



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/241581/>

Version: Published Version

Article:

Shuvo, M.E.R., Schwiening, M., Avramidis, N. et al. (2026) Tracking the evolution of biomarker efficacy in SARS-CoV-2: a global meta-analyses series. Respiratory Research, 27. 265. ISSN: 1465-9921

<https://doi.org/10.1186/s12931-026-03656-9>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Tracking the evolution of biomarker efficacy in SARS-CoV-2: a global meta-analyses series

ME Rahman Shuvo^{*,1} MBBS, Max Schwiening^{*,2} BA, Nikos Avramidis MSc³, Felipe Soares^{4,5} PhD, Oliver Feng⁶ PhD, Susana Abreu² PhD, Niki Veale^{1,2} MBBS, William Thomas¹ MBBS, AA Roger Thompson⁴ PhD, Richard J Samworth⁶ PhD, Nicholas W Morrell¹ PhD, Malcolm G Semple^{7,8} PhD, Kenneth Baillie³ PhD, Stefan J Marciniak^{1,2} PhD, Elaine Soon^{1,2} PhD (0000-0002-5744-5014).

¹Department of Respiratory Medicine, Box 40, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK.

²Cambridge Institute for Medical Research, University of Cambridge, Keith Peters Building, Hills Rd, Cambridge CB2 0XY, UK.

³Baillie Gifford Pandemic Science Hub, Centre for Inflammation Research, University of Edinburgh, 47 Little France Crescent, Edinburgh, EH16 4TJ, UK.
Roslin Institute, University of Edinburgh, Easter Bush, Midlothian EH25 9RG.

⁴Department of Clinical Medicine, University of Sheffield, Beech Hill Road, Sheffield S10 2RX, UK.

⁵Universidade Federal do Rio Grande do Sul, Farroupilha, Porto Alegre - RS, 90010-150, Brazil.

⁶Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WB, UK.

⁷NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Institute of Infection, Veterinary, and Ecological Sciences, University of Liverpool, Ronald Ross Building, 8 West Derby St, Liverpool, L69 7BE, UK.

⁸Liverpool Institute for Child Health and Wellbeing, Alder Hey Children's Hospital, Eaton Road, Liverpool, L12 2AP, UK.

*these authors contributed equally to this work.

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

Abstract

Background: Sophisticated prognostic scores have been proposed for SARS-CoV-2 but do not always perform consistently. We performed this systematic review to discover why and to investigate the impact of vaccination and viral variants.

Methods: We searched the PubMed database for the keywords ‘SARS-CoV-2’ or ‘Covid19’ with ‘biomarker’ and ‘mortality’ for the baseline tranche (01/12/2020-30/06/2021) and either ‘SARS-CoV-2’ or ‘Covid19’ with ‘biomarker’ and either ‘vaccination’ or ‘variant’ from 01/12/2020 to 31/10/2023. To aggregate the data, the *meta* library in R was used, and a random effects model fitted to obtain pooled AUCs and 95% confidence intervals.

Results: We screened 4,688 potential source manuscripts and included 144 in the final analyses. Biomarker effectiveness varies significantly by geographical region. Admission CRP levels were a good prognostic marker for mortality due to wild-type virus 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/Northern America, with a pooled AUC of 0.67 (95%CI 0.63-0.71, $P<0.0001$). The same pattern applies to D-dimer and IL-6. Notably, urea and troponin had pooled AUCs ≥ 0.78 regardless of location, implying that end-organ damage at presentation was a key prognostic factor in wild-type SARS-CoV-2 infection. CRP, D-dimer, and IL-6 have generally declined in effectiveness in the vaccinated and variant cohorts. We note a significant lag from the pandemic advent to data availability and that the type of data varied considerably between studies.

Conclusion: Biomarkers and prognostic scores should be tailored to populations. It is imperative that the infrastructure for collecting clinical data should be put in place ahead of a future

pandemic so that data harvesting, collation and analysis can occur in a robust and timely manner to aid clinical decision making.

Registration: This study was registered with PROSPERO (CRD42022366893).

Keywords:

SARS-CoV-2

Covid19

biomarker

mortality

Clinical Trial Number: Not applicable.

ARTICLE IN PRESS

Background

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel beta coronavirus of zoonotic origin first identified in Wuhan, China at the end of 2019, which led to a Public Health Emergency of International Concern between February 2020 and May 2023. SARS-CoV-2 differs from previous viral threats in showing marked transmissibility during the asymptomatic/very early symptomatic stage[1] and person-to-person transmission by both airborne and fomite routes[2]. At the beginning of the pandemic, there was no previous immunity, no known effective treatment, and no vaccine, resulting in a global death toll of >6.9 million (<https://covid19.who.int/>).

As a large amount of clinical data was amassed quickly (in epidemiological timeframes) in a completely naïve global population, we can track the efficacy of biomarkers and assess the impact of viral mutations and mass vaccination programs. Some prognostic scores do not do well when applied away from their derivation region/population. For example, El-Solh[3] tested 4 peer-reviewed prognostic models constructed to predict in-hospital mortality for SARS-CoV-2 patients; proposed by Chen[4], Shang[5], Yu[6], and Wang[7]. Each model examined had a validation area under curve (AUC) significantly worse than that generated for their derivation cohorts. For example, the AUC of the validation cohort using the model proposed by Chen[4] 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 three models.

Gupta[8] 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 nine 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.

Interestingly, the ISARIC 4C Mortality score has performed remarkably consistently with an initial AUC of 0.78 (0.78-0.79)[9] and follow-up scores of 0.71-0.82[10-13]. We conducted the following systematic review and meta-analyses series to see if we could explain these discrepancies, and to examine the impact of successive rounds of viral mutations and vaccination programs on biomarker efficacy.

