GM-CSF targeting in COVID-19: an approach based on fragile foundations

Katharina Kohler[1] and Andrew Conway Morris ⓘ[1,2,3]

[1]Division of Anaesthesia, Department of Medicine, University of Cambridge, Cambridge, UK. [2]Division of Immunology, Department of Pathology, University of Cambridge, Cambridge, UK. [3]John V Farman Intensive Care Unit, Addenbrooke's Hospital, Cambridge, UK.

Corresponding author: Andrew Conway Morris (ac926@cam.ac.uk)

Shareable abstract (@ERSpublications)
Trials of anti-GM-CSF therapies in COVID-19 show divergent results; this may be explained by underlying biology and the fragility of the study findings. Further investigation of the pathophysiology of COVID-19 is required to better target therapies. http://bit.ly/3O1AuIo

Coronavirus disease 2019 (COVID-19) arises as a result of a pathological inflammatory response following infection with the coronavirus SARS-CoV-2. Although the majority of people infected with this virus will experience minimal or mild symptoms, a proportion will go on to develop more severe disease requiring hospitalisation and oxygen therapy. The most severe forms produce acute respiratory failure, necessitating mechanical ventilation or extracorporeal membrane oxygenation (ECMO). The advent of SARS-CoV-2 vaccination has substantially altered the risk profile of COVID-19, with marked reductions in the severity of illness and hospitalisation. However, for unvaccinated patients and those who do not mount an effective immune response to vaccination, it remains a potentially lethal infection.

Severe COVID-19 is marked by intense immune activation which usually develops days after viral loads have peaked [1]. Early transcriptional responses lead to the release of a broad range of immune mediators with a prominent role for type I, II and III interferons [2]. Patients with impaired interferon responses, due to factors such as autoantibodies [3] or obesity [4], tend towards a more severe course. A key pathological feature of severe COVID-19 is a macrophage-dominant infiltration of the lungs [5], and various investigators have noted similarities between the features of severe COVID-19 and secondary haemophagocytic lymphohistiocytosis [6] where macrophages are thought to play a major causal role [6]. Although incompletely characterised, there is cross-talk between T-lymphocytes and macrophages forming an "inflammatory network" which appears to underpin severe COVID-19 [7]. The early identification of elevations in a broad range of plasma cytokines, including tumour necrosis factor α (TNF-α) and the interleukins IL-1b and IL-6, provided the rationale for successful trials of both pleotropic immunomodulatory therapy with corticosteroids [8] and more selective targeting of IL-6 [9]. However, it is not clear that serum IL-6 levels are predictive of response to IL-6 blockade [10] whilst trials of agents targeting TNF-α [11] and IL-1b [12], as well as those seeking to supplement type I interferons [13], have not been successful.

Granulocyte–macrophage colony-stimulating factor (GM-CSF) plays several critical roles in inflammatory responses. It acts at multiple levels, including on the bone marrow to accelerate emergency granulopoiesis, as well as affecting a range of mature immune cells. Its effects include maintaining healthy function of alveolar macrophages [14] and restoration of function in sepsis-induced immune cell failure [15, 16]. GM-CSF acts in immunological niches, so plasma levels are often undetectable even in systemic inflammation, and a failure to detect elevated plasma levels does not automatically invalidate targeting this molecule therapeutically. Elevated levels have been identified in alveolar fluid from patients with acute respiratory distress syndrome, where elevated levels were associated with better outcomes [17]; conversely, elevated proportions of T-lymphocytes expressing GM-CSF were associated with worse outcomes in sepsis [18]. In COVID-19, serum levels are also generally undetectable, but heightened secretion of GM-CSF by T-lymphocytes in the sickest patients has been noted [19]. Such observational studies cannot distinguish

between a causal role for GM-CSF in driving immunopathology and reactive or even bystander elevation. Animal models of various inflammatory diseases do imply a pathogenic role for GM-CSF [20] and multiple anti-GM-CSF therapies are in clinical trials across a range of, mostly autoimmune, conditions [20]. There are a number of mechanisms whereby GM-CSF may exacerbate or drive pathology in COVID-19, including inhibition of signalling by interferons [21]. Whilst proven therapies such as tocilizumab and corticosteroids can reduce its secretion [22, 23], the data from fundamental and translational biology are unclear as to whether GM-CSF is helpful or harmful in COVID-19 [24]. This uncertainty has led to both studies of both inhaled recombinant GM-CSF [25] and GM-CSF blockade [11, 26–29]. The only phase 3 study published to date, the LIVE-AIR trial of lenzilumab, found a significant improvement in invasive ventilation-free survival in the treated group [27].

