Predicting outcomes of COVID-19 from admission biomarkers: a prospective UK cohort study

David T Arnold, Marie Attwood, Shaney Barratt, Anna Morley, Karen T Elvers, Jorgen McKernon, Charmaine Donald, Adrian Oates, Alan Noel, Alasdair MacGowan, Nick A Maskell, Fergus W Hamilton

ABSTRACT

Introduction COVID-19 has an unpredictable clinical course, so prognostic biomarkers would be invaluable when triaging patients on admission to hospital. Many biomarkers have been suggested using large observational datasets but sample timing is crucial to ensure prognostic relevance. The DISCOVER study prospectively recruited patients with COVID-19 admitted to a UK hospital and analysed a panel of putative prognostic biomarkers on the admission blood sample to identify markers of poor outcome.

Methods Consecutive patients admitted to hospital with proven or clinicoradiological suspected COVID-19 were consented. Admission bloods were extracted from the clinical laboratory. A panel of biomarkers (interleukin-6 (IL-6), soluble urokinase plasminogen activator receptor (suPAR), Krebs von den Lungen 6, troponin, ferritin, lactate dehydrogenase, B-type natriuretic peptide, procalcitonin) were performed in addition to routinely performed markers (C reactive protein (CRP), neutrophils, lymphocytes, neutrophil to lymphocyte ratio). Age, National Early Warning Score (NEWS2), CURB-65 and radiographic severity score on initial chest radiograph were included as comparators. All biomarkers were tested in logistic regression against a composite outcome of non-invasive ventilation, intensive care admission or death, with area under the curve (AUC) (figures calculated).

Results 187 patients had 28-day outcomes at the time of analysis. CRP (AUC: 0.69, 95% CI: 0.59 to 0.78), lymphocyte count (AUC: 0.62, 95% CI: 0.53 to 0.72) and other routine markers did not predict the primary outcome. IL-6 (AUC: 0.77, 0.65 to 0.88) and suPAR (AUC: 0.81, 0.72 to 0.88) showed some promise, but simple clinical features alone such as NEWS2 score (AUC: 0.70, 0.60 to 0.79) or age (AUC: 0.70, 0.62 to 0.77) performed nearly as well.

Discussion Admission blood biomarkers have only moderate predictive value for predicting COVID-19 outcomes, while simple clinical features such as age and NEWS2 score outperform many biomarkers. IL-6 and suPAR had the best performance, and further studies should focus on the additive value of these biomarkers to routine care.

INTRODUCTION

COVID-19 causes a wide spectrum of disease, from asymptomatic to severe respiratory failure. The majority of patients who present to hospital will recover but some develop rapidly progressive respiratory compromise requiring ventilatory support. Biomarkers that might predict this deterioration would be invaluable when triaging patients on hospital admission to inform who can be safely discharged versus those who need careful respiratory monitoring.

The rapidly emerging literature base around biomarkers in COVID-19 has mainly extracted data from electronic health records of large cohort studies. Fewer have prospectively recruited and followed up patients. Those that have tested novel biomarkers have often done so at later time points in patients limiting extrapolation to hospital admission.

METHODS

This study is reported in line with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for observational studies (see online supplemental material).
Study design
This was a prospective cohort study aimed to assess prognostic clinical and blood biomarkers for COVID-19 disease based on the earliest available clinical and biochemical information.

Outcomes
The primary outcome of prediction was a composite outcome of ITU admission, NIV with continuous positive airway pressure (CPAP) or bi-level positive airway pressure (BIPAP) outside the ITU or death (defined below as ‘severe disease’).

The secondary outcome was a composite outcome of ITU admission and death.

Patient recruitment
All patients were recruited via the DISCOVER study, a single-centre observational study at North Bristol NHS Trust recruiting patients with COVID-19, from 30 March 2020 until 29 June 2020. Patients were recruited on the basis of a positive PCR result for SARS-CoV-2, using the established PHE assay in use at the time or a clinicoradiological diagnosis of COVID-19 disease. During the pandemic, community testing became widely available, although the results were not available to hospital staff. As such, later patients were often recruited on the basis of a history of positive testing in the community. The only exclusion criteria was an inability to consent. For patients in ITU, family members were able to consent on behalf of them if too unwell to consent as a personel consultee.

