comparison, a total of 214 adverse events were reported in 37 of the 63 patients in the safety population (60%; 112 events in 17 usual care patients and 102 events in 20 infliximab patients). Of these, 101 (90%) and 78 (77%) were grade 3 or above for usual care and infliximab respectively. There were 7 infection events in usual care and 4 with infliximab. There were 5 serious adverse events in the usual care group and 6 with infliximab, all considered unrelated. There were no deaths in the safety population outside of the mITT population.

Discussion

Our trial clearly demonstrated that the addition of namilumab, but not infliximab to usual care, reduced inflammation as measured by CRP in hospitalised patients with COVID-19, when compared to usual care alone. Importantly, the secondary clinical outcomes are consistent and shared the same directionality as the primary outcome for both interventions, despite not being formally powered to assess for such differences. Our proof-of-concept findings with GM-CSF inhibition is consistent with our hypothesis that recruitment and activation of IMM are important in the pathogenesis of severe COVID-19. This is also consistent with published findings from small non-randomised trials,[24,25] and recent, large randomised trials of other GM-CSF inhibitors in COVID-19. Otilimab showed benefit for the primary endpoint of being alive and free of respiratory failure at day 28 in a predefined subgroup of patients aged 70 or over.[26] Lenzilumab, given as a three dose course in non-ventilated hospitalised patients, showed benefit over standard care in the primary outcome of survival without ventilation,

an effect that seemed more pronounced in patients aged 85 or under and with CRP <150 mg/L.[27] Our data suggest the effect of a single dose of namilumab on CRP and WHO score is independent of age, although this requires confirmation in larger studies. Although it is not possible to directly compare these studies given the differences in sample sizes, inclusion criteria and study designs, the overall RCT data suggest benefit of GM-CSF inhibition in COVID-19. For, example, we observed mortality in the namilumab group of 11% compared to 19% with usual care. In the lenzilumab and otilimab phase 3 trials this was 10% in the active arm compared to 14% (day 28), and 23% versus 24% (day 60) respectively. In two recent phase 2 mavrilimumab trials, mortality was 8% versus 21% [28] , and 5% compared to 16% [29]. Benefit has also been observed with IL-6 inhibition with a recent meta-analysis showing a day 28 mortality of 22% in the active arms compared to 25% with usual care/placebo [30].

In the absence of large treatment effects, small trials using traditional clinical outcomes may give inconclusive or contrary findings in COVID-19, as exemplified by earlier studies of tocilizumab. The CATALYST trial was designed to use a repeatedly collected continuous measure of CRP with a Bayesian adaptive approach that we predicted would require a smaller sample size to show evidence of efficacy or futility. CRP levels, including the rate of decline, have been associated with clinical outcome in COVID-19 (reviewed in[11]) and we hypothesised that an immunomodulatory agent unable to alter CRP would be a less promising candidate to take forward into phase 3 trials. In the face of many options for repurposing immunomodulatory therapies in COVID-19, we contend that such a prioritisation approach will make the most efficient use of phase 3 resource and accelerate development of effective drugs.

In contrast to the observed effect of namilumab, we could not demonstrate a comparable benefit on CRP with infliximab and the arm was stopped for futility. TNF is in important pro-inflammatory cytokine produced by macrophages as well as other cell types, with context-dependent pleiotropic effects including further activation of IMM and up-regulation of inflammatory mediators such as IL-6. One previous non-randomised study of infliximab suggested potential efficacy, albeit with significant limitations including small sample size, use of historical controls, and being conducted prior to routine use of corticosteroids.[31] This, together with circumstantial data, justified our inclusion of infliximab.[18] However, although TNF inhibitors are widely used in inflammatory diseases, not all IMID are responsive, and TNF itself may suppress certain pro-inflammatory factors that may be relevant to COVID-19 such as type 1 interferon expression and Th17 cell differentiation.[32] Inhibition of such cross-regulatory effects may underlie our negative findings, or simply indicate that TNF is not on a critical path to driving inflammatory responses as measured by CRP in patients hospitalised with COVID-19. GM-CSF inhibition might also have an additional benefit in retarding neutrophil recruitment and activation that may be of importance in the pathogenesis of severe COVID-19 and acute respiratory

distress syndrome.[33] Our safety data suggest that the lack of response to infliximab is not due to an increase in secondary infections. We cannot exclude the possibility of benefit with infliximab being seen in a subset of patients, in larger studies, or with a dose higher than the standard dose we employed although this was in large molar excess relative to published concentrations of circulating TNF in COVID-19. It should also be noted that remdesivir use was lower in the infliximab arm when compared to usual care, although the recent negative SOLIDARITY trial for remdesivir suggest this might not unduly influence our results [34], and results of our post-hoc sensitivity analyses were consistent. However, the clear divergence in primary outcome is broadly reflected in the secondary clinical findings and justifies the prioritisation of GM-CSF inhibition over TNF inhibition at this dose for further study in hospitalised COVID-19 patients.

GM-CSF has an important role in the differentiation of alveolar macrophages, and consequently in surfactant clearance, as well as being an important survival factor for lung epithelial cells. Absence of GM-CSF signalling, through genetic defect in the receptor or very high levels of polyclonal autoantibodies to GM-CSF, have been associated with pulmonary alveolar proteinosis (PAP). PAP has been an adverse event of special interest in previous clinical trials of GM-CSF inhibitors but, to our knowledge, has never been observed. It is important to note, (i) that therapeutic monoclonal antibodies will not completely inhibit GM-CSF signalling which appears to be a requirement for PAP,[35] but rather will down-regulate excessive pathway activation, (ii) lack of GM-CSF does not prevent macrophage uptake of surfactant as much as its catabolism, therefore the effect of short-term inhibition is likely to be less pronounced on surfactant clearance when compared with long-term inhibition, (iii) down regulation of monocyte activation, which is the aim of GM-CSF inhibition, should itself lead to a reduction in alveolar epithelial cell damage in COVID-19. However it is also important to note an opposing view that administration of GM-CSF might have therapeutic benefits and the results of clinical trials of inhaled and intravenous sargramostim are awaited.[36]

Our study has a number of limitations. Similar to many other trials in COVID-19 we did not use a placebo control. However, the discordant results of the two active arms, when compared to usual care, as well as the objective nature of CRP data, suggest this does not explain the positive findings we observed with namilumab. Our sample size is too small for a definitive assessment of clinical outcomes and further studies are required for this as well as to understand better the population that may benefit most. Our results may not generalise to hospitalised patients without evidence of pneumonia or raised CRP or patients not requiring hospitalisation. Harms data are difficult to interpret given the small number of participants, lack of blinding, the severity of the background illness and that data was being collected during a pandemic. Overall the number of total adverse events did not differ between namilumab and usual care. However, our data do emphasise the need to monitor secondary

infections in future COVID-19 trials, particularly given the use of combination immune-modulating treatments.

Despite the advances of dexamethasone and tocilizumab in COVID-19, mortality amongst patients with severe disease remain high.[2] There therefore remains considerable unmet medical need, and data pointing to the role of both IMM and GM-CSF in severe COVID-19, together with our findings reported here, strongly suggest that targeted GM-CSF inhibitors such as namilumab should be further investigated in hospitalised patients with COVID-19.

Research in Context

Evidence before this study

We searched Pubmed and medRxiv on 10th May 2021, using the following search terms [(randomised OR trial) AND (anti-GM-CSF OR namilumab OR mavrilimumab OR otilimab OR lenzilumab OR gimsilumab OR TJ003234 OR anti-TNF OR infliximab OR adalimumab OR etanercept OR golimumab OR certolizumab) AND (COVID* OR SARS-CoV-2 OR SARS-CoV)]. Two small non-randomised studies with drugs targeting GM-CSF or its receptor (lenzilumab and mavrilimumab) and one study with a TNF inhibitor (infliximab) have all suggested potential efficacy but with significant limitations of small sample size, use of historical controls, and being conducted prior to routine use of corticosteroids. One RCT with mavrilimumab was small and inconclusive. Two larger RCTs with other anti-GM-CSF inhibitors have recently been published. Otilimab showed benefit for the primary endpoint of being alive and free of respiratory failure at day 28 in a predefined subgroup of patients aged 70 or over. Lenzilumab, given as a three dose course, in non-ventilated hospitalised patients showed benefit over standard care in the primary outcome of survival without ventilation, an effect that seemed more pronounced in patients aged 85 or under and with CRP <150 mg/L. We identified no published randomised trials of TNF inhibitors in COVID-19.

Added value of this study

This is the first randomised trial of namilumab and infliximab in COVID-19. We found that both drugs were safe and that namilumab, but not infliximab, showed proof of concept evidence of reduction in inflammation as measured by CRP in hospitalised patients with COVID-19 pneumonia. Secondary clinical outcomes were concordant with the primary outcome, with trends to improvement in patients recruited from both ward and ICU.

Implications of all the available evidence

Consistent with emerging evidence implicating GM-CSF and inflammatory monocytes/macrophages in the pathogenesis of severe COVID-19, namilumab improved both biological and clinical outcomes. It should be prioritised for further study in COVID-19.

Contributors

BAF, TV, MR, TW, DP, AR, RS. DRT, JB, SG, DR and PK conceived the study. BAF, TV, DS, MR, TW, DP, AR, RS, DRT, JB, HM, LH, PNN, SG, DR and PK designed the clinical trial. BAF and DP were arm leads for namilumab and MR and DR were arm leads for infliximab. TV, TW, JS, DP, MSB, GC, NM, ZG, MPW, JP and AR recruited patients and/or collected data. DS, CG and SG conducted the statistical analysis andhad access to the raw data and verified the data. BAF drafted the manuscript which all authors revised and approved for submission.

Declaration of interests

BAF has undertaken consultancy for Novartis, BMS, Servier, Galapagos and Janssen and received research funding from Servier and Galapagos; MR is currently undertaking a Senior Clinical Fellowship financed by Roche; PK has undertaken consultancy for BMS, and AstraZeneca, and has received research funding from Bayer and Pfizer; DR is a former employee of GSK; all are unrelated to this trial. All other authors declare no competing interests.

Data sharing

Participant data and the associated supporting documentation will be available within six months after the publication of this manuscript. Details of our data request process are available on the Cancer Research UK Clinical Trials Unit (CRCTU) website. Only scientifically sound proposals from appropriately qualified research groups will be considered for data sharing. The decision to release data will be made by the CRCTU Director's Committee, who will consider the scientific validity of the request, the qualifications and resources of the research group, the views of the Chief Investigator and the Trial Steering Committee, consent arrangements, the practicality of anonymising the requested data and contractual obligations. A data sharing agreement will cover the terms and conditions of the release of trial data and will include publication requirements, authorship and acknowledgements and obligations for the responsible use of data. An anonymised encrypted dataset

will be transferred directly using a secure method and in accordance with the University of Birmingham's IT guidance on encryption of datasets.

Acknowledgements

This trial is supported by the Medical Research Council (MRC) grant number MC_PC_20007. SG is supported by a Senior Investigator Award from the National Institute of Health Research. Staff at the CRCTU are supported by core funding grants from Cancer Research UK (C22436/A25354), the NIHR Biomedical Research Centre (BRC-1215-20009), The Kennedy Trust for Rheumatology Research as part of the Arthritis – Trials Acceleration Programme (KENN161704), and Innovate UK as part of the Midlands – Wales Advanced Therapy Treatment Centres (104232). This paper presents independent research supported by the NIHR Birmingham Biomedical Research Centres at Birmingham, Oxford, Imperial College London and University College London. GC is supported by a NIHR Research Professorship. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. Namilumab was provided free of charge by Izana Bioscience, Oxford, UK (now part of Roivant). Infliximab is being provided free of charge by Celltrion.

References

1. Group RC, Horby P, Lim WS, et al. Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report. *N Engl J Med* 2020.
2. Horby PW, Pessoa-Amorim G, Peto L, et al. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): preliminary results of a randomised, controlled, open-label, platform trial. *medRxiv* 2021: 2021.02.11.21249258.
3. Merad M, Martin JC. Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. *Nat Rev Immunol* 2020; **20**(6): 355-62.
4. Ackermann M, Verleden SE, Kuehnel M, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. *N Engl J Med* 2020; **383**(2): 120-8.
5. Filbin MR, Mehta A, Schneider AM, et al. Plasma proteomics reveals tissue-specific cell death and mediators of cell-cell interactions in severe COVID-19 patients. *bioRxiv* 2020.
6. Pairo-Castineira E, Clohisey S, Klaric L, et al. Genetic mechanisms of critical illness in Covid-19. *Nature* 2020.
7. Daamen AR, Bachali P, Owen KA, et al. Comprehensive Transcriptomic Analysis of COVID-19 Blood, Lung, and Airway. *bioRxiv* 2020: 2020.05.28.121889.
8. Mann ER, Menon M, Knight SB, et al. Longitudinal immune profiling reveals distinct features of COVID-19 pathogenesis. *medRxiv* 2020: 2020.06.13.20127605.
9. Schulte-Schrepping J, Reusch N, Paclik D, et al. Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. *Cell* 2020; **182**(6): 1419-40 e23.
10. Hoepel W, Chen H-J, Allahverdiyeva S, et al. Anti-SARS-CoV-2 IgG from severely ill COVID-19 patients promotes macrophage hyper-inflammatory responses. *bioRxiv* 2020: 2020.07.13.190140.

11. Veenith T, Fisher BA, Slade D, et al. CATALYST trial protocol: A multicentre, open-label, phase II, multi-arm trial for an early and accelerated evaluation of the potential treatments for COVID-19 in hospitalised adults. *medRxiv* 2021: 2021.02.10.21251478.
12. Thwaites RS, Sanchez Sevilla Uruchurtu A, Siggins MK, et al. Inflammatory profiles across the spectrum of disease reveal a distinct role for GM-CSF in severe COVID-19. *Science Immunology* 2021; **6**(57): eabg9873.
13. Zhao Y, Kilian C, Turner JE, et al. Clonal expansion and activation of tissue-resident memory-like Th17 cells expressing GM-CSF in the lungs of severe COVID-19 patients. *Sci Immunol* 2021; **6**(56).
14. Williams MA, White SA, Miller JJ, et al. Granulocyte-macrophage colony-stimulating factor induces activation and restores respiratory burst activity in monocytes from septic patients. *J Infect Dis* 1998; **177**(1): 107-15.
15. Channappanavar R, Fehr AR, Vijay R, et al. Dysregulated Type I Interferon and Inflammatory Monocyte-Macrophage Responses Cause Lethal Pneumonia in SARS-CoV-Infected Mice. *Cell Host Microbe* 2016; **19**(2): 181-93.
16. Hussell T, Pennycook A, Openshaw PJ. Inhibition of tumor necrosis factor reduces the severity of virus-specific lung immunopathology. *Eur J Immunol* 2001; **31**(9): 2566-73.
17. Zhang F, Mears JR, Shakib L, et al. IFN-γ and TNF-α drive a CXCL10+ CCL2+ macrophage phenotype expanded in severe COVID-19 and other diseases with tissue inflammation. *bioRxiv* 2020: 2020.08.05.238360.
18. Robinson PC, Liew DFL, Liew JW, et al. The Potential for Repurposing Anti-TNF as a Therapy for the Treatment of COVID-19. *Med (N Y)* 2020; **1**(1): 90-102.
19. Felton T, Pattison N, Fletcher S, Finney S, Walsh T, Dark P. Co-enrolment to UK Critical Care Studies – A 2019 update. *Journal of the Intensive Care Society*; **0**(0): 1751143720971542.
20. Characterisation WHOWGotC, Management of C-i. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis* 2020; **20**(8): e192-e7.
21. Gelman A, Hill J, Vehtari A. Regression and Other Stories. Cambridge: Cambridge University Press; 2020.
22. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *2017* 2017; **80**(1): 28.
23. https://CRAN.R-project.org/package=brms.
24. Temesgen Z, Assi M, Shweta FNU, et al. GM-CSF Neutralization With Lenzilumab in Severe COVID-19 Pneumonia: A Case-Cohort Study. *Mayo Clin Proc* 2020; **95**(11): 2382-94.
25. De Luca G, Cavalli G, Campochiaro C, et al. GM-CSF blockade with mavrilimumab in severe COVID-19 pneumonia and systemic hyperinflammation: a single-centre, prospective cohort study. *Lancet Rheumatol* 2020; **2**(8): e465-e73.
26. Patel J, Beishuizen A, Ruiz XB, et al. A Randomized Trial of Otilimab in Severe COVID-19 Pneumonia (OSCAR). *medRxiv* 2021: 2021.04.14.21255475.
27. Temesgen Z, Burger CD, Baker J, et al. LENZILUMAB EFFICACY AND SAFETY IN NEWLY HOSPITALIZED COVID-19 SUBJECTS: RESULTS FROM THE LIVE-AIR PHASE 3 RANDOMIZED DOUBLE-BLIND PLACEBO-CONTROLLED TRIAL. *medRxiv* 2021: 2021.05.01.21256470.
28. Pupim L, Wang TS, Hudock K, et al. LB0001 MAVRILIMUMAB IMPROVES OUTCOMES IN PHASE 2 TRIAL IN NON-MECHANICALLY-VENTILATED PATIENTS WITH SEVERE COVID-19 PNEUMONIA AND SYSTEMIC HYPERINFLAMMATION. *Annals of the Rheumatic Diseases* 2021; **80**(Suppl 1): 198-9.
29. Cremer PC, Abbate A, Hudock K, et al. Mavrilimumab in patients with severe COVID-19 pneumonia and systemic hyperinflammation (MASH-COVID): an investigator initiated, multicentre, double-blind, randomised, placebo-controlled trial. *Lancet Rheumatol* 2021; **3**(6): e410-e8.
30. Group WHOREAfC-TW, Shankar-Hari M, Vale CL, et al. Association Between Administration of IL-6 Antagonists and Mortality Among Patients Hospitalized for COVID-19: A Meta-analysis. *JAMA* 2021.

