Non-generalizability of biomarkers for mortality in SARS-CoV-2: a meta-analyses series

ME Rahman Shuvo*1 MBBS, Max Schwiening*2 BA, Felipe Soares4,5 PhD, Oliver Feng3 PhD,
Susana Abreu2 PhD, Niki Veale1 MBBS, William Thomas1 MBBS, AA Roger Thompson4 PhD,
Richard J Samworth3 PhD, Nicholas W Morrell1,2 PhD, Stefan J Marciniak1,2 PhD, Elaine Soon1,2
PhD (0000-0002-5744-5014)

1Department of Respiratory Medicine, Level 5, Box 157, Cambridge University Hospitals NHS
Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK - MER Shuvo (Specialist trainee), Niki
Veale (Specialist trainee), W Thomas (Consultant), NW Morrell (Professor), SJ Marciniak
(Professor), E Soon (Clinician Scientist).

2Cambridge Institute for Medical Research, University of Cambridge, Keith Peters Building, Hills
Rd, Cambridge CB2 0XY, UK - M Schwiening (PhD candidate), S Abreu (Postdoctoral researcher),
SJ Marciniak (Professor), E Soon (Clinician Scientist).

3Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3
0WB – O Feng (Postdoctoral researcher), RJ Samworth (Professor).

4Department of Infection, Immunity and Cardiovascular Disease, The University of Sheffield,
Beech Hill Road, Sheffield S10 2RX – F Soares (PhD candidate), AAR Thompson (Clinician
Scientist).

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Correspondence to:

Elaine Soon

Cambridge Institute for Medical Research,
Keith Peters Building, Cambridge Biomedical Campus, Hills Rd, Cambridge CB2 0XY, UK.

Fax number: ++44 (1223) 336846

Telephone number: ++44 (1223) 336788

E-mail: eels2@cam.ac.uk, e_soon@hotmail.co.uk

Author contributions

MERS, FS, WT, and NV collected and helped analyze data. MS, FS and SA contributed to the creation of online resources. OF and RJS performed statistical analyses. RT, NWM and SJM contributed to the planning, organization, and funding of the study. ES conceptualized and organized the study, performed data analyses, funded the study, and takes responsibility for the integrity of the work as a whole. All authors contributed to the writing and critical revision
of the manuscript. The corresponding author attests that all listed authors meet authorship
criteria and that no others meeting the criteria have been omitted.

Disclosures

All authors have completed the ICMJE uniform disclosure form at
http://www.icmje.org/disclosure-of-interest/ and declare: no support from any commercial
organisation for the submitted work; no financial relationships with any organisations that
might have an interest in the submitted work in the previous three years; no other relationships
or activities that could appear to have influenced the submitted work.

Ethics approval

Internal ethical approval was obtained from the Integrated Research Application System
(reference 281880) for analysis of the Cambridge (UK) data. All the other data has been
published and is in the public domain.

Transparency

E Soon affirms that the manuscript is an honest, accurate, and transparent account of the study
being reported and that no important aspects of the study have been omitted.

Funding
ES and MS are supported by the UK Medical Research Council (MR/R008051/1); the British Medical Association (the Josephine Lansdell Award); and the Association of Physicians of Great Britain and Ireland (Young Investigator Award to ES); the Wellcome Trust ISSF and the Cambridge BHF Centre of Research Excellence (RE/18/1/34212). MES and WT are full-time NHS physicians who have volunteered their time for this work. FS received in-kind funding by the AWS Diagnostic Development Initiative and Google TPU Research Cloud. NV is supported by a BLF-Papworth Fellowship from the British Lung Foundation and the Victor Dahdaleh Foundation (VPDCF17-18). AART is supported by a British Heart Foundation Intermediate Clinical Fellowship (FS/18/13/33281). OF is funded by the StatScale programme (EP/N031938/1). RJS is supported by Engineering and Physical Sciences Research Council grants EP/P031447/1 and EP/N031938/1, as well as ERC Advanced Grant 101019498. SA and SJM are funded by the British Lung Foundation (VPDCF17-18), the Medical Research Council, UK (MR/V028669/1), the NIHR Cambridge Biomedical Campus (BRC-1215-20014) and the Royal Papworth NHS Trust.