Clinical information
Clinical information was recorded on a REDCap (Vanderbilt University) database, by the consenting nurse or physician. Routine demographics were recorded and presence of important comorbidities. Comorbidities were defined either by their recording in the admission notes/hospital record (for hypertension, heart disease and chronic lung disease), the presence of a positive serological result (for HIV) or for requirement for dialysis or an estimated glomerular filtration rate (eGFR) of <30 mL/min (for chronic kidney disease). Ethnicity was also recorded.

The earliest admission National Early Warning Score (NEWS2) was extracted from the clinical record. This is a numeric score (from 0 to 20), reflecting the degree of physiological dysfunction. Higher numbers indicate more severe physiological dysfunction. Routine biochemistry and haematology results were extracted from the clinical record, using the earliest available figure. Outcome data recorded was in line with RECOVERY, the national RCT of therapeutic interventions for COVID-19 (https://www.recoverytrial.net/). This included 28-day mortality, requirement for ITU, ventilation, renal replacement therapy, and inotropes.

NIV was defined by the receipt of NIV by CPAP, BIPAP or High Flow Nasal Oxygen (HFNO) any time during the hospital stay. The radiological severity score was calculated using the method described by Wong et al. A score of 0–4 was assigned to each lung depending on the extent of involvement by consolidation or ground glass opacities: 0 = no involvement, 1 = <25%, 2 =25%–49%, 3 =50–75%, 4 = >75% involvement. The scores for each lung were summed to produce a final severity score ranging from 0 to 8. Radiographs were scored by a respiratory and infectious diseases physician.

Biomarkers and samples
For conventional biomarkers of infection (namely C reactive protein (CRP), components of the full blood count and routine renal function), data were prospectively recorded. The first result from that admission was taken or, for established inpatients, the result on the day of the diagnosis. These results were available to the clinician and therefore could impact on prognosis (ie, clinicians were not blinded to these results)

For other potential predictive markers (lactate dehydrogenase (LDH), procalcitonin (PCT), interleukin-6 (IL-6), Krebs von den Lungen 6 (KL-6), ferritin, troponin, B-type natriuretic peptide (NT-pro-BNP), soluble urokinase plasminogen activator receptor (suPAR)), analysis was performed on frozen samples in batch analysis as described below. Laboratory staff were unaware of clinical outcome and were therefore functionally blinded.

The earliest initial sample was extracted from the blood sciences laboratory after routine testing had been performed; in admitted cases this was the initial sample taken in the emergency department (ED); in hospitalised cases, this was the sample from the day of diagnosis. suPAR analysis was performed on the suPARnostic ELISA platform; KL-6 and IL-6 were performed on the Fujirebio Lumipulse. CRP, ferritin, LDH, troponin, NT-pro-BNP and PCT were performed on the Roche Diagnostics cobas platform. The specific analytic platforms were spectrophotometric: c701 (CRP) and c301 (LDH), immunoassay: c501 (ferritin, troponin, NT-pro-BNP, PCT). The full blood count was performed on the Sysmex XN (Sysmex Diagnostics).

### Table 1 Demographics of the study cohort

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Survivor, N=148</th>
<th>Non-survivor/ITU/NIV, N=39</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (18+), years</td>
<td>55 (44–71)</td>
<td>66 (60–76)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td>0.7</td>
</tr>
<tr>
<td>Male</td>
<td>79 (53)</td>
<td>22 (56)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>69 (47)</td>
<td>17 (44)</td>
<td></td>
</tr>
<tr>
<td>PCR proven disease</td>
<td></td>
<td></td>
<td>0.6</td>
</tr>
<tr>
<td>Proven</td>
<td>111 (75)</td>
<td>31 (79)</td>
<td></td>
</tr>
<tr>
<td>Suspected</td>
<td>37 (25)</td>
<td>8 (21)</td>
<td></td>
</tr>
<tr>
<td>Status at time of consent</td>
<td></td>
<td></td>
<td>0.010</td>
</tr>
<tr>
<td>Inpatient</td>
<td>119 (80)</td>
<td>38 (97)</td>
<td></td>
</tr>
<tr>
<td>Outpatient</td>
<td>29 (20)</td>
<td>1 (2.6)</td>
<td></td>
</tr>
<tr>
<td>Diabetes status</td>
<td></td>
<td></td>
<td>0.3</td>
</tr>
<tr>
<td>No</td>
<td>115 (78)</td>
<td>34 (87)</td>
<td></td>
</tr>
<tr>
<td>T1DM</td>
<td>3 (2.0)</td>
<td>1 (2.6)</td>
<td></td>
</tr>
<tr>
<td>T2DM</td>
<td>30 (20)</td>
<td>4 (10)</td>
<td></td>
</tr>
<tr>
<td>Heart disease</td>
<td>34 (23)</td>
<td>12 (31)</td>
<td>0.3</td>
</tr>
<tr>
<td>Chronic lung disease</td>
<td>26 (18)</td>
<td>20 (51)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Severe liver disease</td>
<td>3 (2.0)</td>
<td>2 (5.1)</td>
<td>0.3</td>
</tr>
<tr>
<td>Severe kidney impairment (eGFR &lt;30 or dialysis)</td>
<td>16 (11)</td>
<td>6 (15)</td>
<td>0.4</td>
</tr>
<tr>
<td>Hypertension</td>
<td>37 (25)</td>
<td>15 (38)</td>
<td>0.10</td>
</tr>
<tr>
<td>HIV status</td>
<td>1 (0.7)</td>
<td>1 (2.6)</td>
<td>0.4</td>
</tr>
<tr>
<td>Non-white ethnicity</td>
<td>20 (16)</td>
<td>5 (15)</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Median (IQR); n (%). Wilcoxon rank sum test; Pearson’s χ² test; Fisher’s exact test.