31. Stallmach A, Kortgen A, Gonnert F, Coldewey SM, Reuken P, Bauer M. Infliximab against severe COVID-19-induced cytokine storm syndrome with organ failure-a cautionary case series. *Crit Care* 2020; **24**(1): 444.
32. Notley CA, Inglis JJ, Alzabin S, McCann FE, McNamee KE, Williams RO. Blockade of tumor necrosis factor in collagen-induced arthritis reveals a novel immunoregulatory pathway for Th1 and Th17 cells. *J Exp Med* 2008; **205**(11): 2491-7.
33. Masso-Silva JA, Moshensky A, Lam MTY, et al. Increased peripheral blood neutrophil activation phenotypes and NETosis in critically ill COVID-19 patients: a case series and review of the literature. *Clin Infect Dis* 2021.
34. Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results. *New England Journal of Medicine* 2020; **384**(6): 497-511.
35. Piccoli L, Campo I, Fregni CS, et al. Neutralization and clearance of GM-CSF by autoantibodies in pulmonary alveolar proteinosis. *Nat Commun* 2015; **6**: 7375.
36. Mehta P, Porter JC, Manson JJ, et al. Therapeutic blockade of granulocyte macrophage colony-stimulating factor in COVID-19-associated hyperinflammation: challenges and opportunities. *Lancet Respir Med* 2020; **8**(8): 822-30.

Table 1. Baseline characteristics of all patients randomised.

		Namilumab		Infliximab	
		Usual care alone (n=54)	Active arm (n=57)	Usual care alone (n=34)	Active arm (n=35)
Male n(%)		37 (69)	34 (60)	21 (62)	19 (54)
Age, median (IQR)		62.8 (51.9, 70.5)	56.2 (47.6, 63.3)	64.5 (51.9, 71.9)	55.4 (46.1, 70.5)
Ethnicity	White, n (%)	33 (61)	34 (60)	23 (68)	17 (49)
	Black	1 (2)	1 (2)	0 0.0	2 (6)
	South Asian	7 (13)	8 (14)	3 (9)	4 (11)
	Other	11 (20)	14 (25)	7 (21)	9 (26)
	Not known	2 (4)	0 (0.0)	1 (3)	3 (9)
Clinical Frailty score, level 4-8[1], n (%)		7 (13)	4 (7)	5 (15)	5 (14)
Smoking status	Ever, n (%)	22 (41)	15 (26)	11 (32)	12 (34)
Body mass index, median (IQR)		29.5 (25.4, 34.7)	30.5 (27.1, 35.4)	30.7 (25.2, 34.3)	32.3 (26.9, 35.9)
Background respiratory disease[3], n(%)		13 (24)	13 (23)	10 (29)	8 (23)
Background diabetes, n(%)		22 (41)	17 (30)	12 (35)	11 (31)
Care status	Ward	33 (61)	33 (58)	22 (65)	22 (63)
	ICU	21 (39)	24 (42)	12 (35)	13 (37)
SARS-CoV2 PCR result n(%)	Positive	50 (93)	54 (95)	30 (88)	29 (83)
	Negative	3 (6)	2 (4)	3 (9)	6 (17)
Previous COVID-19 treatment at baseline, n (%)	Corticosteroids	49 (91)	53 (93)	29 (85)	33 (94)
	Remdesivir	29 (54)	32 (56)	21 (62)	10 (29)
	Antibiotics	46 (85)	48 (84)	28 (82)	31 (89)
Time to enrolment (days), median (IQR)		1 (1,3)	1 (1,2)	2 (1, 3)	1 (1, 2)
CRP, median (IQR)		108.0 (60.0, 160.0)	94.6 (55.4, 171.0)	88.0 (48.8, 142.0)	99.0 (46.0, 173.0)
Lymphocyte count, median (IQR)		0.8 (0.6, 1.2)	0.9 (0.6, 1.1)	0.9 (0.6, 1.3)	0.9 (0.6, 1.0)
Neutrophil count, median (IQR)		7.2 (5.4, 10.0)	7.5 (5.0, 10.1)	7.2 (5.5, 11.0)	6.8 (4.5, 9.5)
Ferritin, median (IQR), n=51, 37		750 (490, 1685)	791 (433, 1621)	676 (506, 1022)	642 (435, 1114)
D-dimers, median (IQR), n=57, 47		787 (376, 1822)	592 (227, 1418)	739 (414, 1184)	398 (235, 805)

[1]Vulnerable, mildly frail, moderately frail, severely frail. [2]Time from date of hospital admission to date of randomisation.
[3]The number of patients that have at least one of the following lung disease co-morbidities (chronic obstructive pulmonary disease, asthma, interstitial lung disease).

Table 2 – Median time in days (95% CI) to a two point improvement in the WHO clinical progression scale, for overall and subgroups for both drugs (modified intention to treat population). NR, not recordable.

	Namilumab			Infliximab		
	n	Usual care	Active arm	n	Usual care	Active arm
Whole population	108	10 (7,12)	8 (6,9)	62	10 (6, 14)	15 (6, 21)
Ward	66	9 (6,12)	8 (5,10)	42	9 (5, 12)	15 (5, NR)
ICU	42	14 (5,NR)	8 (6,11)	20	14 (4, NR)	19 (6, 28)

NR, not recordable.

Table 3. Hospital discharge status at day 28. Data was available on all patients (modified intention to treat population), n(%). Difference in proportions (95% CI).

		Namilumab			Infliximab		
	Status	Usual care (n=54)	Active arm (n=55)	Usual Care vs Namilumab	Usual care (n=34)	Active arm (n=29)	Usual Care vs Infliximab
Whole population	Discharge	33 (61)	43 (78)	-0.17 (-0.34, -0.001)	22 (65)	22 (76)	-0.11 (-0.34, 0.11)
	In hospital	11 (20)	6 (11)	0.09 (-0.04, 0.23)	7 (21)	3 (10)	0.10 (-0.07, 0.28)
	Death	10 (19)	6 (11)	0.08 (-0.06, 0.21)	5 (15)	4 (14)	0.01 (-0.16, 0.18)
Ward	Discharge	28 (85)	29 (88)	-0.03 (-0.20, 0.14)	19 (86)	16 (80)	0.06 (-0.16, 0.29)
	In hospital	4 (12)	2 (6)	0.06 (-0.08, 0.20)	2 (9)	1 (5)	0.04 (-0.11, 0.19)
	Death	1 (3)	2 (6)	-0.03 (-0.13, 0.07)	1 (5)	3 (15)	-0.10 (-0.28, 0.07)
ICU	Discharge	5 (24)	14 (64)	-0.40 (-0.67, -0.13)	3 (25)	6 (67)	-0.42 (-0.81, -0.02)
	In hospital	7 (33)	4 (18)	0.15 (-0.11, 0.41)	5 (42)	2 (22)	0.19 (-0.19, 0.58)
	Death	9 (43)	4 (18)	0.25 (-0.02, 0.51)	4 (33)	1 (11)	0.22 (-0.11, 0.56)

610

Figure Legends

Figure 1. Trial profile indicating number of subjects evaluable for the primary outcome.

Figure 2. Conditional effects plots of the natural logarithm of CRP modelled over time in days in patients recruited in ward and ICU for namilumab (A) and infliximab (B).

Figure 3. WHO clinical progression score over 28 days for usual care versus namilumab. A, stacked bar chart of raw data for whole population eligible for comparison. B, conditional effects plots of WHO score modelled over time in days showing the probability of being at each level on each day for patients recruited in ICU and ward.

Figure 1.

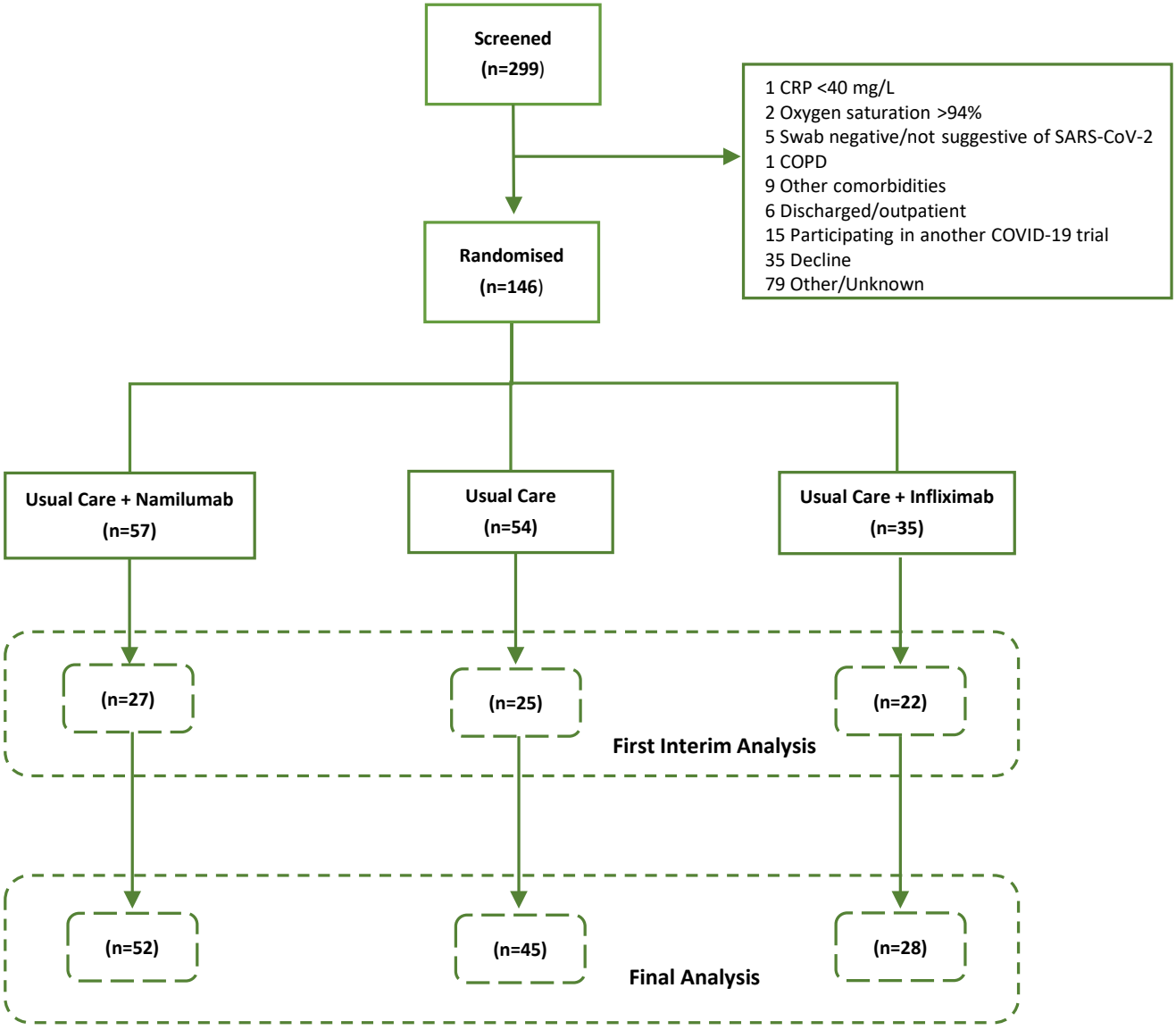
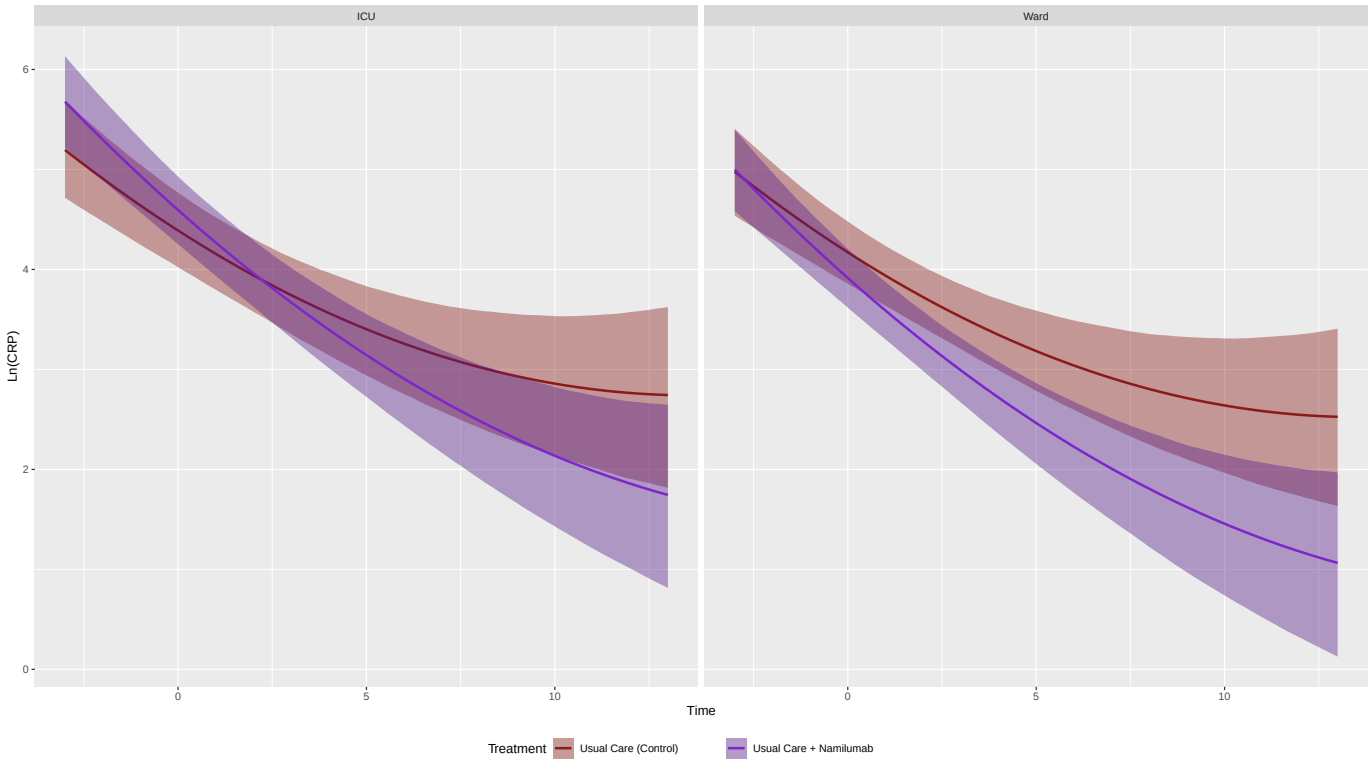
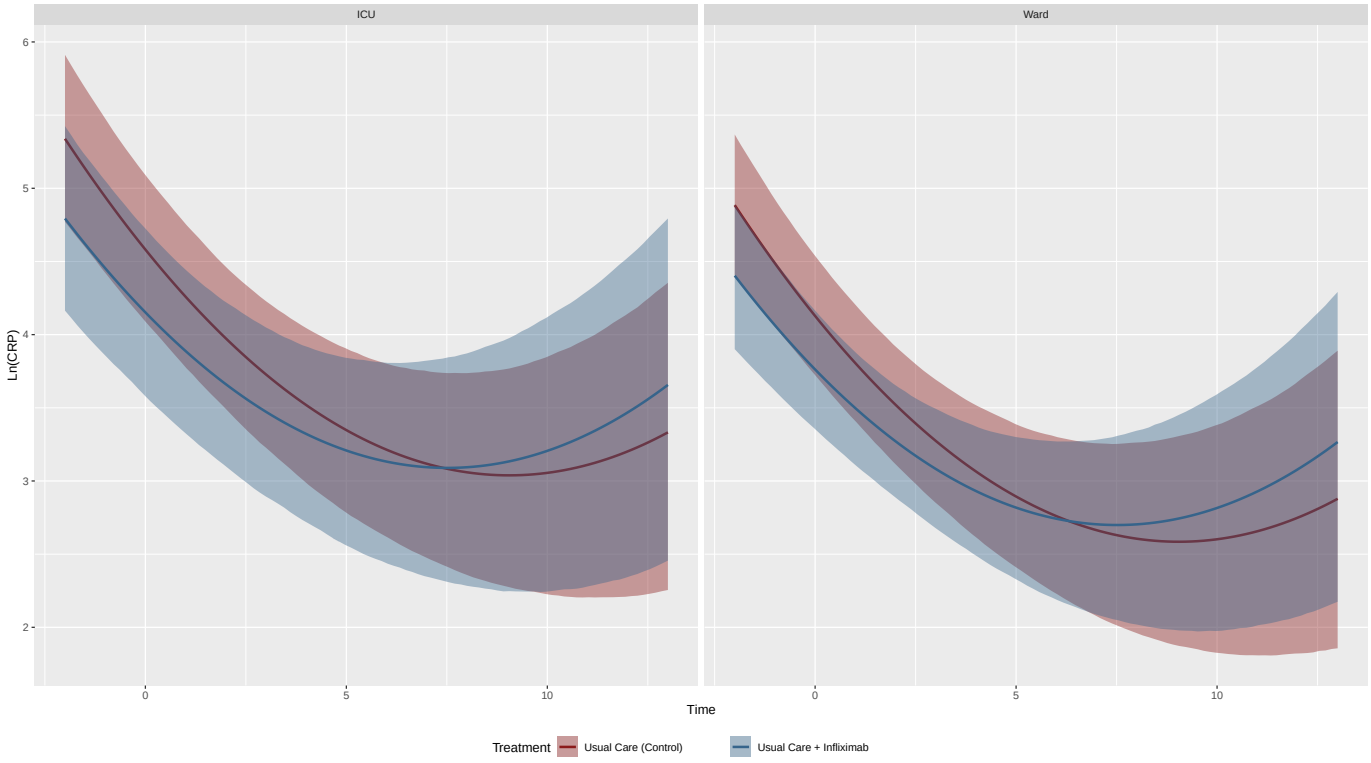