NWM is supported by the British Heart Foundation (SP/12/12/29836), the Cambridge BHF Centre of Research Excellence (RE/18/1/34212), the UK Medical Research Council (MR/K020919/1), the Dinosaur Trust, BHF Programme grants to NWM (RG/13/4/30107), and the NIHR Cambridge Biomedical Research Centre.

**Total word count:** 2,235

**Abstract word count:** 335
Summary box

Section 1: What is already known on this topic

Biomarkers such as CRP, D-dimer, and interleukin-6 have been proven to have prognostic value in SARS-CoV-2. However, prognostic scores using these as building blocks perform unevenly in different locations.

Section 2: What this study adds

Commonly used biomarkers for SARS-CoV-2 have different efficacy in different parts of the world. For example, admission CRP and interleukin-6 levels are good prognostic markers for mortality in Asian countries but only average in Europe and North America. Prognostic markers and scores cannot be ‘transplanted’ from one region to another. This has implications not just for SARS-CoV-2 but also for scores in other conditions.
ABSTRACT

Objectives: Sophisticated scores have been proposed for prognostication of mortality due to SARS-CoV-2 but perform inconsistently. We conducted these meta-analyses to uncover why and to pragmatically seek a single dependable biomarker for mortality.

Design: We searched the PubMed database for the keywords ‘SARS-CoV-2’ with ‘biomarker name’ and ‘mortality’. All studies published from 01st December 2019 to 30th June 2021 were surveyed. To aggregate the data, the meta library in R was used to report overall mean values and 95% confidence intervals. We fitted a random effects model to obtain pooled AUCs and associated 95% confidence intervals for the European/North American, Asian, and overall datasets.

Setting and Participants: Data was collected from 131 studies on SARS-CoV-2 PCR-positive general hospital adult admissions (n=76,169 patients in total).

Main Outcome Measures: We planned a comparison of pooled area under curves (AUCs) from Receiver Operator Characteristic curves plotted for admission D-dimer, CRP, urea, troponin and interleukin-6 levels.

Main Results: Biomarker effectiveness varies significantly in different regions of the world. Admission CRP levels are a good prognostic marker for mortality due to SARS-CoV-2 in Asian countries, with a pooled area under curve (AUC) of 0.83 (95% CI 0.80-0.85), but only an average predictor of mortality in Europe/North America, with a pooled AUC of 0.67 (95% CI 0.63-0.71, P<0.0001). We observed the same pattern for D-dimer and IL-6. This variability explains why the proposed prognostic scores did not perform evenly. Notably, urea and troponin had pooled
AUCs $\geq 0.78$ regardless of location, implying that end-organ damage at presentation is a key prognostic factor. These differences might be due to age, genetic backgrounds, or different modes of death (younger patients in Asia dying of cytokine storm while older patients die of multi-organ failure).

**Conclusions:** Biomarker effectiveness for prognosticating SARS-CoV-2 mortality varies significantly by geographical location. We propose that biomarkers and by extension prognostic scores need to be tailored for specific populations. This also implies that validation of commonly used prognostic scores for other conditions should occur before they are used in different populations.
Introduction

SARS-CoV-2 is a novel beta coronavirus of zoonotic origin that emerged at the end of 2019 in Wuhan, China. SARS-CoV-2 differs from previous viral threats in showing marked transmissibility during the asymptomatic/very early symptomatic stage and person-to-person transmission by both airborne and fomite routes. At the beginning of the pandemic, there was no previous immunity, no known effective antiviral treatment, and no vaccine, resulting in a global death toll of just over six million (https://covid19.who.int/).