In this issue of the *European Respiratory Journal*, PATEL *et al.* [29] report a two part, phase 3 study of the anti-GM-CSF antibody otilimab. They initially included 793 patients in a multicentre randomised, placebo-controlled trial in adults (age ⩾18 years) with severe COVID-19 requiring non-invasive or invasive respiratory support with systemic inflammation (C-reactive protein (CRP) or ferritin above normal range). The primary endpoint, proportion of patients alive and free of respiratory failure at day 28, did not differ significantly between the groups (otilimab 71%, placebo 67%, with a model-adjusted absolute risk reduction of 5.3%; p=0.09). However, a pre-defined subgroup analysis did suggest a statistically and clinically significant benefit in older participants (age ⩾70), where a model-adjusted 19.1% increase in the proportion alive and free of respiratory failure was found following otilimab treatment, with a reduction in the key secondary outcome of day 60 mortality (p=0.04).

To explore this apparent differential effect on age, PATEL *et al.* [29] went on to conduct part two of the study, also reported in the same paper, restricted to patients over the age of 69 years. In this second part, 300 older patients were randomised with no benefit of otilimab found (52% of otilimab *versus* 51% of placebo treated patients reached the primary endpoint; p=0.86), and no significant signal to benefit found in day 60 mortality (43% *versus* 45%; p=0.69).

The failure to reproduce the effect seen in the first part's older subgroup, as well as the discrepancy with demonstration of ventilation-free survival benefit in the LIVE-AIR [27] study, raise a number of questions. The first is whether the initial subgroup finding of benefit in older patients was biologically plausible. Age is one the strongest predictors of mortality in COVID-19 [30], so giving a potentially greater modifiable mortality. Chronic inflammation is also seen in older persons, so called "inflammaging" [31]. This is often accompanied by reduced antiviral responses [31], with impaired viral clearance leading to worsening organ failure, in a manner analogous to that recently reported in obesity [4]. However, it is unclear whether GM-CSF is pivotal to these immunosenescent effects and therefore the biological plausibility of the subgroup analysis found in part one is modest at best.

The second question is, why should two monoclonal antibodies targeting the same molecule (lenzilumab in LIVE-AIR [27] and otilimab in OSCAR [29]) have apparently different effects on the same disease? The OSCAR trialists demonstrated appropriate target engagement, with otilimab engaging high proportion of plasma GM-CSF, so the differences are unlikely to be due to failure of pharmacodynamic effect. The divergence may be due to differences in the patients recruited. Whilst OSCAR patients were all receiving more than simple oxygen therapy, in LIVE-AIR over half the patients were on simple oxygen or room air and therefore less severely unwell at the time of enrolment. Interestingly, a *post hoc* analysis of the LIVE-AIR study suggested that the beneficial effect of lenzilumab was restricted to patients with lower systemic inflammation (defined by a CRP level of <150 mg·dL^{-1}) [32]; however, median levels of CRP did not differ markedly between patients in either part of the OSCAR study and the LIVE-AIR study.

A further plausible explanation for the difference in results is that the LIVE-AIR result [27] occurred by chance. Using frequentist statistics with a conventional cut-off of p<0.05 leads, mathematically, to a 1 in 20 chance of a false positive result. An additional method of trial analysis is the fragility index [33]. The fragility index indicates how many patients would have had to have a different outcome within the trial setting for the result to have changed. Whilst fragility indices have their limitations and can depend on the statistical method used to calculate them [34], a low fragility index indicates that the results are not necessarily robust even if statistically significant results have been reported.

A recent review [35] of 47 randomised controlled trials in COVID-19 including studies on treatments, vaccines and interventions, found a median fragility index of the included trials was 4 (1–11), meaning if four patients had had different outcomes the studies would have lost statistical significance. The median fragility index of randomised controlled trials of pharmacological interventions specifically was lower at

2.5 (1–6) and, overall, the fragility quotient (fragility index divided by trial size) of many studies was less than 1%, indicating a lack of robustness in individual clinical trials.