Screened
(n=299)
1 CRP <40 mg/L
2 Oxygen saturation >94%
5 Swab negative/not suggestive of SARS-CoV-2
1 COPD
9 Other comorbidities
6 Discharged/outpatient
15 Participating in another COVID-19 trial
35 Decline
79 Other/Unknown
Randomised
(n=146)
Usual Care + Namilumab
(n=57)
Usual Care
(n=54)
Usual Care + Infliximab
(n=35)
(n=27)
(n=25)
(n=22)
First Interim Analysis
(n=52)
(n=45)
(n=28)
Final Analysis

Figure 2.

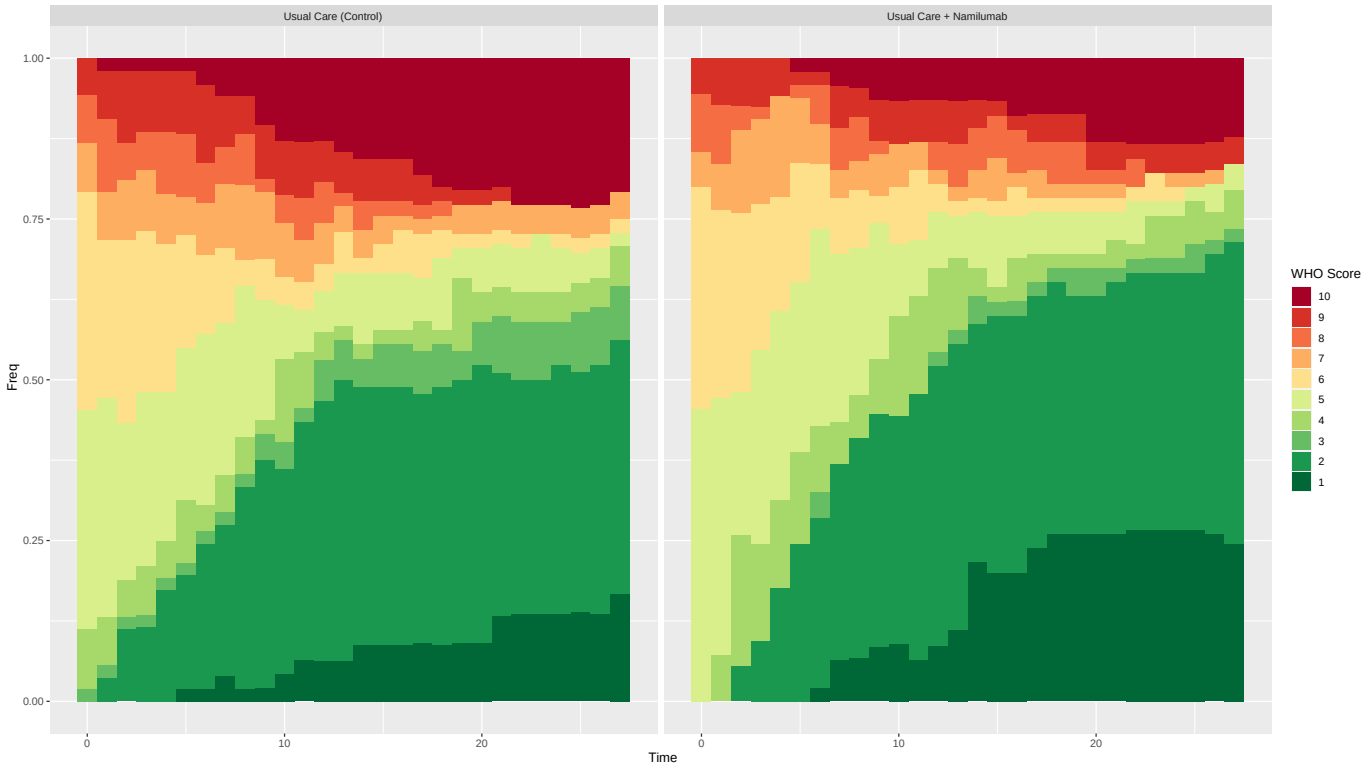

ICU
Ward
Ln(CRP)
Time
Treatment
Usual Care (Control)
Usual Care + Namilumab

(a)

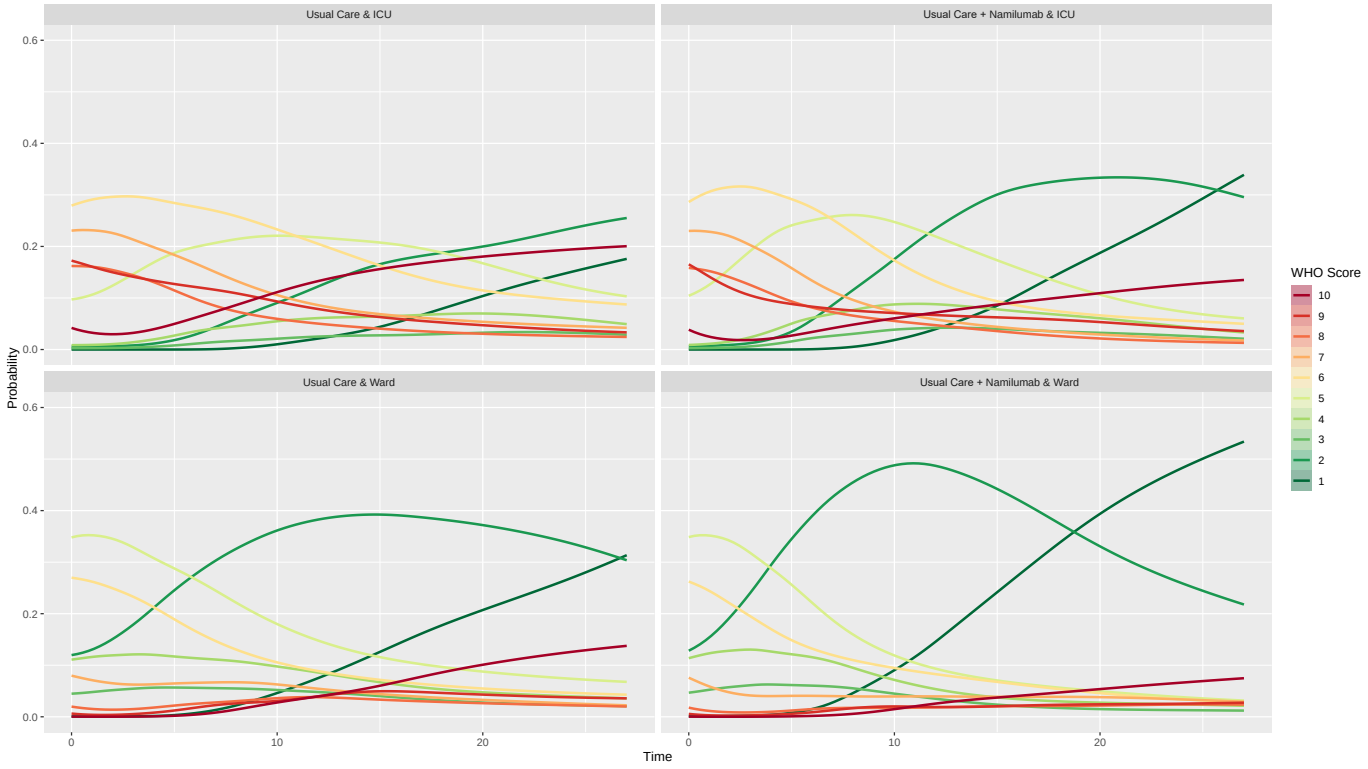

ICU
Ward
Ln(CRP)
Time
Treatment
Usual Care (Control)
Usual Care + Infliximab

(b)

Figure 3.


Usual Care (Control)
Usual Care + Namilumab
Freq
1.00
0.75
0.50
0.25
0.00
0
10
20
Time
WHO Score
10
9
8
7
6
5
4
3
2
1

(a)


Usual Care & ICU
Usual Care + Namilumab & ICU
Usual Care & Ward
Usual Care + Namilumab & Ward
Probability
0.6
0.4
0.2
0.0
0
10
20
Time
WHO Score
10
9
8
7
6
5
4
3
2
1

(b)

Supplementary Appendix

Namilumab or infliximab compared to standard of care in hospitalised patients with COVID-19 (CATALYST): a phase 2 randomised adaptive trial

Benjamin A. Fisher[1,2,3]*, Tonny Veenith[4,5]*, Daniel Slade[3]*, Charlotte Gaskell[3], Matthew Rowland[6], Tony Whitehouse[4,5], James Scriven[7,8], Dhruv Parekh[4,5,9] Madhu S. Balasubramaniam[10], Graham Cooke[11], Nick Morley[12], Zoe Gabriel[13], Matthew P. Wise[14], Joanna Porter[15], Helen McShane[16], Ling-Pei Ho[17,18], Philip N. Newsome[2,19], Anna Rowe[2,3], Rowena Sharpe[3], David R. Thickett[5,9], Julian Bion[4,5], Simon Gates[3], Duncan Richards[20] and Pamela Kearns[2,3] on behalf of CATALYST investigators

* These authors have made an equal contribution.

Corresponding authors contact details:

Dr Benjamin A. Fisher, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, B15 2TT, UK. Email: b.fisher@bham.ac.uk

Prof. Tonny Veenith, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Birmingham, B15 2TT, UK. Email: t.v.veenith@bham.ac.uk

1. Rheumatology Research Group, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, UK

2. National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK.

3. Cancer Research UK Clinical Trials Unit, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK

4. Department of Critical Care Medicine, University Hospitals Birmingham NHS Trust, United Kingdom

5. Birmingham Acute Care Research Group, Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, UK

6. Kadoorie Centre for Critical Care Research, Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, UK

7. Department of Infectious Diseases, University Hospitals Birmingham NHS Trust, UK

8. Institute of Microbiology and Infection, University of Birmingham, UK.

9. Department of Respiratory Medicine, University Hospitals Birmingham NHS Trust, UK

10. Department of Critical Care Medicine, Royal Bolton Hospital, Bolton, UK

11. Department of Infectious Disease, Imperial College London, London, UK

12. Department of Haematology, Royal Hallamshire Hospital, Sheffield, UK

13. Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

14. Department of Critical Care Medicine, University Hospital of Wales, Cardiff, UK

15. Department of Respiratory Medicine, University College Hospital, London, UK

16. The Jenner Institute, University of Oxford, Oxford, UK

17. MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK

18. Oxford Interstitial Lung Disease Service, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

19. Centre for Liver Research, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK

20. Oxford Clinical Trials Research Unit, Botnar Research Centre, University of Oxford, Oxford, UK

The CATALYST investigators include the following:

Prof. Tonny Veenith (Chief Investigator, University Hospitals Birmingham), Dr Benjamin A. Fisher (Deputy Chief Investigator, University of Birmingham), Daniel Slade (Cancer Research UK Clinical Trials Unit (CRCTU), University of Birmingham), Charlotte Gaskell (CRCTU, University of Birmingham), Dr Francis Mussai (University of Birmingham), Prof Gary Middleton (University of Birmingham), Dr Dhruv Parekh (University of Birmingham), Dr Anna Rowe (CRCTU, University of Birmingham), Prof. Duncan Richards (University of Oxford), Dr Mathew Rowland (University of Oxford), Prof. Julian Bion (University of Birmingham), Prof. Simon Gates (Cancer Research UK Clinical Trials Unit (CRCTU), University of Birmingham), Prof. Pam Kearns (CRCTU, University of Birmingham), Dr Rowena Sharpe (CRCTU, University of Birmingham), Dr Sarah Bowden (CRCTU, University of Birmingham), Prof. David Thickett (University Hospitals Birmingham), Dr Mansoor Bangash (University Hospitals Birmingham), Prof. Fang Gao-Smith (University Hospitals Birmingham), Dr Jaimin Patel (University Hospitals Birmingham), Prof. Elizabeth Sapey (University Hospitals Birmingham), Dr Mark Thomas (University of Birmingham), Prof. Mark Coles (University of Oxford), Prof. Peter Watkinson (University of Oxford), Prof. Naj Rahman (University of Oxford), Prof. Ling-Pei Ho, (University of Oxford), Prof. Brian Angus (University of Oxford), Dr Alexander J. Mentzer (University of Oxford), Dr Alex Novak (University of Oxford), Prof. Marc Feldmann (University of Oxford).