Due to the overwhelming number of cases and the significant morbidity and mortality associated with SARS-CoV-2, reliable prognostic scores are critically important to maximize survivorship and optimize the use of limited resources. Sophisticated scoring systems have been proposed but have not performed consistently. For example, El-Solh tested 4 prognostic models constructed to predict in-hospital mortality for SARS-CoV-2 patients; proposed by Chen et al, Shang et al, Yu et al, and Wang et al. All models had been peer reviewed and were based on a cohort size of ≥100. All the models examined had validation area under curves (AUCs) which were significantly worse than the area under curves of their derivation cohorts. For example, the AUC of the validation cohort using the model proposed by Chen et al was at best 0.69 (95% confidence interval [CI] 0.66-0.72) compared to the derivation AUC, which was 0.91 (95% CI 0.85-0.97). A similar pattern was noted in the other 3 models.
Gupta tested 20 candidate prognostic models using data derived from 411 consecutively admitted adults with a PCR-confirmed diagnosis of SARS-CoV-2 in a major London hospital. Five of these models were pre-existing point-based scores not specific for Covid19 (MEWS, REMS, qSOFA, CURB65 and NEWS2) and the remainder of which were a combination of point-based scores and logistic regression models specifically derived from SARS-CoV-2 patients. None of these methods overlapped with those previously tested by El-Solh and 9 of the 15 Covid-specific models had been developed in China. The most discriminating univariable predictor for in-hospital mortality was age (AUC 0.76 [95% CI 0.71-0.81]) and for in-hospital deterioration was oxygen saturation on room air (AUC 0.76 [95% CI 0.71-0.81]). More importantly, none of the models tested performed consistently better than these univariable predictors.

These inconsistencies are an ongoing issue. Bradley concluded that the overall prognostic performance of established clinical scores (CURB-65, NEWS2 and qSOFA) was generally poor with reference to SARS-CoV-2 while Fan concluded the opposite. To illustrate the AUC for CURB65 prognostic score was 0.85 (Fan), 0.75 (Bradley) and 0.698 (Kodama). This begs the question - why are prognostic scores performing so inconsistently even when tested against cohorts who are similar clinically? We ran these meta-analyses to uncover possible reasons for these inconsistencies. As a secondary goal, we also sought an easily measurable, dependable single-parameter biomarker to predict mortality in swab-positive SARS-CoV-2 patients; especially as there is not always time or resources available to calculate a full prognostic model.
Methods

We searched the PubMed database for the keywords ‘SARS-CoV-2’ in combination with ‘biomarker name’ and ‘mortality’. The period for the first data tranche was set from the emergence of the SARS-CoV-2 pathogen on the 01\textsuperscript{st} December 2019 to 30\textsuperscript{th} June 2021. Two independent reviewers analyzed studies for relevance. All papers reporting mortality data for hospitalized patients swab-positive for SARS-CoV-2 with a biomarker level at presentation were examined for a Receiver Operating Characteristic (ROC) analysis and a corresponding Area under Curve (AUC). When studies failed to quote the margin of error for AUCs, corresponding authors were contacted and their AUC data included in pooled AUCs only if confidence intervals or standard deviations were forthcoming. All studies are displayed in summary figures for completeness. Internal ethical approval was obtained from the Integrated Research Application System (reference 281880) for analysis of the Cambridge (UK) data. To ensure biomarkers were applicable to acute adult general admissions, we excluded reports of patients already admitted to intensive care or restricted to specific groups (pregnancy, hemodialysis, or transplant patients). Mortality (30-day and/or in-hospital) was used as the endpoint. The following data was collected from the root studies: a) Area under curve and 95\% confidence intervals for the biomarkers examined (admission D-dimer, CRP, IL-6, troponin, urea); b) age of cohort (mean and standard deviation) and number of patients in cohort; c) geographical location of cohort (if a multi-center study, the location of the hospital of the first author was used). Europe/North America and Asia were the sources of most studies and were therefore the focus of subsequent meta-analyses. This process is summarized in a PRISMA flowsheet depicted in Fig.1. A more detailed explanation is available in the online Supplement.
Statistical analysis