eGFR, estimated glomerular filtration rate; ITU, Intensive Care admission; NIV, non-invasive ventilation; T1DM, type 1 diabetes mellitus; T2DM, type 2 diabetes mellitus.
As the volume of blood available varied and was small, the total number of assays varied slightly; as for some participants only enough blood was available to perform one or two assays, or there was technical failure and an inability to repeat due to lack of sample.

Statistical approach
In this study, we aimed to identify whether any individual biomarker (lymphocyte count, neutrophil count, neutrophil:lymphocyte ratio (NLR), CRP, IL-6, KL-6, suPAR, NT-pro-BNP, LDH, PCT, troponin T, ferritin) had prognostic significance for the primary outcome as an individual marker, when used on the initial blood sample taken.

There was a deliberate focus on relatively simplistic (logistic regression) models: given the sample size, complex models would be at risk of overfitting, and the main target was to identify the additional value of a biomarker, all of which are reported on a linear scale. Complete case analysis was performed for each individual biomarker.

Receiver operating characteristic curve (ROC) curves were generated, and area under the curve (AUC) figures were calculated alongside sensitivity and specificity for each biomarker, at Youden’s index. CIs were generated around the AUC by bootstrapping.

All analysis was performed in R (V4.0.0), using the packages ‘tidyverse’, ‘broom’, ‘tidymodels’ and ‘pROC’. Analytic code is available at: https://github.com/gushamilton/discover_prediction/

Patient and public involvement
Given the difficulty of performing PPI during a pandemic, patients were not involved in the design, or conduct, or reporting, or dissemination plans of the DISCOVER study at the time this manuscript was submitted.

RESULTS
One hundred eighty-seven participants had reached 28 days postdiagnosis and were included in the analysis (prespecified timing of the primary outcome). Table 1 shows the demographics of the cohort. One hundred and one patients (54%) were male, and the median age was 58 (IQR: 46–73) years; 76% of the cohort had positive PCR results for SARS-CoV-2, with the remaining clinically suspected.

Comorbidities were relatively common within the cohort with hypertension and diabetes being the most predominant. Patients with severe disease (death, ITU admission or NIV) were generally older and more comorbid than non-severe patients. Thirty-nine patients had the primary outcome. Fifteen patients went to the ITU, of which 4 died and 11 survived. Seven patients required NIV outside the ITU, 4 of whom died (see figure 1).

There was a significant variation in physiological state (eg, NEWS2), with a median NEWS2 score of 4 (IQR 2–6), with the highest NEWS2 score recorded was 13. Patients that had severe disease had higher NEWS2 scores. There was also wide variation in functional status, with many patients having some degree of frailty, with sicker patients more likely to die or require enhanced care. Escalation status was recorded for most patients, with patients who died more likely to have limitations on ceiling of care, recorded in table 2.

Admission blood tests are recorded in table 3. There was no significant difference for many blood tests, with lymphocyte count being slightly lower in those with severe disease.

---

**Figure 1** Study flowchart. ITU, intensive care admission; NIV, required non-invasive ventilation.
and NLR being slightly higher. Renal dysfunction was more common in the severe disease.