CRCTU staff: Camilla Bathurst, Joseph van de Wiel, Susie Mee, Karen James, Bushra Rahman, Karen Turner.

Data Monitoring Committee

Member	DMC Role	Contact Information
Professor Adam Hill	Chair	University of Edinburgh
Professor Christina Yap	Independent Statistician	The Institute of Cancer Reserarch, London
Professor Anthony Gordon		Imperial College London

Trial Steering Committee

Member	TSC Role	Contact Information
Professor Michael Matthay	Chair	University of California San Francisco
Professor Danny McAuley	Deputy Chair	Queen's University Belfast
Mr Andrew Hall	Independent Statistician	University of Leeds
Professor Paul Dark		Manchester University
Professor Sir Andrew McMichael		University of Oxford

Scientific Advisory Board

Member		Contact Information
Professor Philip Newsome	Chair	University of Birmingham
Professor Pamela Kearns		University of Birmingham
Professor Tonny Veenith		University of Birmingham
Dr Benjamin Fisher		University of Birmingham

Dr Rowena Sharpe		University of Birmingham
Professor Julian Bion		University of Birmingham
Professor Bryan Williams		University College London
Professor Duncan Richards		University of Oxford
Professor Graham Cooke		Imperial College London
Professor Helen McShane		University of Oxford
Professor Ling-Pei Ho		University of Oxford
Dr Rebecca Turner		University College London
Professor Simon Gates		University of Birmingham
Professor Vincenzo Libri		University College London

Supplementary Text Methods

Participant Eligibility

Exclusion criteria included planned palliative care, pregnancy or breastfeeding, women of childbearing potential and non-vasectomised men who were unwilling to use effective contraception for the duration of the trial and throughout the drug-defined post-trial period, known HIV or chronic hepatitis B or C infection, concurrent immunosuppression with biological agents, a history of haematopoietic stem cell or solid organ transplant, known hypersensitivity to drug products or excipients, tuberculosis or other severe infections such as (non-SARS-CoV-2) sepsis, abscesses, and opportunistic infections requiring treatment, moderate or severe heart failure (NYHA class III/IV), or any other indication or medical history, that in the opinion of the patient's local investigator, made the patient unsuitable for trial participation. Co-enrolment into other interventional trials was not permitted with the exception of the RECOVERY-Respiratory Support trial comparing continuous positive airway pressure or high flow nasal oxygen to standard care, as this met current UK guidance on mechanistic independence in co-enrolment.[19]

Additional detail on randomisation

The in-house system for randomisation was managed by the programming team at the CRCTU. Site research staff would enter data via eCRFs. The system was designed with the capability of turning off arms or allowing for the addition of new arms given the platform nature of the trial. Programming were to be informed of any modifications to be made following the outcome of interim analyses and implemented them accordingly.

Data Handling

The data was stored securely within a relational database with the raw datasets only accessible by the trials team. Data was entered onto the system at sites through the use of eCRFs. The full details remain in the protocol.

Recommendations

CRP data was considered by the data monitoring committee (DMC) in the context of the emerging safety data to make a recommendation as outlined below:

a) If there is strong evidence of an additional anti-inflammatory effect (CRP) and a satisfactory safety profile consider progression to clinical endpoint evaluation whether in this trial or in another one;
b) If there is no evidence of additional biological effect or an unfavourable safety signal, then terminate arm and do not proceed.

Simulations to Inform Sample Size

The simulations and tables below demonstrate the operating characteristics of a trial design with the chosen decision criteria, based on a simpler analysis of the area under the curve for sequential CRP data, with effect sizes informed from a dataset from 1026 hospitalised COVID-19 patients at Queen Elizabeth Hospital, Birmingham. In our simulations, we compared a traditional fixed trial design recruiting 120 patients with candidate adaptive designs. We present basic operating characteristics for the fixed design (Table 6A) and the chosen adaptive design (Table 6B). We studied six scenarios of treatment effect, and estimated, through simulation, the probability of a trial stopping early for "success" or "futility," and ultimately concluding success. Simulations were performed in Fixed and Adaptive Clinical Trial Simulator (FACTS) software using default non-informative priors.

Table A. Operating characteristics for a fixed trial design of 120 patients.

Scenario	Probability stopping early for success	Probability stopping early for futility	Overall probability of success	Mean number of patients
Null	0	0	0.101	120
A	0	0	0.537	120
B	0	0	0.926	120
C	0	0	0.997	120
D	0	0	0.008	120
E	0	0	0	120

Scenarios A, B, and C are beneficial effects of the intervention with (true) treatment effects of 0.25, 0.5 and 0.75 standard deviations, "null" is zero treatment effect and D and E are harmful effects of 0.25 and 0.5 standard deviations. "success" and "futility" are defined as above.

Table B. Operating characteristics for an adaptive design with interim analyses at 40 and 80 patients.

Scenario	Probability stopping early for success	Probability stopping early for futility	Overall probability of success	Mean number of patients
Null	0.148	0.624	0.176	66
A	0.455	0.281	0.559	70
B	0.798	0.089	0.890	59
C	0.965	0.012	0.985	48
D	0.03	0.901	0.031	52
E	0.003	0.986	0.003	43

The adaptive design achieves similar probabilities of success in scenarios where the treatment effect is truly beneficial (A, B and C), and increases the probability of success only slightly if the intervention is harmful (D and E). There is some increase in the probability of success if the treatment effect is zero (Type I error) but this is offset by the very substantial reductions in the numbers of patients needed in all scenarios. Moreover, Type I error is not a serious problem as all interventions would be evaluated further in phase III trials.

Additional information on interim analyses

The interim analyses for the DMC were conducted by the trial statisticians. Only the statisticians and the DMC members, who were independent from the operation of the trial, had access to the results in confidence

Supplementary Text Results

Fitted Model – Namilumab

$$Log(CRP) = 3.41 + a - 0.24 * Time + b + 0.01 * Time^2 - 0.22 * CareStatus + 0.02 * Age + 0.20 * Trt - 0.46 * CareStatus * Trt - 0.09 * Time * Trt + \varepsilon$$

Where

$$a \sim N(0, 0.73^2), b \sim N(0, 0.21^2), \varepsilon \sim N(0, 0.58^2)$$

CareStatus = 1 if on the ward and 0 for ICU, Trt = 1 if receiving Namilumab or 0 if usual care alone

Fitted Model – Infliximab

$$Log(CRP) = 3.70 + a - 0.34 * Time + b + 0.02 * Time^2 - 0.45 * CareStatus + 0.01 * Age - 0.43 * Trt + 0.06 * CareStatus * Trt - 0.06 * Time * Trt + \varepsilon$$

Where

$$a \sim N(0, 0.79^2), b \sim N(0, 0.18^2), \varepsilon \sim N(0, 0.65^2)$$

CareStatus = 1 if on the ward and 0 for ICU, Trt = 1 if receiving Infliximab or 0 if usual care alone

Supplementary Table 1. Summary of protocol changes

Amendment number	Date of approval	Protocol version number	Type of amendment	Summary of amendment
1	**REC:** 14-May-20	n/a	Substantial Amendment	Addition of Oxford and UCL as sites
2	**MHRA:** 29-May-20 **HRA:** 01-Jun-20	3.0	Substantial Amendment	Addition of two new IMPs: Namilumab and Infliximab. Update SOE, amendments to inclusion/ exclusion criteria. Specifically: New exclusion criteria relating to the addition of the new drugs: 1) Known hypersensitivity to drug products or excipients 2) Patients with tuberculosis or other severe infections such as (non-COVID-19) sepsis, abscesses, and opportunistic infections requiring treatment 3) Patients with moderate or severe heart failure (NYHA class III/IV)
3	**REC:** 10-Jun-20	n/a	Substantial Amendment	Addition of new sites
4	**MHRA:** 08-Jun-20	n/a	Substantial Amendment	IMPD update
5	**MHRA:** 12-Jun-20 **REC:** 12-Jun-20	4.0	Substantial Amendment	Amendment to inclusion criteria. Specifically: Inclusion criterion 1 changed to: 'Hospitalised adult (≥16 yrs) patients with a clinical picture strongly suggestive of SARS-CoV-2 pneumonia (confirmed by chest X-ray or CT scan, with or without a positive reverse transcription polymerase chain reaction [RT-PCR] assay)' in order to: • Allow CT imaging as evidence for COVID-19 pneumonia • Allow recruitment of patients with strong clinical suspicion for COVID-19 pneumonia but with negative PCR assay Non-substantial amendments to Sample Collection Sub-study text.
6	**MHRA:** 19-Jun-20 **REC:** 20-Jun-20	5.0	Substantial Amendment	Amendment to exclusion criteria. Specifically: 'Concurrent immunosuppression with biological agents or prednisone dose > 20mg' Was changed to 'Concurrent immunosuppression with biological agents' in order to allow patients to be recruited on dexamethasone, following the RECOVERY data
7	**MHRA:** 12-Oct-20 **REC:** 12-Oct-20	6.0	Substantial Amendment	Change of Primary and Secondary Outcomes Specifically: Primary outcome changed to CRP (previously a secondary outcome) from the oxygen saturation to fractional inspired oxygen concentration (SpO2/FiO2) ratio, which now becomes a secondary outcome Hospital free days added as a secondary outcome Overall survival listed as a safety measure (previously death included under hospital survival status as a clinical outcome) Applicable changes to Inclusion/ Exclusion Criteria Specifically: Inclusion criteria changed from 'Oxygen saturation (SaO2) of ≤94% while breathing ambient air or a ratio of the partial pressure of Oxygen (PaO2) to the fraction of inspired oxygen (FiO2) (PaO2:FiO2) ≤ 300 mg Hg (≤40kPa'), to 'CRP ≥40' The following exclusion criteria that relate to the unopened Myelotarg arm were removed from general exclusion and made arm specific: • Known veno-occlusive disease • Neutrophil count < 2 x 109/l or White Blood Cell Count < 4.0 x 109/l The following exclusion criteria was removed as it was felt to be unnecessarily hindering recruitment: • Chronic Obstructive Pulmonary Disease (known FEV1 < 50% predicted or ambulatory or long term oxygen therapy

				Inclusion of Abbreviations list and eCRF table Update to Statistical Analysis section • Justification for CRP, operating characteristics and decision rules

Supplementary Table 2 World Health Organisation Clinical Progression Scale

Patient State	Descriptor	Score
Uninfected	Uninfected; no viral RNA detected	0
Ambulatory	Asymptomatic; viral RNA detected	1
	Symptomatic; independent	2
	Symptomatic; assistance needed	3
Hospitalised; mild disease	Hospitalised; no oxygen therapy	4
	Hospitalised; oxygen by mask or nasal prongs	5
Hospitalised; severe disease	Hospitalised; oxygen by NIV or high flow	6
	Intubated and mechanical ventilation, $pO_2/FiO_2 \geq 150$ or $SpO_2/FiO_2 \geq 200$	7
	Mechanical ventilation $pO_2/FiO_2 < 150$ ($SpO_2/FiO_2 < 200$) **or** vasopressors	8
	Mechanical ventilation $pO_2/FiO_2 < 150$ ($SpO_2/FiO_2 < 200$) **and** vasopressors, dialysis or ECMO	9
Death	Dead	10

Adapted from WHO Working Group on the Clinical Characterisation and Management of COVID-19 infection. A minimal common outcome measure set for COVID-19 clinical research. Lancet Infect Dis. 2020;20:e192-e197.

Footnotes for use in CATALYST

1. If pO_2 not available then use the SpO_2/FiO_2 ratio instead

2. For pO_2 measurements in kPa, use an online calculator e.g. https://www.msdmanuals.com/en-gb/medical-calculators/PaO_2_FiO_2Ratio.htm to calculate a pO_2/FiO_2 ratio equivalent to that obtained with pO_2 measured in mmHg, or else consider an equivalent ratio to 200, when dividing pO_2 in kPa by FiO_2, is 26.7, and an equivalent to 150 is 20.

3. If medically fit for discharge, record status as for ambulatory patient

4. Asymptomatic implies a return to baseline symptomatic state, i.e. no fever, and no cough, shortness of breath, confusion, myalgia, diarrhoea, fatigue, or weakness above what the participant would have experienced on a daily basis before their COVID-19 episode

5. Symptomatic but independent, implies that the participant has some of the additional symptoms as above, but needs no additional help with activities of daily living above what they required prior to their COVID-19 episode.

6. Symptomatic but needs assistance, implies that in addition to having symptoms as above, they require help with activities of daily living i.e. bathing/showering, personal hygiene and combing of hair, dressing, toileting, mobility/transferring and self-feeding, above what they required on a daily basis prior to their COVID-19 episode.

7. Score 0 (uninfected: no viral RNA detected) is not being assessed as part of CATALYST.

Supplementary Table 3. Point estimates and associated 95% credible intervals for mean posterior predictive expected ln(CRP) values on days 1, 7 and 14 for ward and ICU patients allocated to usual care alone or namilumab plus usual care and for the differences between these groups (Δ). Conditional effects data derived from Bayesian multi-level regression.

Care Status	Day	Usual Care (Control)	Usual Care + Namilumab	Δ
Intensive Care Unit	1	4.39 (4.02, 4.77)	4.59 (4.25, 4.93)	0.2 (-0.3, 0.7)
Intensive Care Unit	7	3.26 (2.75, 3.73)	2.91 (2.44, 3.37)	-0.35 (-1.02, 0.35)
Intensive Care Unit	14	2.74 (1.82, 3.62)	1.74 (0.82, 2.65)	-1.00 (-2.26, 0.26)
On Ward	1	4.17 (3.86, 4.48)	3.91 (3.62, 4.21)	-0.26 (-0.67, 0.17)
On Ward	7	3.04 (2.60, 3.49)	2.23 (1.77, 2.68)	-0.81 (-1.45, -0.18)
On Ward	14	2.53 (1.64, 3.41)	1.06 (0.13, 1.97)	-1.96 (-4.92, 0.79)

Supplementary Table 4. Point estimates and associated 95% credible intervals for mean posterior predictive expected ln(CRP) values on days 1, 7 and 14 for ward and ICU patients allocated to usual care alone or infliximab plus usual care and for the differences between these groups (Δ). Conditional effects data derived from Bayesian multi-level regression.

Care Status	Day	Usual Care (Control)	Usual Care + Infliximab	Δ
Intensive Care Unit	1	4.58 (4.09, 5.09)	4.15 (3.58, 4.72)	-0.43 (-1.19, 0.32)
Intensive Care Unit	7	3.21 (2.62, 3.80)	3.13 (2.44, 3.81)	-0.08 (-0.99, 0.82)
Intensive Care Unit	14	3.33 (2.26, 4.35)	3.66 (2.46, 4.79)	0.32 (-1.19, 1.77)
On Ward	1	4.13 (3.72, 4.54)	3.76 (3.35, 4.16)	-0.37 (-0.94, 0.20)
On Ward	7	2.76 (2.23, 3.30)	2.74 (2.19, 3.27)	-0.02 (-0.79, 0.73)
On Ward	14	2.88 (1.86, 3.89)	3.27 (2.17, 4.29)	0.52 (-2.15, 2.94)

Supplementary table 5. Point estimates of the probability of being at each level of the WHO Clinical progression score on days 1, 14 and 28 for ward and ICU patients allocated to usual care alone or namilumab plus usual care. Conditional effects data derived from Bayesian longitudinal proportional odds ordinal regression.