The fragility index for LIVE-AIR is 0 [27], indicating that the result would be rendered non-significant simply by using a different statistical test (figure 1a). We can also use the fragility index to determine how robust "neutral" results, such as those reported in the OSCAR trial, are. The inverse fragility index (number of patients needed to give a statistically significant effect by Fisher's exact test) for part one of the OSCAR study was 11 (fragility quotient 2.7%), indicating an improvement in outcome of 2.7% of the cohort would have produced a positive result for otilimab (figure 1b). Although the part one subgroup analysis of older patients produced a significant result, it was also a fragile result (figure 1c). Subgroup analyses should be approached with caution and, as PATEL *et al.* [29] rightly did, considered hypothesis generating when the overall result is neutral. Part two of the OSCAR trial was robustly negative: it would have required an increase in positive outcomes of 10% (from 52 to 62%) to bring this aspect of the study into the significant outcome category (figure 1d). All studies are samples drawn from a population, and therefore prone to error in measurement; the size of that potential error rather than arbitrary probability cut-offs should determine our interpretation of results.

Headline results without formal publication or preprint have been announced for two further large studies of GM-CSF inhibitory strategies, an 807 patient study of GM-CSF receptor blockade with mavrilimumab [36] and the ACTIV-5/BET-B study [37], which was another trial of lenizilumab, restricted to patients with CRP <150 mg·dL^{-1}. Both studies were reportedly neutral, although full interpretation of these studies requires more details to be released. The OSCAR trialists should be commended for their transparency in rapidly releasing full details of an industry-sponsored neutral trial in a prominent journal. Rapid publication of trial results has been a notable problem in previous pandemics [38].

Although there is a therapeutic rationale in targeting GM-CSF in COVID-19, it is not based on strong mechanistic data from either humans or animal models and relies mostly on inference from other diseases. This is not unusual in COVID-19, given its novelty and the difficulty distinguishing causal from epiphenomenal or consequential effects in observational studies of patients. Although there was an
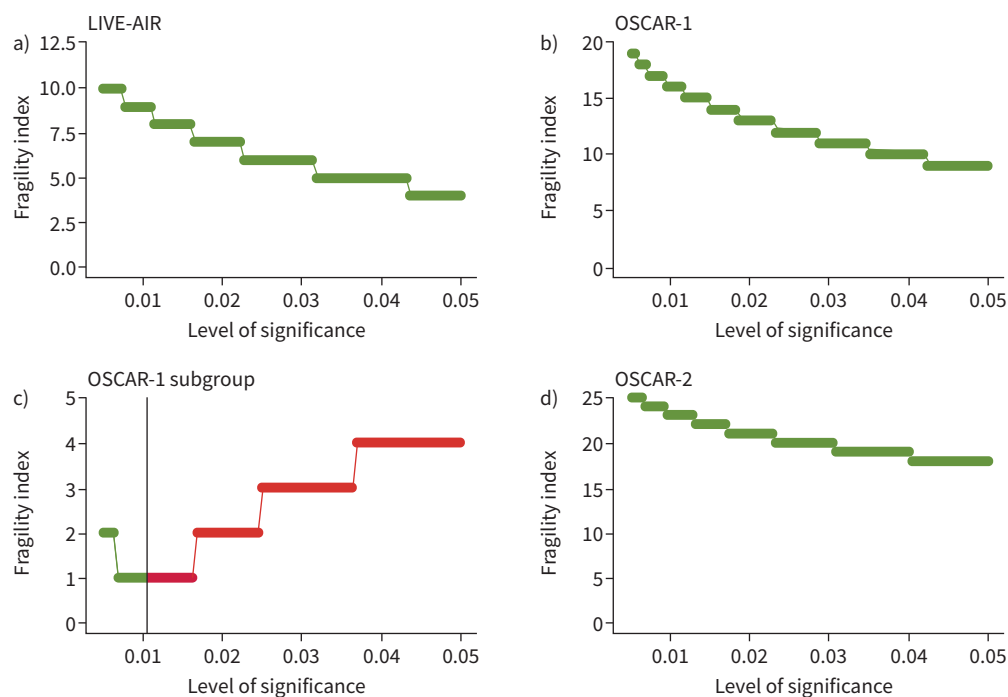

FIGURE 1 Plots of fragility index (*y*-axis) for each level of significance. a) LIVE-AIR study, b) OSCAR trial part one main cohort, c) OSCAR part one, subgroup ≥70 years old, and d) OSCAR part two. Green shows the areas of non-significance and red significance based on the level indicated on the *x*-axis based on Fisher's exact method. In panel c the vertical line indicates the p-value of the subgroup analysis.

understandable hurry to get therapies to patients, we must ensure that the robustness of both the underlying biological plausibility and trial results themselves are considered when evaluating reports of such therapies. With the current state of evidence there is no justification for the routine use of GM-CSF blockade in COVID-19, and we would suggest a more detailed understanding of its role in the biology of this disease is required before such approaches are trialled again.