To aggregate the data on age and biomarkers from individual studies, the meta library in R was used to report overall mean values and 95% confidence intervals and statistical significance of differences between mean values in the joint European and North American cohort and the Asian cohort. This analysis was based on estimates of standard errors for each study, obtained by assuming values for individual subjects were normally distributed in each study with a study-specific mean. In this way, measures of spread (IQR, SD and range) were converted into estimates of within-study standard deviations. Since the estimates of the study-specific means exhibited high levels of heterogeneity within both categories, a random effects model was fitted as opposed to a fixed effects model in the meta-analysis.

Sensitivity analyses

Sensitivity analyses were performed by serially excluding each study to determine the implications of individual studies on the pooled AUC. No individual study had a significant implication for pooled AUCs for either European/North American or the Asian cohorts (Supplementary Tables 1-5). Note that when fitting a random effects meta-analysis model, the individual study means are assumed to be random, and the between-study heterogeneity ($\tau^2$) needs to be estimated. For the pooled AUC, we used a single estimate of $\tau^2$ based on the overall dataset (both European/North American and the Asian studies) due to small sample
sizes, so removing a study affects this and thereby may shift the confidence intervals very slightly for the other subgroup.

Patient and public involvement

After discussion with members of the public we decided to display all the root data underpinning the meta-analyses on a publicly-available website (https://covid19.cimr.cam.ac.uk/), with links back to the original studies. Study authors have also volunteered that statistical software is expensive and hence inaccessible. Therefore we have written a programme in R which allows for calculation for the AUC of a biomarker which is free to download from the same website (tested by Mr Zubkov). Our intention is that everyone will be able to view the most effective biomarkers for their locale from the website. The website was tested for accessibility and ease of understanding by Ms Natalie Doughty and Mr Chris Davies.

Results

We examined 1,930 articles that were published from the beginning of the SARS-CoV-2 pandemic in 01st December 2019 to 30th June 2021, and selected 131 papers which met our pre-specified selection criteria. This process is summarized in Fig.1 and all reference papers are listed in the References (Meta-analyses) section.

Our meta-analyses have revealed differences in the effectiveness of biomarkers in different regions of the world. These are summarized in Figure 2. For example, admission CRP levels are a
good prognostic marker for mortality in Asian countries, with a pooled AUC (area under curve) of 0.83 (95% CI 0.80-0.85) from 34 studies, but only an average predictor of mortality in Europe and North America, with a pooled AUC of 0.67 (95% CI 0.63-0.71) from 21 studies ($P<0.0001$, Fig. 3A, Table 1). We see the same pattern for admission D-dimer and IL-6 levels – they are good predictors of mortality in Asian countries (pooled AUCs of 0.78 [95% CI 0.76-0.82]) and 0.86 [95% CI 0.81-0.90] respectively) but not in Europe and North America (pooled AUCs of 0.69 [95% CI 0.66-0.72] and 0.70 [95% CI 0.64-0.75] respectively; $P<0.0001$ for both compared to Asian counterparts; Fig.3B and Fig.4A). This explains why the prognostic scores that are being proposed for SARS-CoV-2 do not perform evenly in different countries, as the ‘building blocks’ underpinning these prognostic scores have intrinsically different effectiveness in different populations.