			WHO score level									
	Care Status	Day	1	2	3	4	5	6	7	8	9	10
Usual Care (Control)	Intensive Care Unit	1	0.00	0.01	0.00	0.01	0.10	0.28	0.23	0.16	0.17	0.04
	Intensive Care Unit	14	0.03	0.13	0.03	0.06	0.22	0.20	0.08	0.05	0.07	0.14
	Intensive Care Unit	28	0.18	0.26	0.03	0.05	0.10	0.09	0.04	0.02	0.03	0.19
	On Ward	1	0.00	0.12	0.04	0.11	0.35	0.27	0.08	0.02	0.01	0.00
	On Ward	14	0.09	0.39	0.05	0.08	0.14	0.08	0.05	0.04	0.04	0.04
	On Ward	28	0.31	0.31	0.02	0.04	0.07	0.04	0.02	0.02	0.03	0.14
Usual Care + Namilumab	Intensive Care Unit	1	0.00	0.01	0.00	0.01	0.10	0.29	0.23	0.16	0.16	0.04
	Intensive Care Unit	14	0.06	0.26	0.04	0.09	0.21	0.12	0.05	0.04	0.07	0.08
	Intensive Care Unit	28	0.34	0.30	0.02	0.03	0.06	0.05	0.02	0.01	0.03	0.13
	On Ward	1	0.00	0.12	0.05	0.12	0.35	0.26	0.07	0.02	0.01	0.00
	On Ward	14	0.18	0.48	0.03	0.05	0.08	0.07	0.04	0.02	0.02	0.03
	On Ward	28	0.54	0.22	0.01	0.02	0.03	0.03	0.03	0.02	0.03	0.07

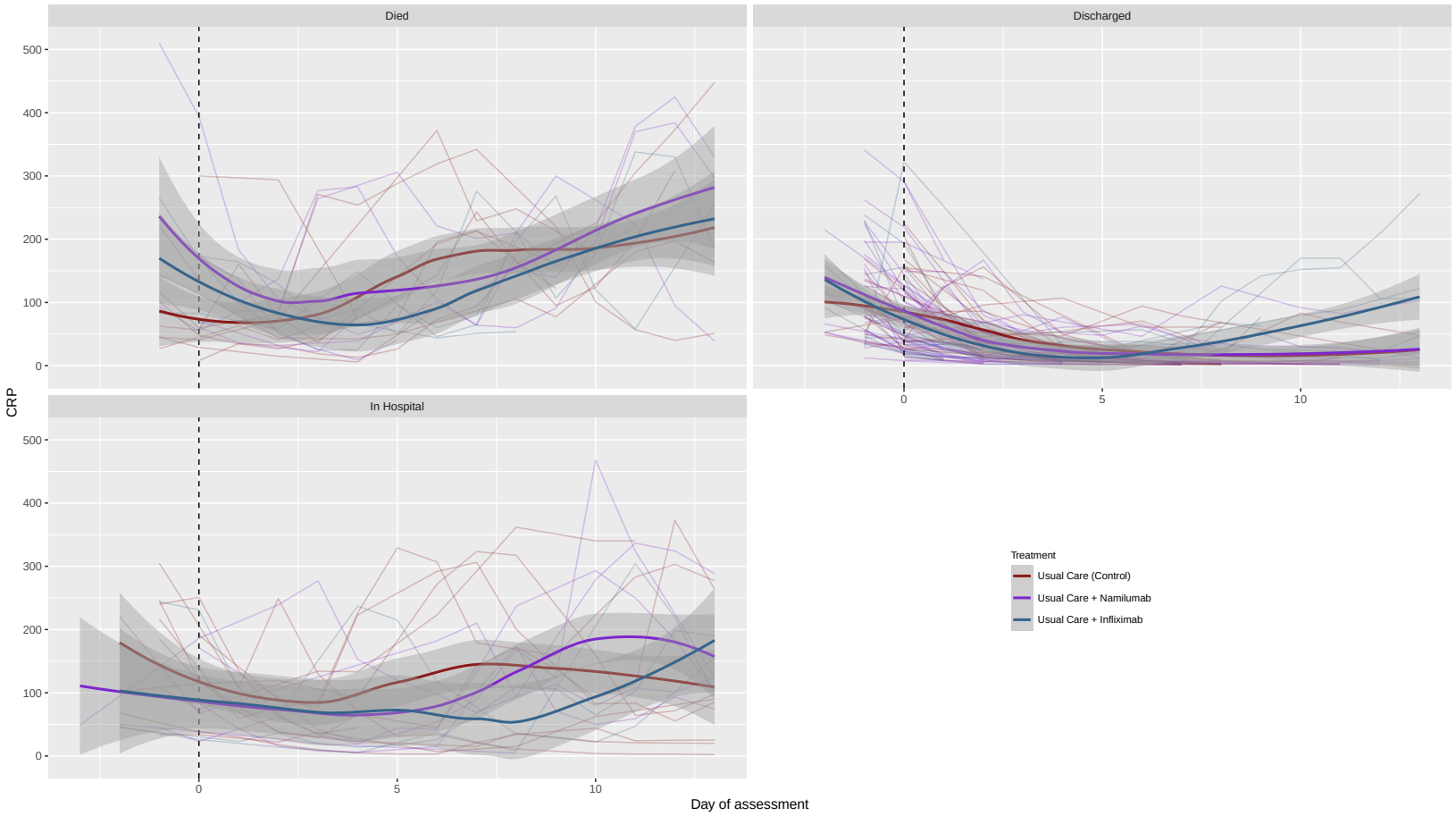
Supplementary table 6. Point estimates of the probability of being at each level of the WHO Clinical progression score on days 1, 14 and 28 for ward and ICU patients allocated to usual care alone or infliximab plus usual care. Conditional effects data derived from Bayesian longitudinal proportional odds ordinal regression.

			WHO score level									
	Care Status	Day	1	2	3	4	5	6	7	8	9	10
Usual Care (Control)	Intensive Care Unit	1	0.00	0.00	0.00	0.01	0.06	0.21	0.26	0.25	0.14	0.07
	Intensive Care Unit	14	0.03	0.15	0.07	0.08	0.18	0.18	0.09	0.05	0.04	0.14
	Intensive Care Unit	28	0.21	0.28	0.06	0.05	0.07	0.08	0.04	0.02	0.01	0.17
	On Ward	1	0.00	0.09	0.09	0.13	0.28	0.27	0.10	0.02	0.00	0.00
	On Ward	14	0.11	0.42	0.09	0.07	0.09	0.05	0.04	0.05	0.04	0.03
	On Ward	28	0.40	0.28	0.04	0.03	0.04	0.02	0.02	0.02	0.01	0.13
Usual Care + Infliximab	Intensive Care Unit	1	0.00	0.00	0.00	0.01	0.08	0.23	0.26	0.24	0.12	0.05
	Intensive Care Unit	14	0.02	0.12	0.05	0.07	0.18	0.19	0.10	0.07	0.04	0.15
	Intensive Care Unit	28	0.16	0.24	0.06	0.06	0.09	0.09	0.05	0.04	0.02	0.19
	On Ward	1	0.00	0.12	0.10	0.14	0.29	0.25	0.08	0.02	0.00	0.00
	On Ward	14	0.09	0.39	0.10	0.08	0.10	0.07	0.05	0.04	0.04	0.05
	On Ward	28	0.31	0.31	0.05	0.04	0.05	0.04	0.02	0.02	0.02	0.14

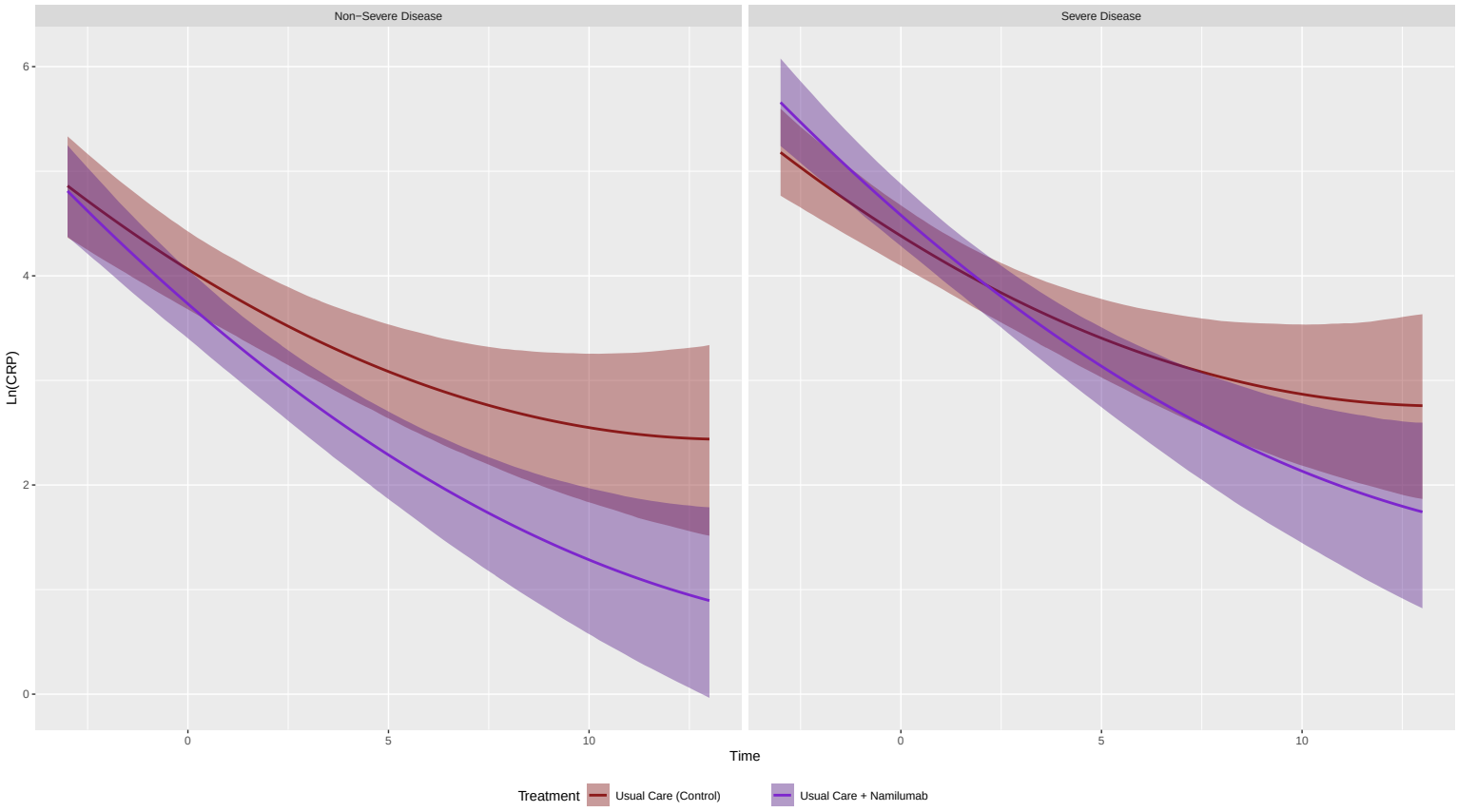
Supplementary Table 7. Total recorded adverse events, SAEs and deaths due to any cause. Specific diagnoses relate to adverse events that are CTCAE grade ≥3, secondary infection or allergic reaction. Only events occurring at least twice within an active drug/usual care comparison are shown. Data shown are number of adverse event occurrences (number of patients affected) in the safety population.

	Namilumab		**Infliximab**	
	Usual care n=54	Active arm n=55	Usual care n=34	Active arm n=29
Total reported adverse events (all grades)	145 (29)	134 (30)	112 (17)	102 (20)
Total adverse events (CTCAE grade ≥3, secondary infection or allergic reaction)	115 (24)	132 (25)	102 (16)	79 (16)
Total infection events	10 (7)	20 (8)	7 (4)	4 (4)
SAEs	10 (10)	10 (10)	5 (5)	6 (6)
Deaths	10 (10)	6 (6)	5 (5)	4 (4)
Anaemia	10 (6)	10(6)	8 (5)	2 (2)
Sinus bradycardia	2 (1)	0 (0)	2 (1)	0 (0)
Multiorgan failure	3 (3)	1 (1)	-	-
Covid pneumonia/pneumonitis	5(5)	4(4)	2 (2)	2 (2)
Lung infection	2(2)	1(1)	-	-
Pleural infection	2(1)	0(0)	2 (1)	0 (0)
Sepsis	1(1)	2(2)	-	-
Raised ALT	3(3)	5(5)	1 (1)	1 (1)
Raised Troponin I	0(0)	2(1)	-	-
Raised Creatinine	2(2)	4(3)	2(2)	2(1)
Raised CRP	5(4)	2(2)	5(4)	0(0)
Raised d-dimers	6(5)	3(3)	5(4)	1(1)
Raised ferritin	7(6)	5(5)	5(4)	11(7)
Low lymphocytes	16(12)	5(3)	11 (8)	4 (2)
Raised monocytes	0(0)	3(2)	-	-
Raised neutrophils	9(5)	5(3)	9 (5)	4 (4)
Raised white cells	9(6)	7(4)	9 (6)	9 (5)
Low platelets	1(1)	1(1)	-	-
Raised urea	9(7)	11(5)	8(6)	8(5)
Raised potassium	6(5)	1(1)	4(3)	2(1)
Raised sodium	1(1)	2(1)	-	-
Raised triglycerides	3(3)	1(1)	2(2)	6(4)
Low albumin	15 (13)	11 (7)	12 (10)	13 (8)
Low sodium	2(2)	1(1)	2(2)	2(2)
ARDS	-	-	1(1)	1(1)
Hypotension	-	-	0(0)	2(2)