References

1   Walsh KA, Jordan K, Clyne B, et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. J Infection 2020; 81: 357–371.

2   Sposito B, Broggi A, Pandolfi L, et al. The interferon landscape along the respiratory tract impacts the severity of COVID-19. Cell 2021; 184: 4953–4968.e16.

3   Bastard P, Rosen LB, Zhang Q, et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. Science 2020; 370: eabd4585.

4   Guo SA, Bowyer GS, Ferdinand JR, et al. Obesity is associated with attenuated tissue immunity in COVID-19. Am J Respir Crit Care Med 2022; in press [https://doi.org/10.1164/rccm.202204-0751OC].

5   Velu PP, Lucas CD, Conway Morris A. Post-mortem dissection of COVID-19: a pathogenic role for macrophages? Intensive Care Med 2021; 47: 1322–1325.

6   Retamozo S, Brito-Zerón P, Sisó-Almirall A, et al. Haemophagocytic syndrome and COVID-19. Clin Rheumatol 2021; 40: 1233–1244.

7   Bose P, Sunita P, Pattanayak SP. Molecular insights into the crosstalk between immune inflammation nexus and SARS-CoV-2 virus. Curr Microbiol 2021; 78: 3813–3828.

8   WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group, Sterne JAC, Murthy S, et al. Association between administration of systemic corticosteroids and mortality among critically ill patients with COVID-19. JAMA 2020; 324: 1330–1341.

9   REMAP-CAP Investigators, Gordon AC, Mouncey PR, et al. Interleukin-6 receptor antagonists in critically ill patients with Covid-19. N Engl J Med 2021; 384: 1491–1502.

10   San-Juan R, Fernández-Ruiz M, López-Medrano F, et al. Analysis of the factors predicting clinical response to tocilizumab therapy in patients with severe COVID-19. Int J Infect Dis 2022; 117: 56–64.

11   Fisher BA, Veenith T, Slade D, et al. 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. Lancet Respir Med 2022; 10: 255–266.

12   CORIMUNO-19 Collaborative group. Effect of anakinra versus usual care in adults in hospital with COVID-19 and mild-to-moderate pneumonia (CORIMUNO-ANA-1): a randomised controlled trial. Lancet Respir Med 2021; 9: 295–304.

13   Monk P, Wilkinson T, Brookes J, et al. SPRINTER: a randomized, double-blind, placebo-controlled, phase 3 trial to determine the efficacy and safety of inhaled interferon beta-1a (SNG001) for the treatment of patients hospitalized due to COVID-19 (NCT04732949). Am J Respir Crit Care Med 2022; 205: A5768.

14   Trapnell BC. Granulocyte macrophage-colony stimulating factor augmentation therapy in sepsis: is there a role? Am J Respir Crit Care Med 2002; 166: 129–130.

15   Pinder EM, Rostron AJ, Hellyer TP, et al. Randomised controlled trial of GM-CSF in critically ill patients with impaired neutrophil phagocytosis. Thorax 2018; 73: 918–925.

16   Meisel C, Schefold JC, Pschowski R, et al. Granulocyte-macrophage colony-stimulating factor to reverse sepsis-associated immunosuppression: a double-blind, randomized, placebo-controlled multicenter trial. Am J Respir Crit Care Med 2009; 180: 640–648.

17   Matute-Bello G, Liles WC, Radella F, et al. Modulation of neutrophil apoptosis by granulocyte colony-stimulating factor and granulocyte/macrophage colony-stimulating factor during the course of acute respiratory distress syndrome. Crit Care Med 2000; 28: 1–7.

18   Huang H, Wang S, Jiang T, et al. High levels of circulating GM-CSF+CD4+ T cells are predictive of poor outcomes in sepsis patients: a prospective cohort study. Cell Mol Immunol 2019; 16: 602–610.

19   Zhou Y, Fu B, Zheng X, *et al.* Pathogenic T-cells and inflammatory monocytes incite inflammatory storms in severe COVID-19 patients. *Natl Sci Rev* 2020; 7: 998–1002.