There are two biomarkers that performed well in all cohorts regardless of geographical location. Admission troponin levels had a pooled AUC of 0.81 [95% CI 0.77-0.85] in Asian countries and a pooled AUC of 0.79 [95% CI 0.74-0.83] in European and North American countries (Fig.4B). Similarly, urea levels on admission had a pooled AUC of 0.79 [95% CI 0.70-0.85] in Asian countries and a pooled AUC of 0.78 [95% CI 0.74-0.81] in European and North American countries (Fig.4C). This implies that end-organ damage at the time of presentation is a key prognostic indicator of severity for SARS-CoV-2.
Pooling all the results from Asian, European, and North American studies gives a false impression of overall effectiveness for CRP, D-dimer, and IL-6 (Table 1). As an example, the pooled AUC for CRP for the entire dataset is 0.78 (95% CI 0.74-0.81). When separated into the regional blocks as previously described it becomes obvious that the Asian studies are skewing the results and masking the fact that admission levels of CRP, D-dimer and IL-6 are simply not very effective in predicting mortality in European and North American countries.

Discussion

We demonstrate for the first time that biomarker effectiveness for mortality in SARS-CoV-2 varies significantly by geographical location. This important finding has impact for clinicians using biomarkers and/or prognostic scores derived in other regions to assist the process of decision-making (e.g. whether to admit to intensive care) particularly when ‘waves’ of infection risk overwhelming local health resources.

We propose that biomarkers need to be tailored for specific populations in specific locations. Consistent with our findings, Marino et al\textsuperscript{15} demonstrated that a prognostic score developed in the same country (PREDI-CO, Bartoletti et al\textsuperscript{16}, Italy) had reasonable predictive power (AUC of 0.76, 95% CI 0.58-0.93) while a prognostic score developed in another country (Yan-XGBoost, Yan et al\textsuperscript{17}, China) did not perform satisfactorily (AUC of 0.57, 95% CI 0.37-0.76) when applied to their cohort.
Our observations are likely to apply to other conditions. CURB-65 is well-known and validated as a tool for predicting mortality in community-acquired pneumonia\(^{18}\), developed in the UK, New Zealand and the Netherlands. However it performs less satisfactorily in older populations\(^{19}\). For example, Shirata et al\(^{20}\) demonstrated that CURB-65 had an AUC of 0.672 (95% CI 0.607-0.732) when applied to patients ≥65 years. Since Japan has one of the highest life expectancies in the world it is likely that CURB-65 would not perform as well if applied to a Japanese cohort. Interestingly, CURB-65 also performs relatively poorly when applied to Colombian patients (AUCs of 0.629-0.669 when tested against 3 cohorts)\(^{21}\). Hincapie suggested that this may be due to the factors underpinning a significant difference in community-acquired pneumonia associated mortality (9.5%\(^{18}\) versus 17-32%\(^{21}\)).

It is not possible to know from these descriptive meta-analyses why there are these regional differences in biomarker effectiveness. The differences might be due to cohort age, different modes of death, genetic backgrounds, treatment effects, and/or various combinations of the above. The Asian cohorts were universally younger than the European/North American cohorts in all five parameters we investigated (CRP, D-dimer, troponin, urea, and IL-6; Table 1). It is possible that in Asia younger patients were dying from cytokine storm (hence the marked prognostic value of the ‘inflammatory’ markers such as CRP, D-dimer, and IL-6), while in Europe older people were dying from multi-organ failure.
It is also possible that there has been a ‘training effect’, with the West having had prior warning from the Asian experience. The earlier use of specific anti-inflammatory approaches, in particular steroids and tocilizumab, has most probably blunted the effectiveness of markers such as IL-6 and CRP as predictors of death. The use of social contact-limiting measures (‘lockdowns’) has likely changed the composition of people falling ill and hence seeking hospital admission.

This study has a number of limitations. First, a significant number of studies did not quote 95% confidence intervals (14 of 78 for D-dimer, 18 of 75 for CRP, 3 of 35 for troponin, 1 of 16 for urea, 8 of 38 for IL-6) and we were unable to obtain them despite best efforts to communicate with the authors. These studies are included in Figs.3-4 but are not included in calculation of the pooled AUCs. Second, insufficient numbers of studies were located in other continents to perform an adequate meta-analysis. Finally, the majority of studies in the Asian section were from China (so 34 of 47 studies for CRP were on the Chinese population) and so the result may be representative of the Chinese population rather than of Asian populations in general.