**Logistic regression and AUC calculation**

Table 4 shows biomarker performance to predict severe COVID-19 outcomes. Blood was not available for all participants for all tests, with the number included in each model listed. Most biomarkers had modest predictive value, with suPAR (AUC 0.81) and IL-6 (AUC 0.77) having the best performance. Many biomarkers had weak performance (CRP, neutrophils, lymphocytes, KL-6), with AUC figures between 0.5 and 0.65. ROC plots are available in the online supplemental material for all biomarkers. Of note, both age and NEWS2 score performed as well as most biomarkers.

Results of routine biomarkers (CRP, neutrophils, lymphocytes and NLR) are shown in figure 2. Best performing novel biomarkers (IL-6, suPAR, LDH and PCT) box plots are shown in figure 3 (PCT values have been logged prior to plotting to aid visualisation as they had very wide ranges).

**Secondary outcomes**

For the composite outcome of ITU and mortality, results are reported in the online supplemental material. Due to a small number of events (18 deaths, 15 ITU admissions), CIs were wider, but results were similar although perhaps slightly worse. Again, suPAR and IL-6 were the best performing biomarkers (AUC 0.81 for suPAR, AUC 0.70 for IL-6), with most other biomarkers having an AUC between 0.5 and 0.65.

**DISCUSSION**

COVID-19 remains a clinical challenge. The majority of patients who present to hospital will recover but some develop rapidly progressive respiratory failure. Biomarkers that might predict this deterioration would be invaluable when triaging patients on hospital admission to inform who can be safely discharged versus those who might need intensive care support in the near future. This paper presents a prospectively recruited UK cohort of patients with COVID-19 with targeted biomarker sampling at presentation.

**Previous literature**

A large observational study recruiting from the majority of NHS hospitals (ISARIC) estimates that of the 34,608 patients with COVID-19 outcomes.
outcome data, 7374 (17%) required admission to ITU and 11 659 (33.7%) died within follow-up.12 There have been numerous studies and meta-analyses on potential predictors of respiratory decline in COVID-19 pneumonia with the development of a ‘cytokine storm’ in specific patients cited as a major determinant.13–14 As a result, there has been a focus on biomarkers that rise in other similar conditions such as ferritin, LDH, IL-6 and suPAR.15 A rise in cardiovascular events and coagulopathies has also been seen in patients with COVID-19 pneumonia so biomarkers such as troponin, NT-pro-BNPs fibrinogen and D-dimer have been studied.16–19 Given that mortality from COVID-19 increases with age and frailty, biomarkers of frailty have been shown to be related to worse outcomes (albumin, eGFR).2,12,20 Finally, given this is a respiratory infection, blood-based biomarkers that prognosticate bacterial pneumonia have been included in treatment guidelines and even entry criteria for clinical trials, including lymphopenia, neutrophilia, PCT and C reactive protein.6,21

Study findings

In this study, conventionally performed blood biomarkers, in isolation, did not predict outcome when performed on admission. Neutrophilia and C reactive protein had AUC close to 0.5, with lymphopenia and NLR having only marginally better discriminative value. Cardiac markers were on average slightly higher in patients with worse clinical outcomes but should not be relied on to make treatment decisions at baseline. Literature suggests that they may have more utility when measured serially, especially in the very unwell patient.22–24

Markers of immune activation had more promise at the admission timepoint with IL-6 and suPAR the best-performing within this cohort. It has been hypothesised that an exaggerated immune response or ‘cytokine storm’ plays a significant role in COVID-19 and several immunomodulatory therapies are being trialled. IL-6 is a proinflammatory mediator for the induction of the acute phase response and shown to be a predictive marker of deterioration in other serious lung pathologies. Specific to COVID-19, Han and colleagues22 assessed the cytokine profile of 102 patients, admitted to a single hospital in Wuhan, by disease severity on admission. From a panel of six cytokines (TNF-α, IFN-γ, IL-2, IL-4, IL-6 and IL-10), serum IL-6 was predictive of disease severity at the time with an AUC of 0.84, although longer term outcomes were not reported. Another prospective study of 89 patients admitted to a German hospital demonstrated that admission IL-6 was superior to other blood-based biomarkers at predicting the need for mechanical ventilation (using a cut-off of 35 pg/mL).22 Another prospective study of 89 patients admitted to a German hospital demonstrated that admission IL-6 was superior to other blood-based biomarkers at predicting the need for mechanical ventilation (using a cut-off of 35 pg/mL).23 In our study, the admission IL-6 (at a cut-off of 74.9 pg/mL) had a sensitivity of 0.70 and specificity of 0.84 (AUC 0.77). Randomised trial evidence is emerging showing that treatment with the IL-6 receptor antagonists, tocilizumab and sarilumab, improves outcomes, including survival in severely unwell patients with COVID-19.24