Methods

Our primary intention with these meta-analyses was to examine which biomarkers would be most useful in predicting mortality for the adult patient who presented to the emergency department and was then admitted under the general medical team from the community. This

was done so that the biomarkers would have the widest applicability. To do this, 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 01st December 2019 to 30th June 2021. The second search covered 01st December 2019 to 31st October 2023 and used the following keywords:

- a) ‘SARS-CoV-2’ with ‘vaccination’ and ‘biomarker name’
- b) ‘SARS-CoV-2’ with ‘variant’ and ‘biomarker name’
- c) ‘Covid19’ with ‘vaccination’ and ‘biomarker name’
- d) ‘Covid19’ with ‘variant’ and ‘biomarker name’

We examined all papers which fulfilled the following criteria:

- a) Age of participants ≥ 18 years, with confirmed positivity for SARS-CoV-2,
- b) Availability of biomarker levels within 48 hours of presentation to the emergency department from the community,
- c) Availability of outcome data (survival versus death), and
- d) Being written in English or having an English translation available.

If one of these datasets was missing the corresponding author on the study was contacted and the missing information requested. Ethical approval was obtained from the UK Integrated Research Application System (reference 281880) for analysis of the Cambridge (UK) data, in accordance with the Declaration of Helsinki. All other data has either been published and/or is held in the public domain as an anonymized resource.

We excluded reports of patients who were limited to those already admitted to intensive care. We would emphasize that we did not exclude studies where patients who were admitted generally and then sent to intensive care. We also included studies with critically ill patients who were not admitted to intensive care, whether for reasons of ITU availability or otherwise. We excluded cohorts that focused exclusively on a single disease process *e.g.* studies done in cancer or transplant patients. To see if we could include pregnant women and paediatric cases, the search algorithm was re-interrogated and all manuscripts regarding pregnant women and children extracted. 24 papers met this criterion and were examined (summarised in Supplementary Table 1 in the online supplement). However, none of them had sufficient information to add to the meta-analyses. Mortality (30-day or in-hospital) was used as the endpoint.

Our first survey revealed that the majority of studies were based in Asia, Europe or Northern America. We defined Asian countries as being bounded by the Indian Ocean, the Pacific Ocean, the Arctic Ocean, the Ural Mountains and the Caspian Sea and separated from the Middle East along Iran and the Caucasus Mountains. European countries were defined as being bounded by the Arctic Ocean, Atlantic Ocean, Mediterranean Sea and the Ural Mountains and the Caspian Sea. Northern American countries included Canada, the United States, Greenland and Bermuda. There were insufficient numbers of studies originating from other locations to perform adequate meta-analyses and therefore we have concentrated on two main groups: Asian countries and European/Northern American countries. This study was registered with PROSPERO (CRD42022366893); and the PRISMA flow diagram is shown in Figure 1. The root source articles and further information are available in the online Supplement (Supplementary Figures 1-5 and references) and at <https://covid19.cimr.cam.ac.uk/>.

The area under curve (AUC) was chosen as the test parameter as most of the studies analyzed had different methods of measuring their biomarkers, used different units for measurement and reported their results in different ways (e.g. by reporting odds ratios or hazard ratios per unit increase in biomarker). Therefore we pragmatically adopted the AUC as a way of estimating the efficacy of a particular biomarker for a cohort which was easy to generate, independent of the method of measurement, and which could be pooled statistically for comparison. If the AUC was not reported, we contacted the corresponding authors to request this.

Statistical analysis

Meta-analysis of diagnostic discrimination is ideally performed using threshold-specific sensitivity and specificity, which enables hierarchical ROC modelling. However, the primary studies reported only AUC estimates and confidence intervals, precluding such approaches. We were unable to get more data as there were limits on the time and effort we could request of the source manuscript authors due to the ongoing pandemic. In this setting, transformation-based inverse-variance meta-analysis of AUCs was the only feasible quantitative synthesis.

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. Further details are available in the online Supplement.

To summarize, we analyzed logit-transformed AUCs, which respected the bounded nature of the statistic and leveraged the asymptotic normality underlying the confidence intervals reported in the original studies. Therefore, the pooled estimate should be interpreted as a descriptive summary of discrimination across heterogeneous populations rather than as an estimate of expected performance in a new setting. Due to the limitations inherent in this approach we have not performed further statistical analyses.

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 2-6). Note that when fitting a random effects meta-analysis model, the individual study means are assumed to be random, and the between-study heterogeneity (τ^2) needs to be estimated. For the pooled AUC, we used a single estimate of τ^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 the root data underpinning the meta-analyses on a publicly available website (<https://covid19.cimr.cam.ac.uk/>). Study authors have also volunteered that statistical software is expensive and hence inaccessible. Therefore we have written a program in R which allows for calculation for the AUC of a biomarker which is free to download from the same website. Our intention is that everyone will be able to view the most effective biomarkers for their locale from the website. The website was assessed 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 on 01st December 2019 to 30th June 2021 for the baseline study and 2,758 papers from 01st December 2019 to 31st October 2023 to obtain vaccination and variant data. The first phase meta-analyses revealed different patterns in the effectiveness of biomarkers in different regions of the world. For example, admission CRP levels were a good prognostic marker for mortality in Asian countries, with a pooled area under curve (AUC) of 0.83 (95% confidence intervals [CI] 0.80-0.85), but only an average predictor of mortality in Europe and Northern America, with a pooled AUC of 0.67 (95% CI 0.63-0.71, $P < 0.0001$, Table 1, Fig.2A). This was also true for

admission D-dimer (Fig.2B), with a pooled AUC of 0.78 (95% CI 0.76-0.82) in Asian countries and a pooled AUC of 0.69 (95% CI of 0.66-0.72) in European and Northern American countries and for IL-6 (Fig.3A) levels. This may explain why some 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.

Interestingly, urea and troponin levels had universally ‘good’ pooled AUCs (Fig.3B-C). For example, the pooled AUC for troponin is 0.81 (95% CI 0.77-0.85) in Asia and 0.79 (95% CI 0.74-0.83) in Europe and Northern America. This implies that end-organ damage at the time of presentation was a key prognostic indicator of severity for wild-type SARS-CoV-2 infection. Age is also a reasonable prognostic biomarker with an AUC of 0.78 (95% CI 0.77-0.80) in Europe and Northern America and 0.73 (95% CI 0.65-0.79) in Asia (Fig.3D). The results of meta-analyses for all five biomarkers and age in the first tranche (relating to wild-type virus) are summarized in Table 1 and Fig.4A. We note that age performed better or at least as well as any of the five biochemical parameters measured (D-dimer, CRP, urea, troponin, and interleukin-6) in European and Northern American countries.