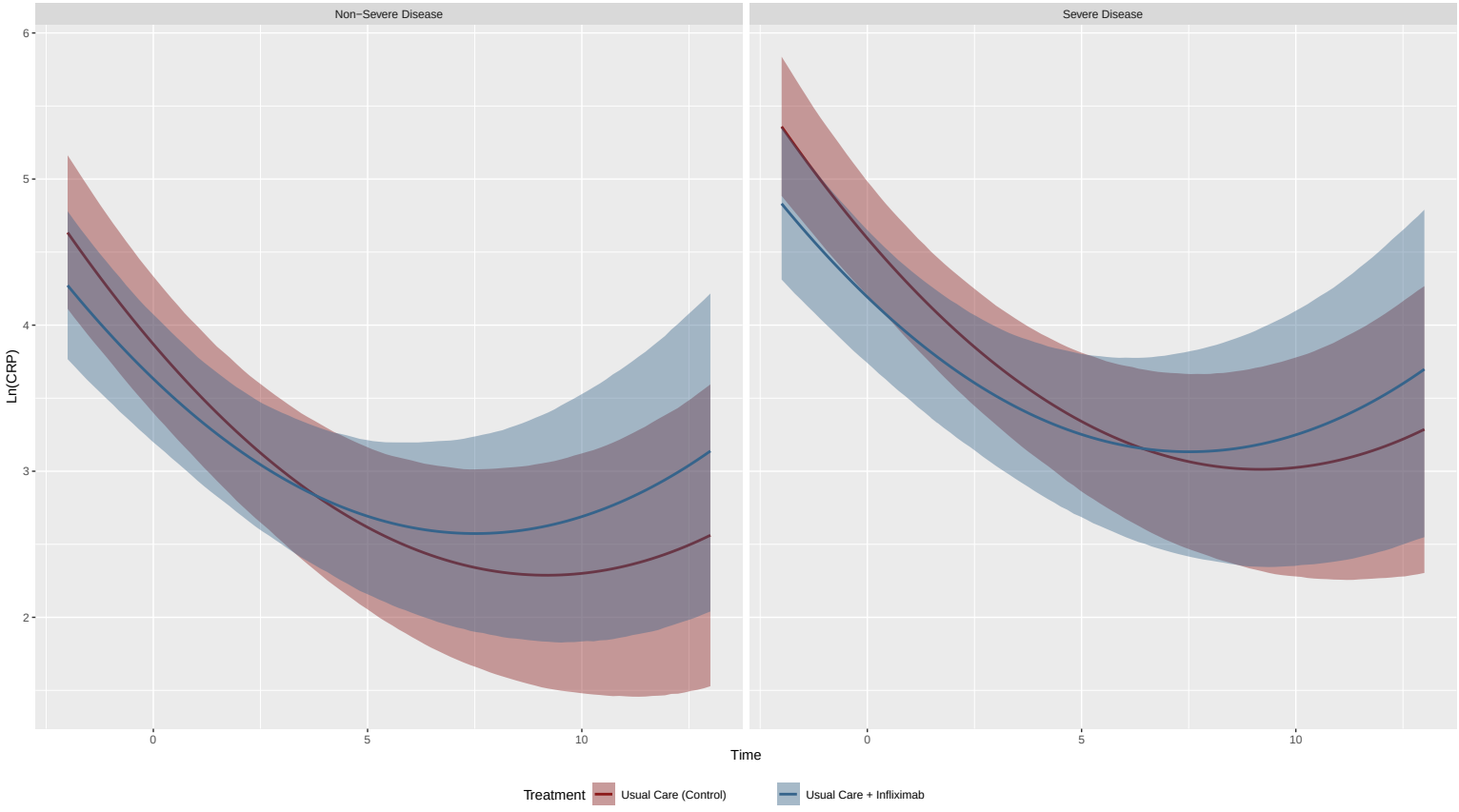
Supplementary Figure 1. CRP over time in relation to day 28 outcomes of death, discharge, and ongoing hospitalisation within the whole CATALYST modified intention to treat population.



Supplementary Figure 2. Conditional effects plots of CRP modelled over time in patients recruited in with non-severe and severe disease at baseline for namilumab (A) and infliximab (B). Severe disease was defined as the use of non-invasive or invasive ventilation or high flow nasal oxygen at baseline. Time 0 is day 1 of assessment.

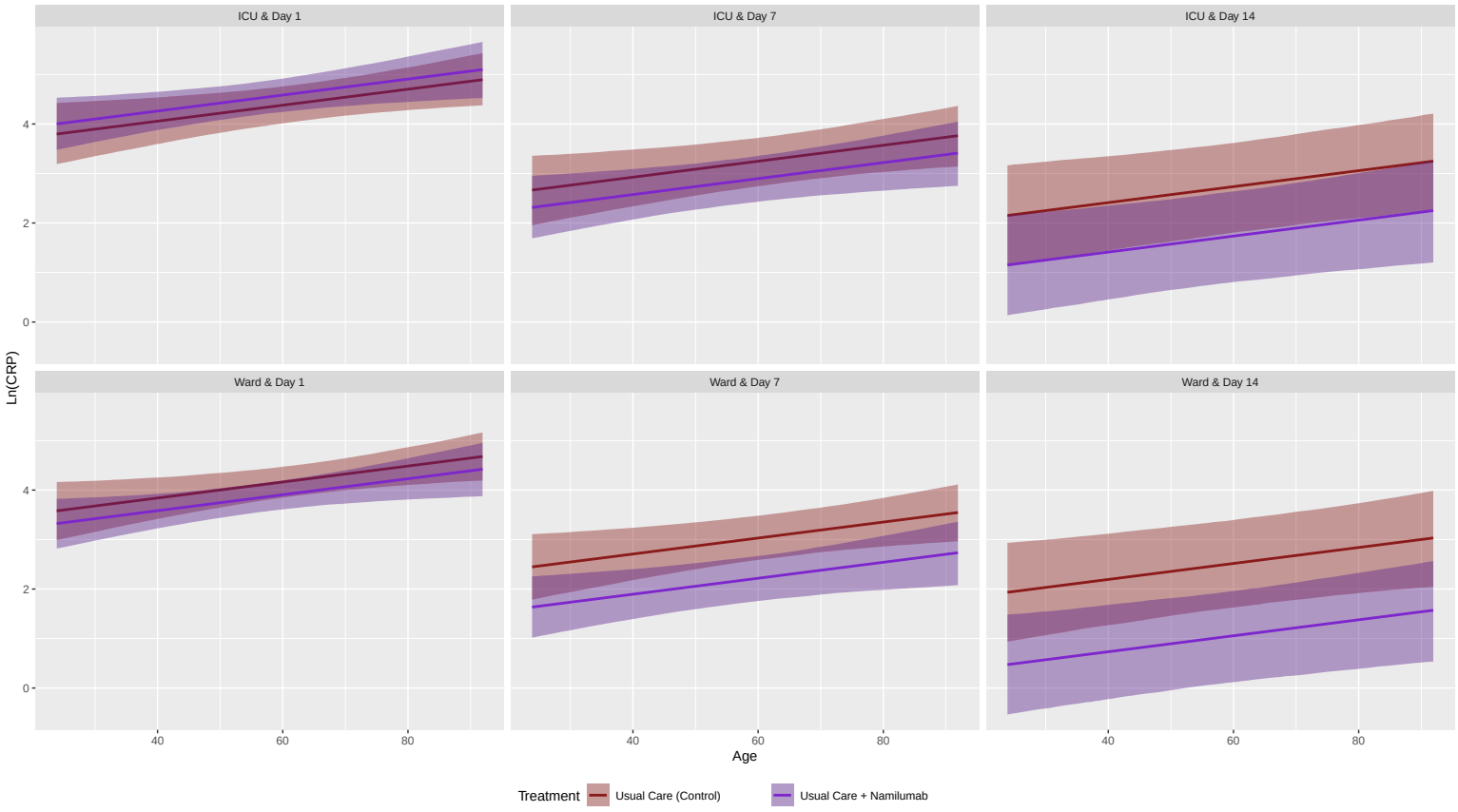


(a)

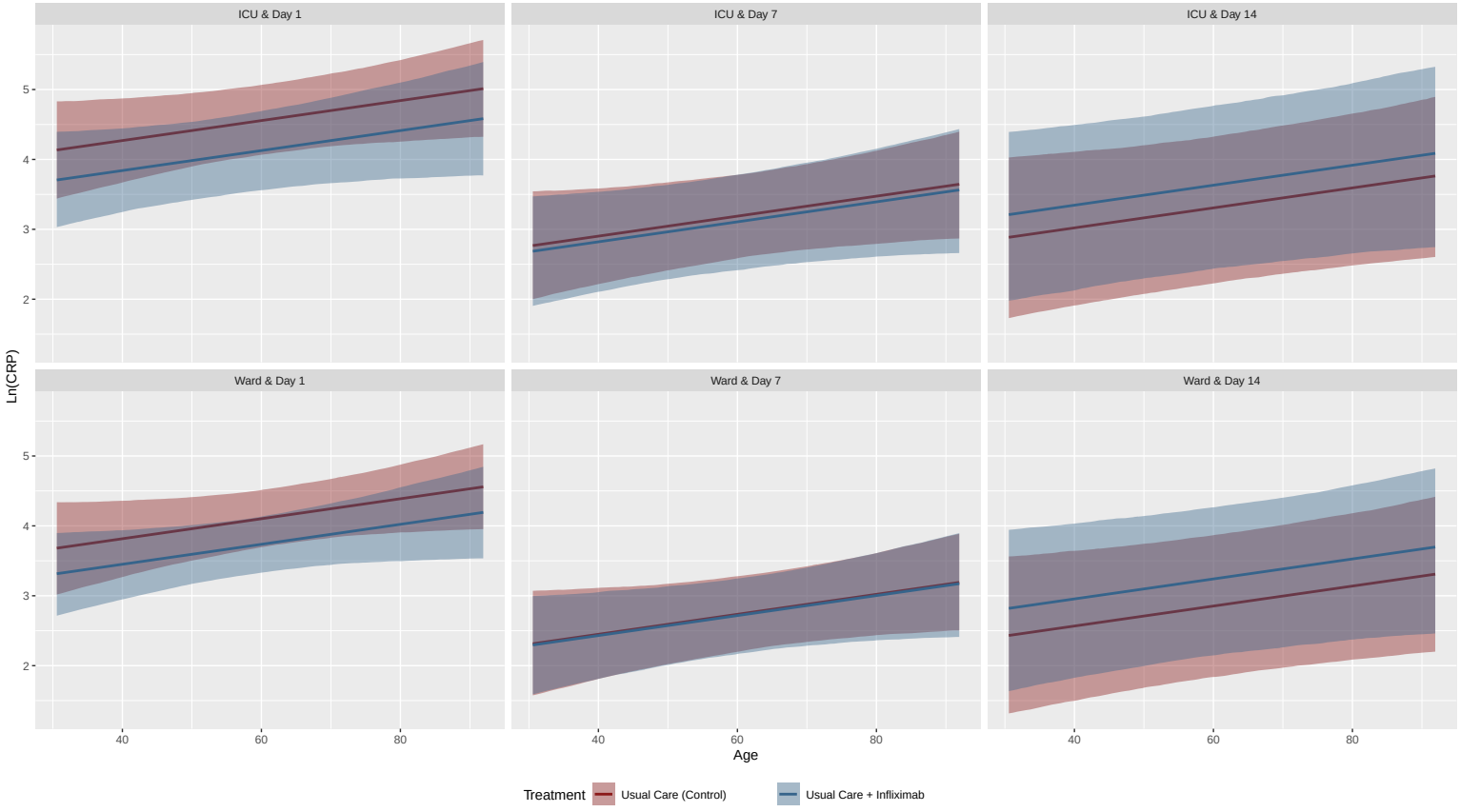


(b)

Supplementary Figure 3. Conditional effects model of CRP in relation to age and treatment at days 0, 7 and 14 in ward and ICU groups. CRP is associated with age but the effect of (A) namilumab and (B) infliximab on CRP is independent of age.

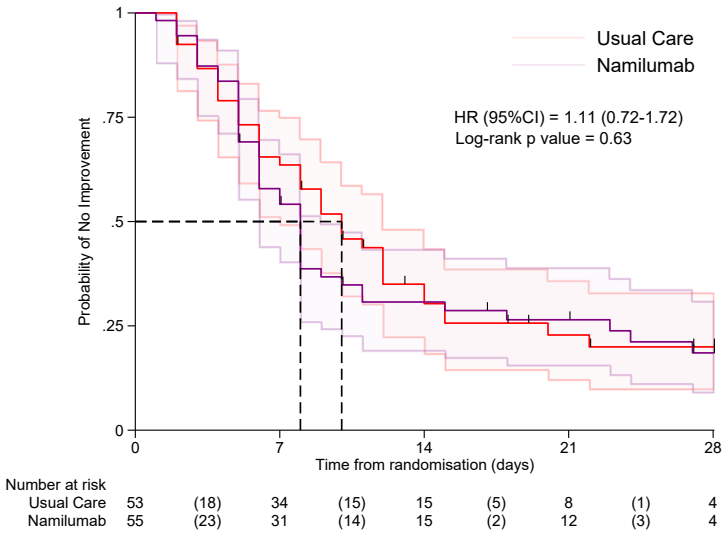


(a)

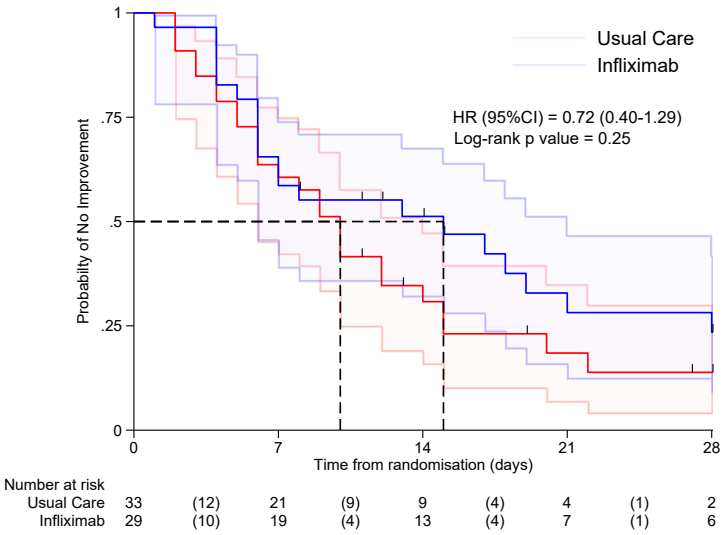


(b)

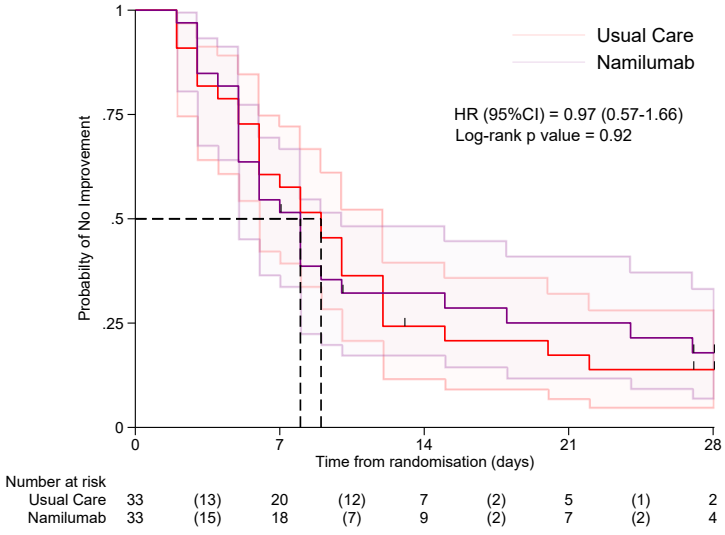
Supplementary Figure 4. Kaplan-Meier plots for time to 2 point improvement for whole population (A, B), ward (C, D) and ICU (E, F) for namilumab (A, C, E) and infliximab (B, D, F).