20   Lang FM, Lee KMC, Teijaro JR, *et al.* GM-CSF-based treatments in COVID-19: reconciling opposing therapeutic approaches. *Nat Rev Immunol* 2020; 20: 507–514.

21   Kasper S, Kindler T, Sonnenschein S, *et al.* Cross-inhibition of interferon-induced signals by GM-CSF through a block in Stat1 activation. *J Interferon Cytokine Res* 2007; 27: 947–960.

22   Rollwagen FM, Davis TA, Li YY, *et al.* Orally administered IL-6 induces elevated intestinal GM-CSF gene expression and splenic CFU-GM. *Cytokine* 2004; 27: 107–112.

23   Adkins KK, Levan TD, Miesfeld RL, *et al.* Glucocorticoid regulation of GM-CSF: evidence for transcriptional mechanisms in airway epithelial cells. *Am J Physiol* 1998; 275: L372–L378.

24   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: 822–830.

25   Bosteels C, Maes B, Damme KV, *et al.* Sargramostim to treat patients with acute hypoxic respiratory failure due to COVID-19 (SARPAC): a structured summary of a study protocol for a randomised controlled trial. *Trials* 2020; 21: 491.

26   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 Rheum* 2021; 3: e410–e418.

27   Temesgen Z, Burger CD, Baker J, *et al.* Lenzilumab in hospitalised patients with COVID-19 pneumonia (LIVE-AIR): a phase 3, randomised, placebo-controlled trial. *Lancet Respir Med* 2021; 10: 237–246.

28   Criner GJ, Lang FM, Gottlieb RL, *et al.* Anti-granulocyte-macrophage colony-stimulating factor monoclonal antibody gimsilumab for COVID-19 pneumonia: a randomized, double-blind, placebo-controlled trial. *Am J Respir Crit Care Med* 2022; 205: 1290–1299.

29   Patel J, Bass D, Beishuizen A, *et al.* A randomised trial of anti-GM-CSF otilimab in severe COVID-19 pneumonia (OSCAR). *Eur Respir J* 2023; 61: 2101870.

30   Knight SR, Ho A, Pius R, *et al.* Risk stratification of patients admitted to hospital with Covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ* 2020; 370: m3339.

31   Bartleson JM, Radenkovic D, Covarrubias AJ, *et al.* SARS-CoV-2, COVID-19 and the ageing immune system. *Nat Aging* 2021; 1: 769–782.

32   Temesgen Z, Kelley CF, Cerasoli F, *et al.* C reactive protein utilisation, a biomarker for early COVID-19 treatment, improves lenzilumab efficacy: results from the randomised phase 3 "LIVE-AIR" trial. *Thorax* 2022; in press [https://doi.org/10.1136/thoraxjnl-2022-218744].

33   Walsh M, Srinathan SK, McAuley DF, *et al.* The statistical significance of randomized controlled trial results is frequently fragile: a case for a fragility index. *J Clin Epidemiol* 2014; 67: 622–628.

34   Potter GE. Dismantling the fragility index: a demonstration of statistical reasoning. *Stat Med* 2020; 39: 3720–3731.

35   Itaya T, Isobe Y, Suzuki S, *et al.* The fragility of statistically significant results in randomized clinical trials for COVID-19. *JAMA Netw Open* 2022; 5: e222973.

36   Kiniksa Pharmaceuticals. Kiniksa Announces Results from Phase 3 Trial of Mavrilimumab in COVID-19-Related ARDS. 2021. Date last accessed: 27 November 2022. https://investors.kiniksa.com/news-releases/news-release-details/kiniksa-announces-results-phase-3-trial-mavrilimumab-covid-19

37   Humanigen Inc. Humanigen Receives Preliminary Topline Data From NIH/NIAID Study of Lenzilumab in ACTIV-5/BET-B. 2022. Date last accessed: 27 November 2022. https://ir.humanigen.com/English/news/news-details/2022/Humanigen-Receives-Preliminary-Topline-Data-From-NIHNIAID-Study-of-Lenzilumab-in-ACTIV-5BET-B/default.aspx

38   Jones CW, Adams AC, Murphy E, *et al.* Delays in reporting and publishing trial results during pandemics: cross sectional analysis of 2009 H1N1, 2014 Ebola, and 2016 Zika clinical trials. *BMC Med Res Methodol* 2021; 21: 120.