Pooling all the results from Asian, European, and Northern 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 skew the

results and mask the fact that admission levels of CRP, D-dimer and IL-6 were only average at predicting mortality in European and North American countries.

We expected that multiple rounds of vaccinations and ongoing mutations into different strains would significantly impact biomarker efficacy. There were insufficient numbers of studies to conduct a meta-analysis and so all the studies are collated and shown in Table 2, Table 3 and Figure 5. The ‘inflammatory’ biomarkers (CRP, D-dimer, and IL-6) have generally declined in effectiveness in the vaccinated cohorts (Fig.5A-B). This is particularly illustrated by our largest cohort (ISARIC4C), where the AUC for CRP in the unvaccinated Alpha cohort is 0.61 (95% CI 0.61-0.62) and in the corresponding vaccinated cohort is 0.51 (95% CI 0.39-0.63).

The same pattern is seen in the variant cohorts, with most studies showing that the inflammatory biomarkers have lower AUCs with variant cohorts (Table 3). Interestingly, the Jalali cohort[14] (Delta variant), who are completely unvaccinated, have AUCs of 0.576 (95% CI 0.434-0.713) for CRP and 0.620 (95% CI 0.481-0.760) for D-dimer. This is consistent with reports that the pathogenicity of SARS-CoV-2 has decreased in the order wild-type>Delta>Omicron and that the Omicron variant in particular is associated with decreased IL-6 production[15, 16]. It is also likely that the targeting of inflammation generally (*e.g.* dexamethasone) and IL-6 pathways specifically (*e.g.* tocilizumab) has had an effect; in that we are now able to effectively neutralize at least some of the inflammatory effects of severe SARS CoV-2.

To examine this in greater detail we have interrogated data collected from the ISARIC4C[17], GenOMICC[18] and COG-UK[19] consortiums. Fig.4B shows increasing loss of efficacy for CRP from all unvaccinated in the cohort to those vaccinated to intensive care admissions. The AUCs for urea vary less between vaccinated and unvaccinated and have greater prognostic capability in every patient category examined, as does age. We think this shows that co-morbidity and/or frailty are now the leading factors in predicting susceptibility to fatal SARS-CoV-2 in the UK.

Discussion

We demonstrate conclusively for the first time that biomarker effectiveness for mortality in SARS-CoV-2 varied significantly by geographical location. We hypothesize that biomarker efficacy in any population at this given time hinges on many factors including:

- A) Population factors – the numbers vaccinated and/or previously infected, the age and co-morbidities present,
- B) Variant factors such as pathogenicity and transmissibility,
- C) Socio-economic factors such as access to healthcare and nutritional status,
- D) Health systems factors that affect the time and severity of illness at presentation.

Prognostic biomarkers/scores are likely to perform better if individualized to specific populations. Consistent with our findings, Marino *et al*[20] demonstrated that a prognostic score developed in the same country (PREDI-CO, Bartoletti *et al*[21], 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*[22], 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. The CURB-65 score is well-known and has been validated as a tool for predicting mortality in community-acquired pneumonia[23]. It was developed in the UK, New Zealand and the Netherlands. European studies confirm an AUC of 0.759-0.780[24, 25]. However it performs less satisfactorily in older populations[26]. For example, Shirata *et al*[27] 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 three cohorts[28]). Hincapié suggested that this may be due to the factors underpinning a significant difference in community-acquired pneumonia associated mortality (9.5%[23] *versus* 17-32%[28]).

It is not possible to know from these descriptive meta-analyses exactly why these regional differences in biomarker effectiveness existed for the first wave of SARS-CoV-2. This was likely due to a combination of multiple factors, including differences in populations (age, co-morbidities, genetics), treatment strategies, health system structures, and public health policy; which are summarized in Table 4. In the first, wild-type wave; Asian cohorts were universally younger than the European and Northern American cohorts in all five parameters we investigated (CRP, D-dimer, troponin, urea, and IL-6; Table 1). Gomez-Mesa *et al*[29] showed that the point

prevalence of every single co-morbidity bar diabetes (including hypertension, heart failure, chronic kidney disease, coronary artery disease and cancer) was higher in European men *versus* Asian men. It is possible that in the first wave, younger Asian patients were dying predominantly from cytokine storm (hence the marked prognostic value of the ‘inflammatory’ markers such as CRP, D-dimer, and IL-6), while in Europe and Northern America older people were dying from multi-organ failure. We note that age remained a significant prognostic factor in all ISARIC4C cohorts (all vaccinated, all unvaccinated, vaccinated in ITU, and unvaccinated in ITU; Fig.4B) and indeed in every single cohort where reported. It is probable that age is essentially a proxy for co-morbidity and/or frailty.

Other critical regional differences include different treatment strategies, health system structures, and public health policy. Asian policies were pragmatic and incorporated early use of antivirals, repurposing of other drugs, and integration with containment measures including widespread masking protocols. Zhang *et al*[30] described the use of specialized ‘fever clinics’ to triage patients with respiratory symptoms using a combination of blood biomarkers (CRP, lymphocyte counts) and CT scanning to diagnose viral pneumonia which were isolated and then sent for specific SARS-CoV-2 testing. This approach was adopted due to its success in containing the severe acute respiratory syndrome (SARS) epidemic in 2003. The focus in Europe was on supportive care, mass testing including the use of home lateral flow tests[31], and episodic social contact-limiting measures (‘lockdowns’) in order to prevent local healthcare providers from being overwhelmed. In particular, use of repurposed drugs was more restricted to trial formats in Europe, with strong emphasis on large RCTs and widespread adoption of dexamethasone after publication of the RECOVERY trial[32].