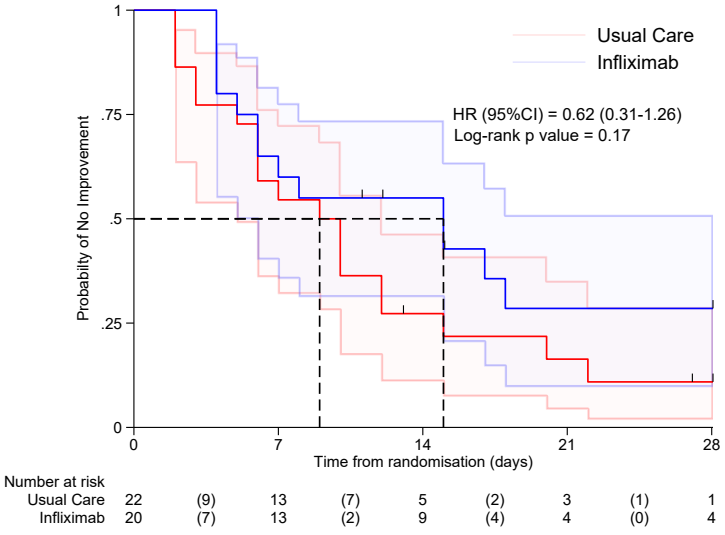
(a)



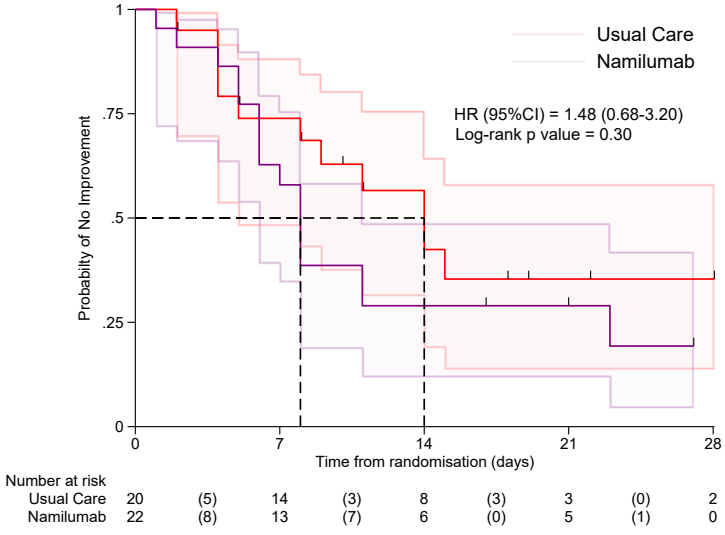
(b)



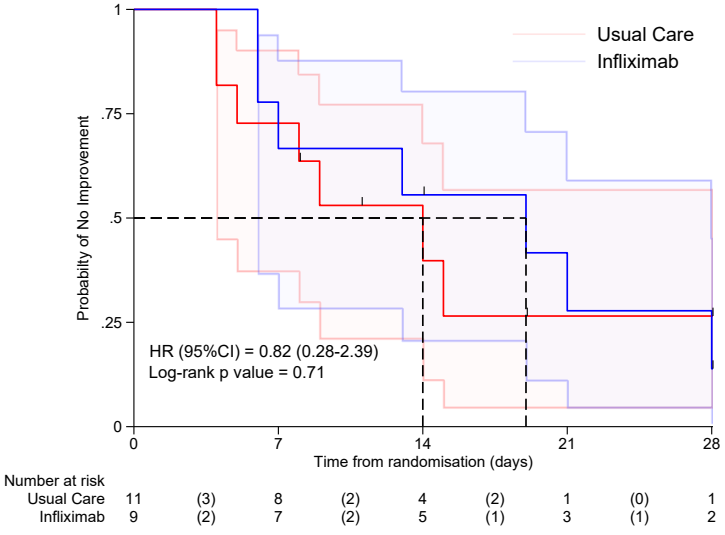
(c)



(d)

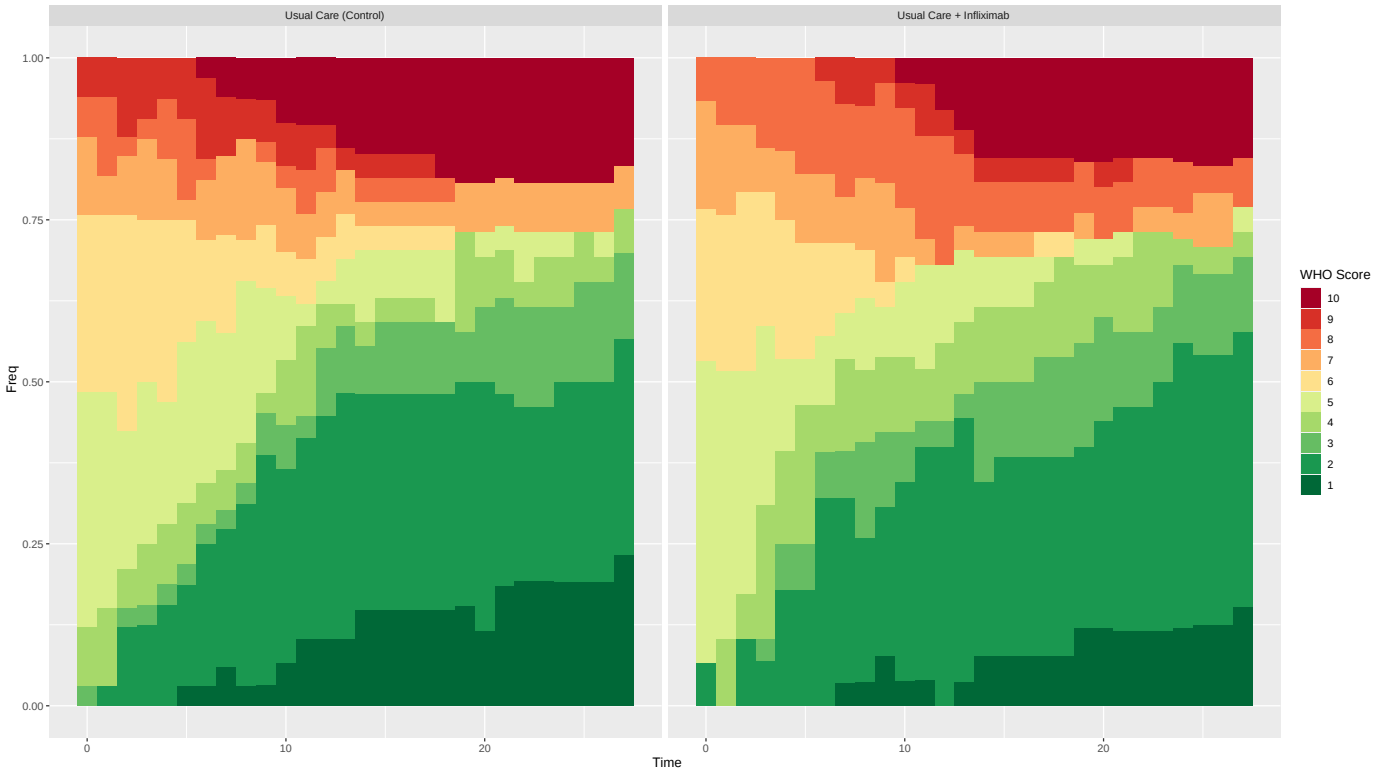


(e)

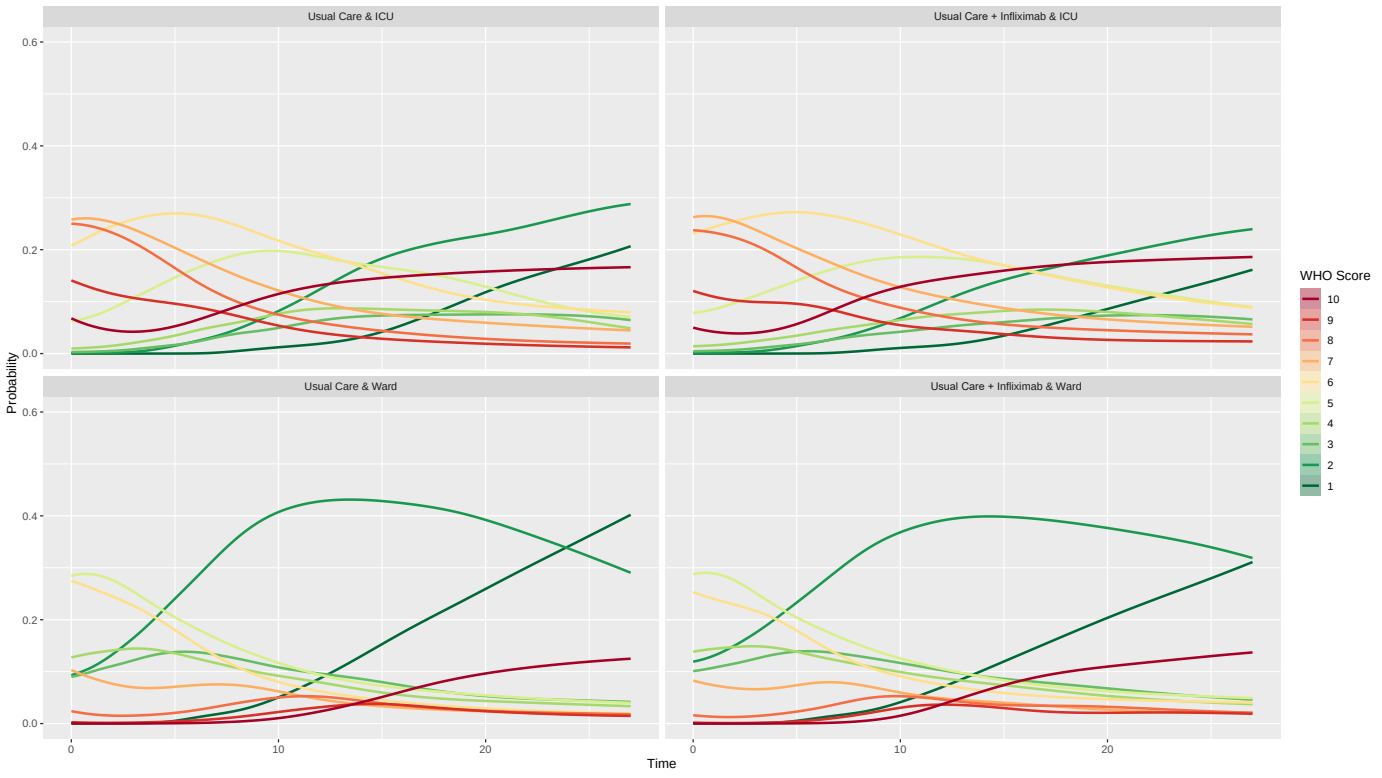


(f)

Supplementary Figure 5. WHO clinical progression score over 28 days for usual care versus infliximab. A, stacked bar chart of raw data for whole population eligible for comparison. B, conditional effects plots of WHO score modelled over time in days showing the probability of being at each level on each day for patients recruited in ICU and ward.

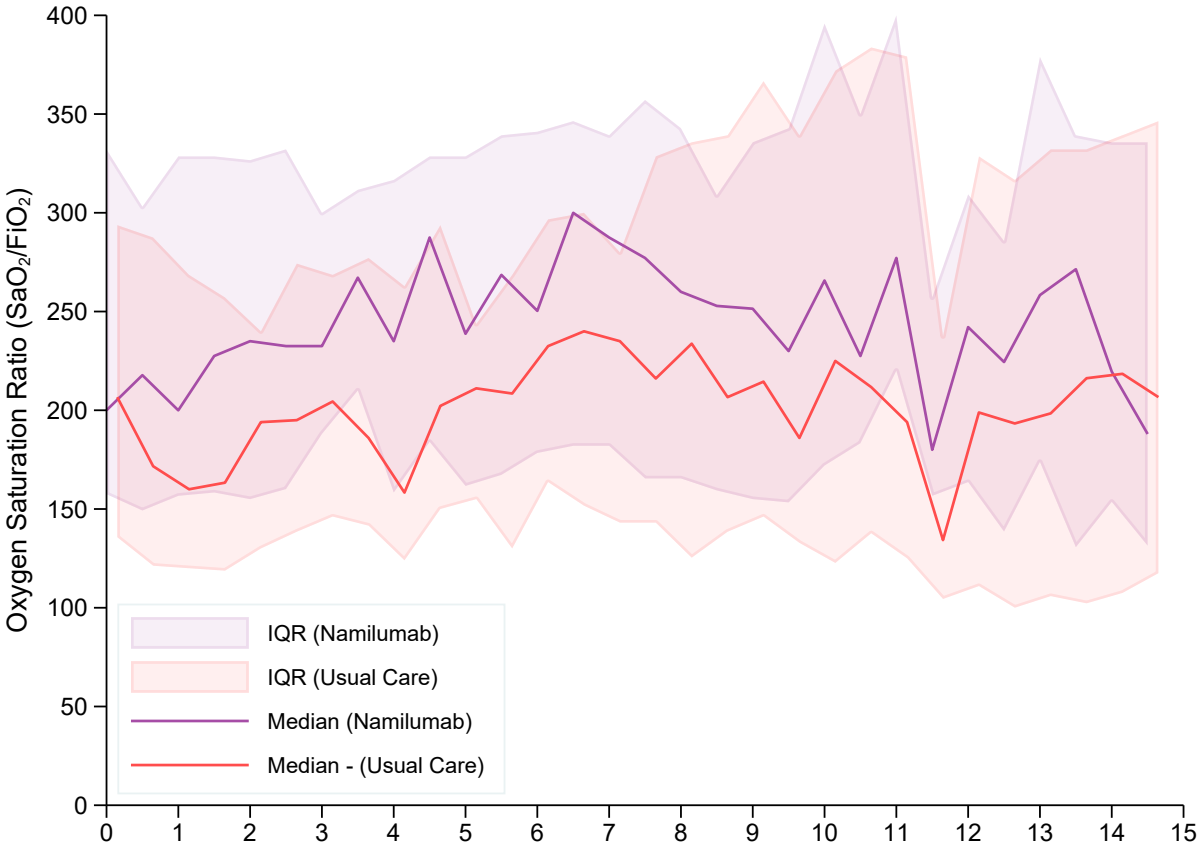


(a)

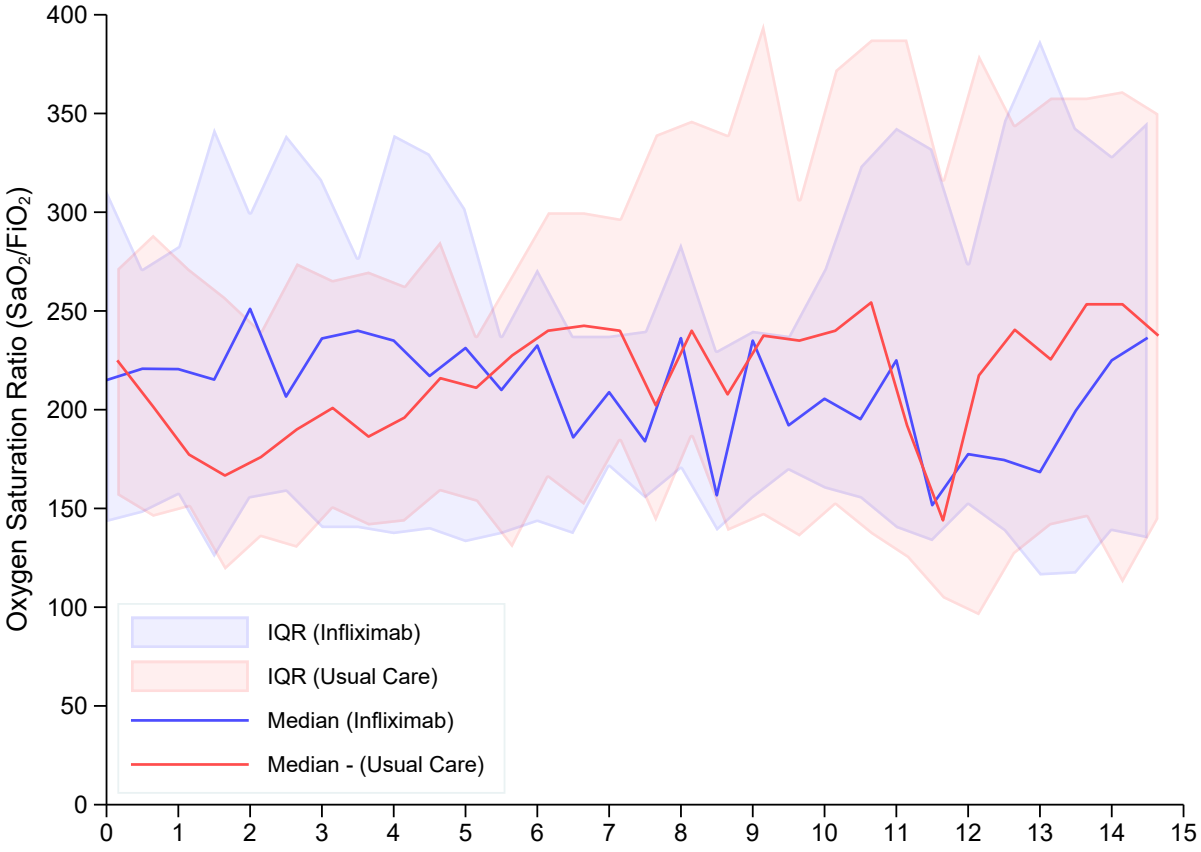


(b)

Supplementary Figure 6. Median oxygen saturation to fraction of inspired oxygen ratio (SF ratio) over time (days) for (A) namilumab (n=55 namilumab and n=54 usual care) and (B) infliximab (n=34 usual care and n=29 infliximab). Higher values indicate better oxygenation status.