We acknowledge that SARS-CoV-2 is a rapidly evolving pathogen and the rise of different strains and advent of mass vaccination programs will likely change biomarker effectiveness. To track these changes, we have mapped the root studies on the following website

(https://covid19.cimr.cam.ac.uk/) so that all may see which biomarkers perform well in their locale. We propose a free-to-use software program that the healthcare community can use to
check whether their biomarker of choice is effective in their population. We would request that results from this be uploaded so we can periodically update the website. Critically, published information on SARS-CoV-2 lags behind the immediate need for such. We aim to display shared data in real time to inform regional practice and identify trends in biomarker utility following vaccination and viral mutation.
References (body of manuscript):


the 4C Mortality Score. *Bmj* 2020;370:m3339. doi: 10.1136/bmj.m3339 [published Online First: 2020/09/11]


<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Location</th>
<th>Pooled AUC</th>
<th>95% CI</th>
<th>P-value</th>
<th>Mean age</th>
<th>95% CI</th>
<th>P-value</th>
<th>No. of patients</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-6</td>
<td>Asia</td>
<td>0.86</td>
<td>0.81-0.90</td>
<td>&lt;0.00001</td>
<td>58.2</td>
<td>53.6-62.9</td>
<td>0.033</td>
<td>2,993</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Europe/</td>
<td>0.7</td>
<td>0.64-0.75</td>
<td>63.5</td>
<td>61.1-65.9</td>
<td>6,362</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N. America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.78</td>
<td>0.73-0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td>Asia</td>
<td>0.79</td>
<td>0.70-0.85</td>
<td>0.86</td>
<td>60.6</td>
<td>56.6-64.6</td>
<td>0.027</td>
<td>3,123</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Europe/</td>
<td>0.78</td>
<td>0.74-0.81</td>
<td>66.1</td>
<td>61.5-70.7</td>
<td>2,880</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N. America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.77</td>
<td>0.72-0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Troponin</td>
<td>Asia</td>
<td>0.81</td>
<td>0.77-0.85</td>
<td>0.42</td>
<td>61.1</td>
<td>58.3-63.8</td>
<td>0.011</td>
<td>7,308</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Europe/</td>
<td>0.79</td>
<td>0.74-0.83</td>
<td>65.0</td>
<td>63.2-66.7</td>
<td>8,690</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N. America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.8</td>
<td>0.77-0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-dimer</td>
<td>Asia</td>
<td>0.78</td>
<td>0.76-0.82</td>
<td>&lt;0.00001</td>
<td>57.9</td>
<td>55.8-59.9</td>
<td>&lt;0.00001</td>
<td>14,076</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Europe/</td>
<td>0.69</td>
<td>0.66-0.72</td>
<td>64.6</td>
<td>63.4-65.8</td>
<td>29,741</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N. America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.76</td>
<td>0.73-0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP</td>
<td>Asia</td>
<td>0.83</td>
<td>0.80-0.85</td>
<td>&lt;0.00001</td>
<td>57.8</td>
<td>55.7-59.9</td>
<td>&lt;0.00001</td>
<td>10,407</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Europe/</td>
<td>0.67</td>
<td>0.63-0.71</td>
<td>64.8</td>
<td>62.8-66.7</td>
<td>28,693</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N. America</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>0.78</td>
<td>0.74-0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1:** Summary of pooled area under curves and mean ages (+/SD) for the 5 biomarkers investigated.
AUC: area under curve

CI: confidence interval

SD: standard deviation

IL-6: interleukin-6

CRP: C-reactive protein
**Figure legends**

**Figure 1**: PRISMA flow diagram for paper review, selection, and inclusion in these meta-analyses.

**Figure 2**: Summary forest plot demonstrating pooled area under curves for the five biomarkers being meta-analyzed (CRP, D-dimer, troponin, urea, and IL-6). For ease of comparison values for Asian countries are shown in red and values for European/North American countries are shown in blue. There was insufficient data to accurately construct pooled AUCs for other geographical regions.