Multinational practice data have also demonstrated variation in respiratory support strategies and ITU admission rates. These include higher reported use of prone positioning in parts of Asia and Europe ($\approx 70\text{--}85\%$) compared with North America ($\approx 50\text{--}70\%$). ITU admission rates also differed: 35% in the USA, 17% in the UK, and 14% in China[33]; and in many countries was influenced by ITU capacity due to an overwhelming need. The corresponding ITU mortality was 29% (USA), 33% (UK) and 24% in China. In parallel, pandemic-associated disruption of non-COVID care was more pronounced in Asia, where Teglia *et al*[34] indicate treatment reductions in cancer services approaching -42% (95% CI: -50 to -35%) compared with -35% (95% CI: -47 to -22%) in North America and -8% (95% CI: -16 to -0.2%) in Europe, suggesting greater system-level interruption and diagnostic or therapeutic delay. Mechanistically, these differences likely arise from (i) variable ICU capacity and escalation criteria, (ii) differential early reliance on imaging versus molecular testing, (iii) intensity of lockdown measures affecting care continuity, and (iv) population-level risk structures. Together, the literature supports a region-adaptive framework for clinical management and biomarker selection, recognizing that inflammatory phenotyping may be more informative in settings with younger populations and hyperinflammatory presentations, whereas end organ-damage-oriented risk stratification may be prioritized in systems with older patients and high comorbidity burden.

This study has several limitations. First, insufficient numbers of studies were located in continents other than Asia and Northern America/Europe to perform an adequate meta-analysis for the first tranche (wild-type virus). There were also inadequate numbers of published studies in vaccinated and variant cohorts for meta-analyses and so we have elected to collate and show

the results of each study individually. We also do not have any data to make inferences for the pregnant or paediatric populations. Due to the nature of our representative characteristic (the area under curve for the receiver operator characteristic) we are unable to suggest threshold values for the biomarkers in question. Our pooled estimate of AUCs should be interpreted as a descriptive summary of discrimination across heterogeneous populations rather than as an estimate of expected performance in a new setting; since the robustness of our statistical approach has been limited by the data available. The biomarker detection time varies in the studies we have analysed and collated and this also introduces further uncertainty.

Finally and most importantly, **we need to be better prepared to face another pandemic.** There was a significant lag between data availability and the clinical need for it; and the data itself was not collected in a consistent manner which would have allowed for a more robust analysis. For example it took over a year for data collection for our initial wild-type tranche to be completed, by which time SARS-CoV-2 had evolved sufficiently for the original analyses to be outdated. Access to statistical software was also limited and this led us to write a free-to-use program in R to calculate AUC (available to download at <https://covid19.cimr.cam.ac.uk/>).

The simplest and most practical way to overcome these issues would be to set up information frameworks ahead of time, ideally online, in an automated manner and with ethical permissions already in place. In this way, data collection can be standardized, rather than growing organically as was the case with the SARS-CoV-2 pandemic. The availability of such standardized data would facilitate rapid collation and robust meta-analyses and would reduce the time and effort

required from clinicians and researchers, which will be in short supply during a pandemic. We strongly recommend the use of open-source tools to facilitate this, such as the WHO/ISARIC Clinical Characterization Protocol (<https://isaric.net/ccp/>)[35, 36]. This protocol is free to use and investigators will fully control their data and samples. Specific centers could also be selected to gather information on groups such as pregnant women and children, who are too often left out of studies. This system would allow the rapid identification of population-specific biomarkers and the organization of clinical trials by pinpointing ‘at-risk’ groups who could be targeted early for vaccination or intervention programs.

Conclusions

Biomarkers and prognostic scores should be tailored to particular populations. It is imperative that the infrastructure for collecting clinical data should be put in place ahead of a future pandemic so that data harvesting, collation and analysis can occur in a robust manner, and the results available in time to assist clinical decision making.

List of abbreviations

AUC: area under curve

CI: confidence interval

COG-UK: Covid-19 Genomics UK Consortium

CRP: C-reactive protein

IL-6: interleukin-6

IQR: interquartile range

ISARIC: International Severe Acute Respiratory and emerging Infection Consortium

ITU: intensive care unit

SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2

WHO: World Health Organization

Biomarker	Location	Pooled AUC	95% CI	P-value for AUC	Mean age	IQR	P-value for age	No. of patients	No. of studies
D-dimer	Asia	0.78	0.76-0.82	<0.00001	57.9	55.8-59.9	<0.00001	14,076	38
	Europe/ N. America	0.69	0.66-0.72		64.6	63.4-65.8		29,741	22
	All	0.76	0.73-0.78					43,817	60
CRP	Asia	0.83	0.80-0.85	<0.00001	57.8	55.7-59.9	<0.00001	10,407	34
	Europe/ N. America	0.67	0.63-0.71		64.8	62.8-66.7		28,693	21
	All	0.78	0.74-0.81					39,100	55
Urea	Asia	0.79	0.70-0.85	0.86	60.6	56.6-64.6	0.027	3,123	10
	Europe/ N. America	0.78	0.74-0.81		66.1	61.5-70.7		2,880	6
	All	0.77	0.72-0.82					6,003	16
Troponin	Asia	0.81	0.77-0.85	0.42	61.1	58.3-63.8	0.011	7,308	16
	Europe/ N. America	0.79	0.74-0.83		65.0	63.2-66.7		8,690	16
	All	0.8	0.77-0.83					15,998	32
IL-6	Asia	0.86	0.81-0.90	<0.00001	58.2	53.6-62.9	0.033	2,993	17
	Europe/ N. America	0.70	0.64-0.75		63.5	61.1-65.9		6,362	20
	All	0.78	0.73-0.83					9,355	37
Age	Asia	0.73	0.65-0.79	0.0408	57.3	6.92*	0.028	2,652	9
	Europe/ N. America	0.78	0.77-0.80		62.9	2.75*		18,127	5
	All	0.75	0.70-0.90					20,779	14

Table 1: Results from the first meta-analyses series of wild-type virus showing pooled AUCs from the five biomarkers selected and age, with the ages and numbers of participants from each cohort.