(a)



Statistical Analysis Plan

A randomised phase II proof of principle multi-arm multi-stage trial designed to guide the selection of interventions for phase III trials in hospitalised patients with COVID-19 infection.

Version: 2.0
March 17, 2021

Sponsor(s): University of Birmingham
EudraCT Number: 2020-001684-89
Sponsor Reference Number: RG_20-030

Author(s):	
Name:	Mr Daniel Slade
Trial role:	Statistical Lead
Name:	Miss Charlotte Gaskell
Trial role:	Trial Statistician
Name:	Professor Simon Gates
Trial role:	Statistical Lead & Co-Investigator

Reviewed & Approved by:	
Name:	Dr. Benjamin Fisher
Trial role:	Acting Chief Investigator

Key personnel involved in the Statistical Analysis Plan:

Name	Trial role
Mr Daniel Slade	Statistical Lead
Miss Charlotte Gaskell	Trial Statistician
Dr. Tonny Venith	Chief Investigator
Dr. Benjamin Fisher	Deputy Chief Investigator
Professor. Simon Gates	Statistical Lead & Co-Investigator

Document Control Sheet:

Statistical Analysis Plan version:	Reason for update:
v1.0	Initial Version
v2.0	Incorporation of joint-modelling and AUC approach for the primary analysis, complimentary analyses to attempt to take account of censoring. Addition of ITT analysis for secondary endpoints and MITT definition for secondary endpoints. Modification of primary analysis to state inference will be based on only the interaction term of the model. Specification of subgroup analyses based on disease severity, a request of the DMC.

Contents

1 INTRODUCTION

1.1 Purpose of the Statistical Analysis Plan

This Statistical Analysis Plan (SAP) provides guidelines for the analysis and presentation of results for the Catalyst trial. This plan, along with all other documents relating to the analysis of this trial, will be stored in the 'Statistical Documentation' section of the Trial Master File. The statistical analysis will be carried out by the trial Statisticians.

1.2 Summary of the Trial

Trial Design
Catalyst is a rapid, open-label, phase II, multi-arm, multi-stage trial permitting an efficient evaluation of the potential efficacy of these targeted drugs which can then be considered for larger-scale testing by one of the current national platform trials.

Objectives
Primary Objectives

- To investigate whether candidate treatments demonstrate evidence of greater attenuation of inflammation as defined by an improvement in C-reactive protein (CRP) concentrations compared with usual care in COVID-19 patients.
- To recommend drugs that should be evaluated further in one of the phase III trials.

Outcome Measures
Primary Outcome Measures

- C-reactive protein measured over time up to day 14 for each patient.

Secondary Outcome Measures

- World Health Organisation (WHO) Clinical Progression improvement Scale (1-10 scale; for the purposes of this trial level 0, no viral RNA detected, will not be assessed)
- The ratio of the oxygen saturation to fractional inspired oxygen concentration (SpO2/FiO2), measured from randomisation to day 14, hospital discharge or death. SpO2 and FiO2 are measured as part of routine clinical care
- Respiratory rate
- Body temperature
- NEWS-2 score
- Length of hospital stay
- Hospital survival status at day 28 / hospital free days
- Proportion of patients discharged at day 28
- Destination of discharge
- Lymphocyte and Neutrophil counts and ratios
- Ferritin, D-Dimer and LDH
- Adverse events (AEs) and Serious Adverse Events (SAEs) as recorded by Common Terminology Criteria for Adverse Events (CTCAE), version 4.03 of grade $\geq$3 with interest in veno-occlusive disease (VOD), secondary infection and allergic reaction
- Overall Survival

Exploratory Outcome Measures

- Blood inflammatory mediators, biomarkers, transciptome and cellular immunology in relation to COVID-19 infection
- Viral load
- Host DNA assessed at baseline to assess for predictors of disease severity and drug response
- blood biomarkers of aveolar epithelial cell damage to include surfactant D and RAGE

Patient Population
This trial seeks to recruit hospitalised patients with COVID-19 who are hypoxic, admitted to either a hospital ward or ICU, and are at risk of deterioration.

Sample Size
A total of up to 60 patients per treatment arm will be recruited.

2 TIMING AND REPORTING OF INTERIM AND FINAL ANALYSES

There two planned interim analyses for the primary endpoint at n=20 and n=40 per arm respectively. Data analyses pertaining to trial conduct, data quality and patient safety will be supplied in confidence to an independent DMC throughout the period the trial is running. The DMC shall review the available data on a proposed 3 monthly basis.

The final analyses for the trial will be conducted once the end of trial has been reached. The final analyses will incorporate the primary, secondary and all exploratory outcomes as detailed in this analysis plan. The end of trial is defined as 6 months after the last data capture.

3 RECRUITMENT AND RANDOMISATION

3.1 Recruitment

At the point of analysis the following data will be reported:

- Date of the database snapshot used for recruitment analysis
- Total number of patients who have been recruited into the trial and randomised to each treatment arm
- Recruitment over time (monthly and cumulative)
- Recruitment by site

3.2 Randomisation

Patients will be randomised 1:1 between Usual Care (Control Arm) and interventional arms using the minimisation procedure described by Pocock and Simon, with a single stratification variable with two levels; Care status: 'On ward' or 'ICU'. Patients will be randomised into either a control group or to receive interventional treatments that are available at their site.

3.3 Ineligible Patients

Ineligible patients are defined as those registered patients who are subsequently found to not meet the eligibility criteria of the trial after being recruited. The proportion of ineligible patients and reasons for their ineligibility will be reported for each treatment arm. In addition the number of patients who were screened in total will be reported along with the number of patients not recruited to the trial and their associated reasons e.g. ineligible.

4 DATA QUALITY

4.1 Data Quality: CRFs

Patient data is collected using case report forms (CRFs) and electronic case report forms (eCRFs). Data collected in this way will be stored on a trial database. The trial database will be checked for missing data and any discrepancies at least annually but prior to any analysis as according to the trial specific data validation plan, which will be developed by both the trial statisticians and the trial coordinator.

4.2 Return Rates: CRFs

The proportion of returned CRFs compared to those that were expected will be reported for each case report form.

5 TRIAL POPULATION

5.1 Patient Characteristics

A summary of patient characteristics will be reported. Descriptive statistics will be provided in the summary including counts and percentages for categorical data items and mean (sd), median and ranges for continuous data items.

5.2 Definition of Populations for Analysis

Safety Population - Safety population will include all patients who receive any trial treatment. For interventional arms this requires the patient to have received some IMP.

MITT Population - The Modified Intention-To-Treat population for the primary analyses will include all patients who receive any trial treatment and who have a baseline CRP measurement and at least one further CRP measurement post baseline. For the secondary endpoints, this includes all patients who receive any trial treatment and have available data for the respective outcome measure.

ITT Population - This includes all randomised patients in their treatment arms, that have available data for the respective outcome measure.

6 TREATMENT RECEIVED

For each treatment arm, the proportion of participants who received treatment as per protocol will be reported. The proportion of participants who discontinued treatment early will also be reported along with a tabulation of the reasons. Summary statistics for all participants on treatment arms will be reported e.g. median/mean time on treatment, these statistics will be tailored for the specific arms as naturally the treatments may widely differ and thus different summary measures will be relevant.

7 SAFETY ANALYSIS

The number of serious adverse events (including SARs and SUSUARs), and the number of treatment-related deaths will be reported for each treatment arm. The reporting period for Adverse Events/Serious Adverse Events (SAE's) will commence from the date of consent. Safety will be assessed by looking at adverse events (CTCAE).

The following details will be reported for each treatment arm for all patients who are part of the safety population:

- Adverse events at baseline, summarised by event and number of patients experiencing such events.
- Max grade experienced for all patients.
- A summary of number of events and patients for all toxicities by event and grade.
- The number of events and patients for all grades of toxicities.
- All serious adverse events will be reported, details to be presented include but are not limited to; admitting event, other events, reason for SAE, outcome, sequel and relatedness.

8 ANALYSIS

For all analyses data will be analysed for each intervention against the control group, including in each analysis only those participants who were eligible for the those treatment arms at the point of randomisation. The primary analysis will be conducted on the MITT population and all secondary analyses will be conducted on both the MITT population and ITT unless otherwise specified.

New intervention arms may be added as new interventions become available. All comparisons will be performed temporally with regards to control arm data.

8.1 Analysis of Primary Outcome Measure

The CRP data will be modelled using Bayesian multi-level models that allow for nesting of the repeated measures data within patient, and allowing for non-linear responses. This approach will facilitate an assessment of the effects of the treatments on the CRP. Specifically, posterior probabilities for the treatment/time interaction term will be used to conduct decision making. Care status as a randomisation stratification factor will be incorporated accordingly into the model structure along with age as a known prognostic indicator.

At the specified decision points, with interim analyses at n=20 and n=40 and a final analysis at n=60 per arm, the CRP data will be considered in the context of the emerging safety data to make a recommendation as outlined below:

a) If there is strong evidence of an additional anti-inflammatory effect (CRP) and a satisfactory safety profile consider progression to clinical endpoint evaluation whether in this trial or in another one

b) Terminate arm and do not proceed (based on lack of evidence of an additional biological effect or of an unfavourable safety signal)

We will define that 'strong evidence' or 'success' will be if there is an 90% probability that the intervention arm is better than usual care in reducing CRP as seen by the treatment/time interaction covariate. 'lack of evidence' or 'futility' is defined as less than 50% probability of the intervention being better than usual care. However, given the large number of agents being investigated in various phase II trials, the size of effect and the totality of data will be reviewed before recommending adoption by a phase III platform.

In addition to the above analysis we will analyse the data using two further approaches, namely, modelling AUC and an additional joint-modelling approach for CRP and discharge/death, this is to ascertain if censoring events for CRP; discharge/death, have had any impact on inference and if so to model accordingly.

8.2 Analysis for Secondary Outcome Measures

Outcome measures

- World Health Organisation (WHO) Clinical Progression improvement Scale
  - Time to improvement, measured from the date of randomisation, an event here is defined as at least a one-point improvement on the Time to Clinical Improvement Scale. A Kaplan-Meier plot will be produced for each treatment and control arm comparison, estimates of median time to improvement will be reported along with associated confidence intervals (where they can be estimated). In addition to the one-point improvement an additional analysis utilising a two-point improvement will be conducted, to be comparable with other studies.
  - Patients' scores on the Clinical Improvement Scale for each day will be displayed graphically, and modelled using Bayesian longitudinal ordinal regression, as described by Harrell (`http://hbiostat.org/proj/covid19/bayesplan.html`).
- The ratio of the oxygen saturation to fractional inspired oxygen concentration (SpO2/FiO2) will be presented graphically over time.
- Length of hospital stay will be summarised via descriptive statistics, stratified by treatment group. Reasons for such lengths of stay will be reported and summarised accordingly.
- Respiratory rate, body temperature and NEWS-2, will be plotted over time and summarised through descriptive statistics. These measures may also be modelled over time using multilevel modelling. Exploratory data analysis will drive model formulation, assumptions will be tested accordingly. All modelling will be exploratory in nature.
- The proportion of patients discharged at day 28 along with destination of discharge will be presented accordingly.
- Hospital survival status at 28 days will be reported as a tabulation of the proportion of patients who have died, been discharged or are still in hospital by day 28. Hospital-free days will be summarised through descriptive statistics, patients still in hospital or who have died will be incorporated having 0 hospital-free days.

- Lymphocyte, neutrophil and full blood counts with lymphocyte: neutrophil ratios and ferritin, D-Dimer and Triglycerides LDH values will be plotted over time and summarised through descriptive statistics. These measures may also be modelled over time using multilevel modelling. Exploratory data analysis will drive model formulation, assumptions will be tested accordingly. All modelling will be exploratory in nature.
- AEs and SAEs will be analysed as per section 7
- Overall Survival - Measured from the date of registration, an event here is defined as death. Patients are followed up until they have either died or are censored at date last seen. A Kaplan-Meier plot will be produced for each comparison, estimates of median survival will be reported along with associated confidence intervals (where they can be estimated)

8.3 Subgroup Analysis

Exploratory subgroup analyses will be conducted to attempt to ascertain the effect of disease severity on outcomes. The subgroups of 'non-severe disease' and 'severe disease' are defined as those that have a baseline WHO score of < 6 and ≥ 6 respectively. Other exploratory subgroup analyses may be conducted based on known prognostic indicators e.g. age group.

9 SAMPLE SIZE

The tables below demonstrate the operating characteristics of a trial design with the chosen decision criteria, based on a simpler analysis of area under the curve for sequential CRP data, with effect sizes informed from a dataset from 1026 hospitalised COVID-19 patients at Queen Elizabeth Hospital, Birmingham.

It is anticipated that our proposed hierarchical analysis will have superior operating characteristics. In our simulations, we compared a traditional fixed trial design recruiting 120 patients with candidate adaptive designs. We present basic operating characteristics for the fixed design (Table 1) and the chosen adaptive design (Table 2). We studied six scenarios of treatment effect, and estimated, through simulation, the probability of a trial stopping early for "success" or "fultility," and ultimately concluding success. Scenarios A, B, and C are beneficial effects of the intervention with (true) treatment effects of 0.25, 0.5 and 0.75 standard deviations, "null" is zero treatment effect and D and E are harmful effects of 0.25 and 0.5 standard deviations. "success" and "futility" are defined as above.

Table 1: Operating characteristics for a fixed trial design of 120 patients

Scenario	Probability stopping early for success	Probability stopping early for futility	Overall probability of success	Mean number of patients
Null	0	0	0.101	120
A	0	0	0.537	120
B	0	0	0.926	120
C	0	0	0.997	120
D	0	0	0.008	120
E	0	0	0	120

The adaptive design achieves similar probabilities of success in scenarios where the treatment effect is truly beneficial (A, B and C), and increases the probability of success only slightly if the intervention is harmful (D and E). There is some increase in the probability of success if the treatment effect is zero (Type I error) but this is offset by the very substantial reducitions in the numbers of patients needed in all scenarios. Moreover, Type I error is not seen as a serious problem as all interventions would be evaluated further in Phase 3 trials.

Table 2: operating characteristics for an adaptive design with interim analyses at 40 and 80 patients

Scenario	Probability stopping early for success	Probability stopping early for futility	Overall probability of success	Mean number of patients
Null	0.140	0.607	0.143	70
A	0.471	0.254	0.573	74
B	0.847	0.062	0.910	61
C	0.974	0.010	0.989	51
D	0.030	0.918	0.031	54
E	0.003	0.985	0.003	47

10 STATISTICAL SOFTWARE

Statistical analyses will be carried out using relevant statistical software; SAS , Stata or R respectively. Version numbers and session details will be stated and logged with any analysis.

11 STORAGE AND ARCHIVING

Catalyst files are stored in a restricted access directory on a secure server and will be saved for archive purposes according to CRCTU policy and procedure.