**Figure 3**: Forest plot demonstrating all individual studies contributing to the meta-analyses for (A) CRP and (B) D-dimer in Asian and European/North American countries for the first tranche of data (from Jan 2020 – June 2021).

Blue diamond: pooled AUC for European/North American countries

Red diamond: pooled AUC for Asian countries

The size of each square representing an individual study corresponds with the size of the study population:

Font 4: 0-100

Font 5: 100-1,000

Font 6: 1,001-10,000

Font 7: >10,001
For the reader’s ease the data sources are arranged alphabetically both within the figure and in the references (meta-analyses) sections.

Figure 4: Forest plot demonstrating all individual studies contributing to the meta-analyses for (A) IL-6, (B) troponin, and (C) urea in Asian and European/North American countries for the first tranche of data (from Jan 2020 – June 2021).

Blue diamond: pooled AUC for European/North American countries

Red diamond: pooled AUC for Asian countries

The size of each square representing an individual study corresponds with the size of the study population:

Font 4: 0-100

Font 5: 100-1,000

Font 6: 1,001-10,000

Font 7: >10,001
References (Meta-analyses)


patients, with a focus on diabetes mellitus - A retrospective study in Andhra Pradesh, India. Diabetes Metab Syndr. 2021;15(3):739-45.


Identification of studies via databases and registers

Identification:
- Records identified from PubMed: (n = 1,930)
  - Records removed for lack of original data; e.g. identified on PubMed as book chapter, review or systematic review, meta-analyses (n = 279)

Screening:
- Records screened (n = 1,651)
  - Reports excluded (n = 1,130)
    - Specific patient groups e.g. ITU (n = 451)
    - Specific focus e.g. VTE (n = 238)
    - Different geographical locations (n = 107)
    - No patient data e.g. editorial (n = 107)
    - Other e.g. cell/animal model (n = 273)

Included:
- Reports assessed for eligibility (n = 521)
  - Reports excluded:
    - No AUC data regarding biomarkers of interest (n = 395)
  - Reports included:
    - Original data from Cambridge, UK (Soon) and Spain (Soares)
    - Reports referred from authors contacted

- Studies included in review (n = 130)
Pooled AUCs for biomarkers in SARS-CoV-2

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Region</th>
<th>Area under the curve (95% CI)</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-6</td>
<td>Asia</td>
<td>0.86 (0.81-0.90)</td>
<td>2,192</td>
</tr>
<tr>
<td></td>
<td>Europe/N.A.</td>
<td>0.70 (0.64-0.75)</td>
<td>6,145</td>
</tr>
<tr>
<td>Urea</td>
<td>Asia</td>
<td>0.79 (0.70-0.85)</td>
<td>3,021</td>
</tr>
<tr>
<td></td>
<td>Europe/N.A.</td>
<td>0.78 (0.74-0.81)</td>
<td>2,880</td>
</tr>
<tr>
<td>Troponin</td>
<td>Asia</td>
<td>0.81 (0.77-0.85)</td>
<td>7,308</td>
</tr>
<tr>
<td></td>
<td>Europe/N.A.</td>
<td>0.79 (0.74-0.83)</td>
<td>8,690</td>
</tr>
<tr>
<td>D-dimer</td>
<td>Asia</td>
<td>0.78 (0.76-0.82)</td>
<td>14,076</td>
</tr>
<tr>
<td></td>
<td>Europe/N.A.</td>
<td>0.69 (0.66-0.72)</td>
<td>29,741</td>
</tr>
<tr>
<td>CRP</td>
<td>Asia</td>
<td>0.83 (0.80-0.85)</td>
<td>10,407</td>
</tr>
<tr>
<td></td>
<td>Europe/N.A.</td>
<td>0.67 (0.63-0.71)</td>
<td>28,693</td>
</tr>
</tbody>
</table>