*indicates standard deviation

P-values shown are for comparison between Asian and European/North American cohorts.

AUC: Area under curve

CI: confidence interval

IQR: interquartile range

IL-6: interleukin-6

ARTICLE IN PRESS

Source Country	Parameter	Pooled AUC reference	2020	2021	2022
ISARIC4C [11] UK Age: 53.3±14.2y	AUC CRP unvaccinated	<i>Wild-type</i> 0.67 (0.63-0.71)	<i>Alpha</i> 0.61 (0.61, 0.62) n=68,220 Age:55.8±12.9	<i>Delta</i> 0.60 (0.59, 0.61) n=47,217 Age:49.7±15.2	N/A
	AUC CRP fully vaccinated	N/A	0.51 (0.39,0.63) n=489 Age:60.9±9.2	0.59 (0.58, 0.61) n=11,891 Age:49.7±15.2	N/A
	AUC urea unvaccinated	0.78 (0.74-0.81)	0.73 (0.72, 0.73) n=85,387 Age:56.0±12.9	0.75 (0.74,0.76) n=50,390 Age:49.6±15.1	N/A
	AUC urea fully vaccinated	N/A	0.74 (0.64,0.84) n=551 Age:60.5±9.6	0.73 (0.72, 0.74) n=12,531 Age:49.6±15.0	N/A
	AUC age unvaccinated	0.78 (0.77-0.80)	0.71 (0.71,0.72) n=85,387 Age:56.0±12.9	0.76 (0.76, 0.77) n=50,390 Age:49.6±15.1	N/A
	AUC age fully vaccinated	N/A	0.80 (0.73, 0.87) n=551 Age:60.5±9.6	0.71 (0.70, 0.72) n=12,531 Age:49.6±15.0	N/A
Wang[7] Australia Age: 64.1±26.0y	AUC CRP all-cause mortality	<i>Wild-type</i> 0.67 (0.63-0.71)*	<i>B.1.338</i> 0.6 (0.49, 0.70) n=196 Age: 69.5 (48-84)	<i>Delta</i> 0.55 (0.46, 0.64) n=419 Age: 55 (40-71)	<i>Omicron</i> 0.61 (0.51, 0.70) n=467 Age: 70 (47-82)
	AUC CRP unvaccinated	N/A	N/A	0.58 (0.49, 0.68)	0.61 (0.45, 0.78)
	AUC CRP fully vaccinated	N/A	N/A	N/A	0.57 (0.44, 0.69)
	AUC D-dimer all-cause mortality	0.69 (0.66-0.72)*	0.69 (0.56, 0.82)	0.69 (0.59, 0.78)	0.63 (0.52, 0.73)
	AUC D-dimer unvaccinated	N/A	N/A	0.69 (0.58, 0.79)	0.66 (0.47, 0.85)
	AUC D-dimer fully vaccinated	N/A	N/A	0.58 (0.14, 1.00)	0.57 (0.43, 0.72)

Rzymiski [37] Poland Age: 68.2±15.8y	AUC CRP fully vaccinated	<i>Wild-type</i> 0.67 (0.63-0.71)	N/A	<i>Delta</i> 0.675 (0.640 - 0.709) n=741 Age: 68.2±15.8y	N/A
	AUC D-dimer fully vaccinated	0.69 (0.66-0.72)	N/A	0.649 (0.613 – 0.685) n=741	N/A
	AUC IL-6 fully vaccinated	0.7 (0.64-0.75)	N/A	0.738 (0.695 – 0.778) n=741	N/A
Patel[38] India Age: 54y (37-67)	AUC CRP Unvaccinated	<i>Wild-type</i> 0.83 (0.80-0.85)	N/A	<i>Omicron</i> 0.627 (0.389-0.865) n=79 Age: 61 (35-75)	
	AUC CRP fully vaccinated	N/A	N/A	0.771 (0.653-0.889) n=709 Age: 53 (37-67)	
	AUC D-dimer Unvaccinated	0.78 (0.76-0.82)	N/A	0.868 (0.715-1.00) n=79 Age: 61 (35-75)	
	AUC D-dimer fully vaccinated	N/A	N/A	0.667 (0.471-0.868) n=709 Age: 53 (37-67)	
Altintop [39] Türkiye Age: 72y (61-97)	AUC CRP unvaccinated	<i>Wild-type</i> 0.83 (0.80-0.85)	N/A	<i>Original/Alpha</i> 0.668 (0.592-0.739) n=79	N/A
	AUC CRP fully vaccinated	N/A	N/A	0.715 (0.603-0.811) n=169	N/A

Table 2: Comparison of biomarker efficacy in vaccinated versus unvaccinated cohorts, with the relevant pooled AUC from the first tranche meta-analyses (Table 1) as a reference, with values from Asian countries in red and from European and Northern American countries in blue.

*Note that the European/Northern American pooled AUCs were used as the references for the Australian cohort.

AUC: Area under curve

ISARIC4C: International Severe Acute Respiratory Infection Consortium Comprehensive Clinical Characterisation Collaboration