suPAR, a marker of immune activation, has been shown to predict deterioration in several infectious and inflammatory disorders. It forms part of the fibrinolysis cascade and increased levels have been shown to predispose to clotting abnormality and renal dysfunction, both of which are important drivers of morbidity and mortality in COVID-19. A prospective study of 57 patients presenting with COVID-19 demonstrated that admission levels of suPAR were significantly greater among patients who eventually required ventilatory support.15 A cut-off of 6 ng/mL had a sensitivity of 86% and specificity of 92%. A prospective study of 57 patients presenting with COVID-19 demonstrated that admission levels of suPAR were significantly greater among patients who eventually required ventilatory support.15 A cut-off of 6 ng/mL had a sensitivity of 86% and specificity of 92%. In our study, it performed reasonably compared with other blood biomarkers but at a cut-off of 5.2 ng/dL only had a sensitivity of 82% and specificity of 65%, with lower cutoffs increasing sensitivity at relatively little cost.

Clinical markers

The focus of this study was the additional role of blood biomarkers when initially assessing patients presenting with COVID-19. The DISCOVER cohort also collected routinely recorded clinical data including demographics, baseline observations and initial radiography. It is notable that this easily accessible information outperformed many of the blood-based biomarkers tested.

The NEWS2 is used throughout the UK. There is currently limited evidence supporting its use in COVID-19.24 In China, Liao and colleagues25 created a non-validated COVID-19-specific NEWS2 score (where age over 65 added 3 points). We could not find any prospective UK-based studies accessing the utility of NEWS2 in COVID-19 admission.25 In the DISCOVER cohort, NEWS2 score had an AUC of 0.70. The CURB-65 score is a severity scoring system for community-acquired pneumonia and is often used to make decisions around need for hospital admission. Although previous retrospective literature from our centre has shown that CURB-65 scores are higher in patients with a worse outcome from COVID-19,26 this prospective study has demonstrated that it cannot be relied on to make treatment decisions on an individual patient level.

Strengths

The DISCOVER cohort was prospectively designed and analysed the earliest ED blood sample, which is a key strength. Clinical metadata was robustly recorded, and novel biomarkers were performed in batch assays, maximising replicability and reducing bias as results could not influence patient care. Unlike other datasets (eg, Electronic Health Record extractions), we included patients who had clinically suspected COVID-19, as the currently available assays still have limited sensitivity. Technicians were employed to extract blood out of autoanalyser machines to get the earliest possible sample, enabling recruitment of patients after admission or even
as outpatients. Missing data around comorbidities was rare, as patients were prospectively recruited by research nurses.

Limitations
The major limitation of this study is the limited sample size, leading to imprecise estimates of biomarker performance. A second limitation is the composite outcome of NIV, ITU admission and death, which has been used in major interventional COVID-19 trials. Although this is clinically useful and may aid differentiation of those who require specialist care, the provision and use of NIV are more clinician and hospital dependent and may be harder to extrapolate from. However, our secondary analysis of ITU admission and death was largely similar, supporting this approach. Finally, although around 80% of patients had blood available, a proportion did not have excess sample available for additional testing, so only the routinely performed biomarkers (CRP, neutrophils, lymphocytes, NLR) are recorded for those patients. It is also important to note that these results were available to clinicians and therefore may have biased the outcome (ie, patients with high CRP may have received more aggressive therapy). At the time of recruitment to this study, there were however no approved therapies for COVID-19 and it is therefore unlikely this strongly biased the results.

CONCLUSION
To our knowledge, this is the largest recruited UK cohort of consecutive patients presenting with COVID-19. Blood biomarkers, when performed on admission bloods, had only moderate predictive value for COVID-19 similar to age and routine clinical scores (eg, NEWS2 score). IL-6 and suPAR had the best performance, and further, large prospective studies should validate the additional value of these biomarkers to routinely collected clinical information.

REFERENCES