ARTICLE IN PRESS

Source Country	Parameter	Pooled AUC Reference	Alpha	Delta	Omicron
Homen-Fernandez[40] Spain Age: 80±14y	AUC CRP	0.67 (0.63-0.71)	0.521 0.324-0.712 n=97	0.586 0.425-0.748 n=98	0.631 0.450-0.811 n=99
	AUC D-dimer	0.69 (0.66-0.72)	0.650 0.328-0.921 n=90	0.596 0.290-0.902 n=70	0.651 0.452-0.850 n=57
	AUC trop	0.81 (0.77-0.85)	0.94 0.876-1.000 n=87	0.788 0.657-0.921 n=74	0.883 0.754-1.000 n=34
	AUC age	0.78 (0.77-0.80)	0.886 0.790-0.982 n=100	0.877 0.806-0.947 n=100	0.749 0.642-0.856 n=100
Meletis[41] Greece Age: 55.2±23.9y	AUC CRP	0.67 (0.63-0.71)	0.764 0.660-0.867 n=129	0.767 0.655-0.878 n=130	0.964 0.914-1.000 n=50
	AUC D-dimer	0.69 (0.66-0.72)	0.729 0.600-0.858 n=119	0.800 0.699-0.901 n=122	0.638 0.494-0.782 n=45
	AUC urea	0.77 (0.72-0.82)	0.729 0.612-0.846 n=139	0.638 0.494-0.781 n=136	0.89 0.793-0.986 n=34
	AUC IL-6	0.7 (0.64-0.75)	0.809 0.708-0.911 n=105	0.818 0.707-0.930 n=112	0.883 0.754-1.000 n=60
Jalali[14] Iran Age: 52.4±16.1y	AUC CRP*	0.83 (0.80-0.85)	N/A	0.576 0.434-0.713 n=75	N/A
	AUC D-dimer*	0.78 (0.76-0.82)	N/A	0.620 0.481-0.760 n=67	N/A
Liu[42] China Age: 50.2±17.4y	AUC CRP*	0.83 (0.80-0.85)	N/A	N/A	0.737 0.671-0.804 n=311
	AUC urea*	0.79 (0.70-0.85)	N/A	N/A	0.574 0.482-0.665 n=311

Table 3: Comparison of biomarker efficacy in viral variant cohorts, with the relevant pooled AUC from the initial meta-analyses (Table 1) as a reference, with values from Asian countries in red and from European and Northern American countries in blue.

*Indicate that the AUCs shown were for critical illness rather than mortality.

AUC: Area under curve

IL-6: interleukin-6

ARTICLE IN PRESS

Aspect	Asia	Europe/Northern America
Viral exposure	First exposure	'Conditioned' by prior Asian experience
Population	Younger, fewer co-morbidities[29]	Older, more co-morbidities
Initial treatment strategy	Integrated antiviral + protocol bundles[30]	Focused on supportive care
Evidence threshold	More willingness to use observational / pragmatic evidence	Strong emphasis on large RCTs
Public health integration	Treatment integrated with containment strategy	Treatment separate from containment, emphasis on vaccination
Drug repurposing	Early broad repurposing common	Initially variable, later standardized
Guideline speed	Faster early protocol deployment	Slower adoption, frequent revisions

Table 4: Summary of potential mechanistic differences between Asian and European/Northern American regions.

References (body of manuscript):

1. Yanes-Lane, M., et al., *Proportion of asymptomatic infection among COVID-19 positive persons and their transmission potential: A systematic review and meta-analysis*. PLoS One, 2020. **15**(11): p. e0241536.
2. Bak, A., et al., *SARS-CoV-2 routes of transmission and recommendations for preventing acquisition: joint British Infection Association (BIA), Healthcare Infection Society (HIS), Infection Prevention Society (IPS) and Royal College of Pathologists (RCPath) guidance*. J Hosp Infect, 2021. **114**: p. 79-103.
3. El-Solh, A.A., et al., *Comparison of in-hospital mortality risk prediction models from COVID-19*. PLoS One, 2020. **15**(12): p. e0244629.
4. Chen, R., et al., *Risk Factors of Fatal Outcome in Hospitalized Subjects With Coronavirus Disease 2019 From a Nationwide Analysis in China*. Chest, 2020. **158**(1): p. 97-105.
5. Shang, Y., et al., *Scoring systems for predicting mortality for severe patients with COVID-19*. EClinicalMedicine, 2020. **24**: p. 100426.
6. Yu, C., et al., *Clinical Characteristics, Associated Factors, and Predicting COVID-19 Mortality Risk: A Retrospective Study in Wuhan, China*. Am J Prev Med, 2020. **59**(2): p. 168-175.
7. Wang, K., et al., *Clinical and Laboratory Predictors of In-hospital Mortality in Patients With Coronavirus Disease-2019: A Cohort Study in Wuhan, China*. Clin Infect Dis, 2020. **71**(16): p. 2079-2088.

8. Gupta, R.K., et al., *Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study*. Eur Respir J, 2020. **56**(6).
9. Knight, S.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**: p. m33339.
10. Aletreby, W.T., et al., *External Validation of 4C ISARIC Mortality Score in Critically ill COVID-19 Patients from Saudi Arabia*. Saudi J Med Med Sci, 2022. **10**(1): p. 19-24.
11. Crocker-Buque, T., et al., *Using ISARIC 4C mortality score to predict dynamic changes in mortality risk in COVID-19 patients during hospital admission*. PLoS One, 2022. **17**(10): p. e0274158.
12. Vallipuram, T., et al., *External validation of the ISARIC 4C Mortality Score to predict in-hospital mortality among patients with COVID-19 in a Canadian intensive care unit: a single-centre historical cohort study*. Can J Anaesth, 2023. **70**(8): p. 1362-1370.
13. Ong, S.W.X., et al., *Validation of ISARIC 4C Mortality and Deterioration Scores in a Mixed Vaccination Status Cohort of Hospitalized Coronavirus Disease 2019 (COVID-19) Patients in Singapore*. Clin Infect Dis, 2022. **75**(1): p. e874-e877.
14. Azarfar, F., et al., *Investigation of the relationship between monocyte chemoattractant protein 1 rs1024611 variant and severity of COVID-19*. Cytokine, 2023. **171**: p. 156367.
15. Barh, D., et al., *SARS-CoV-2 Variants Show a Gradual Declining Pathogenicity and Pro-Inflammatory Cytokine Stimulation, an Increasing Antigenic and Anti-Inflammatory Cytokine Induction, and Rising Structural Protein Instability: A Minimal Number Genome-Based Approach*. Inflammation, 2023. **46**(1): p. 297-312.

16. Korobova, Z.R., et al., *Cytokine Profiling in Different SARS-CoV-2 Genetic Variants*. Int J Mol Sci, 2022. **23**(22).
17. Garcia-Gallo, E., et al., *ISARIC-COVID-19 dataset: A Prospective, Standardized, Global Dataset of Patients Hospitalized with COVID-19*. Sci Data, 2022. **9**(1): p. 454.
18. Pairo-Castineira, E., et al., *Genetic mechanisms of critical illness in COVID-19*. Nature, 2021. **591**(7848): p. 92-98.
19. Wright, D.W., et al., *Tracking SARS-CoV-2 mutations and variants through the COG-UK-Mutation Explorer*. Virus Evol, 2022. **8**(1): p. veac023.
20. Marino, L., et al., *Time to hospitalisation, CT pulmonary involvement and in-hospital death in COVID-19 patients in an Emergency Medicine Unit*. Int J Clin Pract, 2021. **75**(9): p. e14426.
21. Bartoletti, M., et al., *Development and validation of a prediction model for severe respiratory failure in hospitalized patients with SARS-CoV-2 infection: a multicentre cohort study (PREDI-CO study)*. Clin Microbiol Infect, 2020. **26**(11): p. 1545-1553.
22. Yan, L., et al., *An interpretable mortality prediction model for COVID-19 patients*. Nature Machine Intelligence, 2020. **2**(5): p. 283-288.
23. Lim, W.S., et al., *Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study*. Thorax, 2003. **58**(5): p. 377-82.
24. Cilloniz, C., et al., *Machine-Learning Model for Mortality Prediction in Patients With Community-Acquired Pneumonia: Development and Validation Study*. Chest, 2023. **163**(1): p. 77-88.

25. Barlow, G., D. Nathwani, and P. Davey, *The CURB65 pneumonia severity score outperforms generic sepsis and early warning scores in predicting mortality in community-acquired pneumonia*. Thorax, 2007. **62**(3): p. 253-9.
26. Parsonage, M., et al., *Evaluation of the performance of CURB-65 with increasing age*. Clin Microbiol Infect, 2009. **15**(9): p. 858-64.
27. Shirata, M., et al., *Development and validation of a new scoring system for prognostic prediction of community-acquired pneumonia in older adults*. Sci Rep, 2021. **11**(1): p. 23878.
28. Hincapié, C., et al., *Community-acquired pneumonia: comparison of three mortality prediction scores in the emergency department*. Colomb Med (Cali), 2021. **52**(4): p. e2044287.
29. Gómez-Mesa, J.E., et al., *Marked Global Differences in Mortality in Male Patients with COVID-19: An Analysis of the CARDIO COVID 19-20 and WHF COVID-19 CVD Studies*. Glob Heart, 2025. **20**(1): p. 21.
30. Zhang, J., et al., *Therapeutic and triage strategies for 2019 novel coronavirus disease in fever clinics*. The Lancet Respiratory Medicine, 2020. **8**(3): p. e11-e12.
31. Iacobucci, G. and R. Coombes, *Covid-19: Government plans to spend £100bn on expanding testing to 10 million a day*. BMJ, 2020. **370**: p. m3520.
32. Horby, P., et al., *Dexamethasone in Hospitalized Patients with Covid-19*. N Engl J Med, 2021. **384**(8): p. 693-704.
33. Chang, R., et al., *COVID-19 ICU and mechanical ventilation patient characteristics and outcomes-A systematic review and meta-analysis*. PLoS One, 2021. **16**(2): p. e0246318.

34. Teglia, F., et al., *Global Association of COVID-19 Pandemic Measures with Cancer Treatment: A Systematic Review and Meta-Analysis*. *Cancers*, 2022. **14**(22): p. 5490.
35. Dunning, J.W., et al., *Open source clinical science for emerging infections*. *Lancet Infect Dis*, 2014. **14**(1): p. 8-9.
36. *Global outbreak research: harmony not hegemony*. *Lancet Infect Dis*, 2020. **20**(7): p. 770-772.
37. Rzymiski, P., et al., *Severe Breakthrough COVID-19 Cases during Six Months of Delta Variant (B.1.617.2) Domination in Poland*. *Vaccines (Basel)*, 2022. **10**(4).
38. Patel, A.K., et al., *COVID-19 patients' clinical profile and outcome with respect to their vaccination status: A prospective observational multicentre cohort study during third wave in Western India*. *Indian J Med Microbiol*, 2023. **41**: p. 28-32.
39. Altintop, S.E., et al., *COVID-19 in elderly: Correlations of viral load, clinical course, laboratory parameters, among patients vaccinated with CoronaVac*. *Acta Microbiol Immunol Hung*, 2022. **69**(4): p. 277-282.
40. Homen-Fernandez, J.R., et al., *Differences in the inflammatory response among hospitalized patients with distinct variants of SARS-CoV-2*. *Front Immunol*, 2023. **14**: p. 1267991.
41. Meletis, G., et al., *Variant-Related Differences in Laboratory Biomarkers among Patients Affected with Alpha, Delta and Omicron: A Retrospective Whole Viral Genome Sequencing and Hospital-Setting Cohort Study*. *Biomedicines*, 2023. **11**(4).
42. Liu, H., et al., *Characteristics of the Severe Acute Respiratory Syndrome Coronavirus 2 Omicron BA.2 Subvariant in Jilin, China from March to May 2022*. *J Transl Int Med*, 2022. **10**(4): p. 349-358.

ARTICLE IN PRESS

Figure Legends

Figure 1: A modified PRISMA flow diagram for paper review, selection, and inclusion in these meta-analyses for both the original/wild-type virus (in purple) and for variant and vaccination-related data (in orange).

Figure 2: Forest plots demonstrating all individual studies contributing to the meta-analyses for (A) CRP and (B) D-dimer in Asian and European/Northern American countries for the first wild-type wave (data from Dec 2019 – June 2021). Note a significant number of studies did not quote 95% confidence intervals (18 of 75 for CRP, 14 of 78 for D-dimer, 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 Fig.2 and Fig.3 for completeness but are not included in calculation of the pooled AUCs.

Blue diamond: pooled AUC for European/Northern American countries

Red diamond: pooled AUC for Asian countries

The size of each symbol

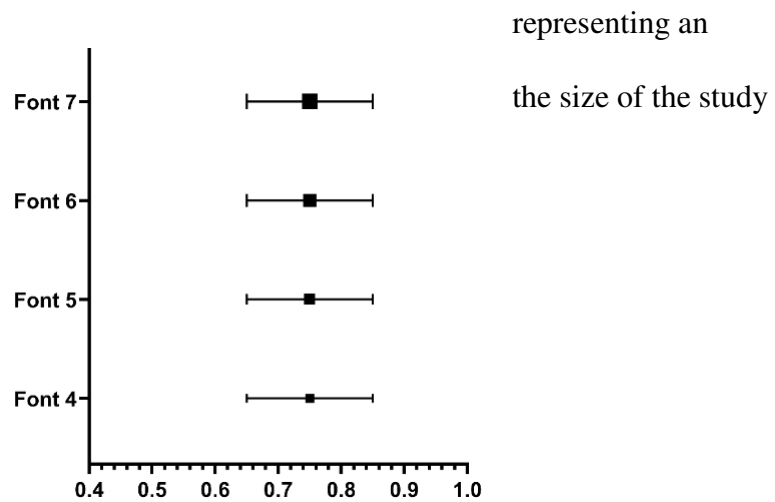
individual study corresponds with population:

Font 7: >10,001

Font 6: 1,001-10,000

Font 5: 100-1,000

Font 4: 0-100



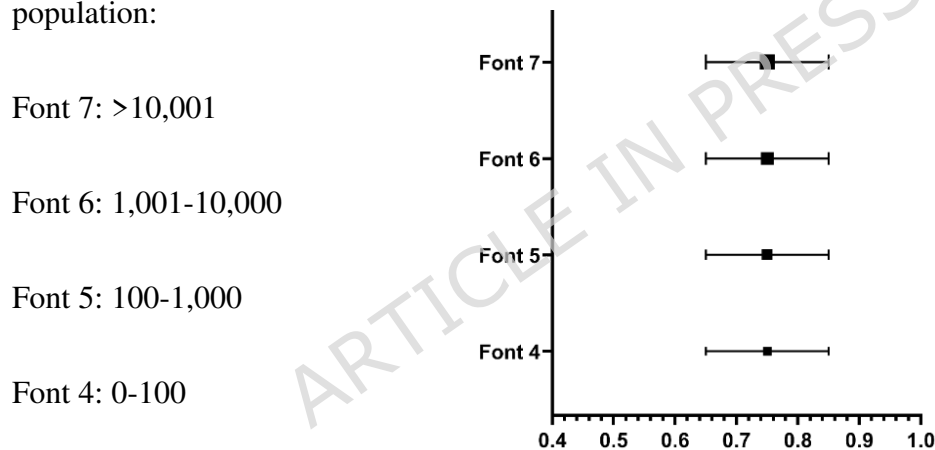
For the reader's ease the data sources are arranged alphabetically within the figure.

Figure 3: Forest plots demonstrating all individual studies contributing to the meta-analyses for (A) IL-6, (B) urea, (C) troponin, and (D) age in Asian and European and Northern American countries for the first tranche of data (from Dec 2019 – June 2021).

Blue diamond: pooled AUC for European and North American countries

Red diamond: pooled AUC for Asian countries

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



For the reader's ease the data sources are arranged alphabetically within the figure.

Figure 4: Panel (A) shows a summary forest plot from the meta-analyses of the first tranche of data (original/wild-type virus) with comparison of pooled area under curves for the five biomarkers being meta-analyzed (CRP, D-dimer, troponin, urea, and IL-6) and age. Panel (B) shows results from the ISARIC4C, GenOMICC and COG-UK consortiums (summarized in figure as 'ISARIC'); who have data relating to intensive care admissions and vaccination status.

Throughout the panels, for ease of comparison, pooled values for Asian countries are shown in red and values for European/Northern American countries are shown in blue.

The size of each symbol representing a patient cohort corresponds with the size of the cohort population:

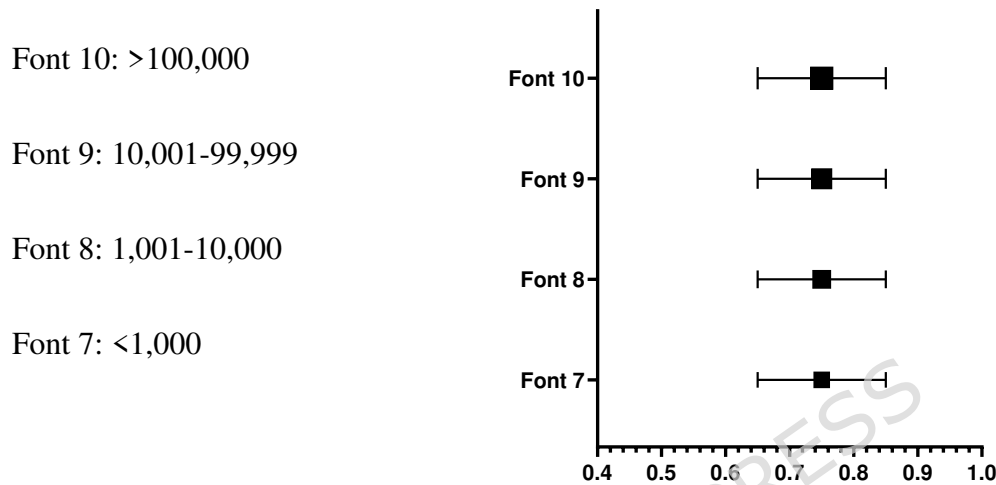


Figure 5: Forest plots demonstrating all individual studies with vaccination and/or variant data for (A) CRP, (B) D-dimer, (C) urea, and (D) age.

Data for the Alpha variant is shown in red.

Data for the Delta variant is shown in magenta.

Data for the Omicron variant is shown in blue.

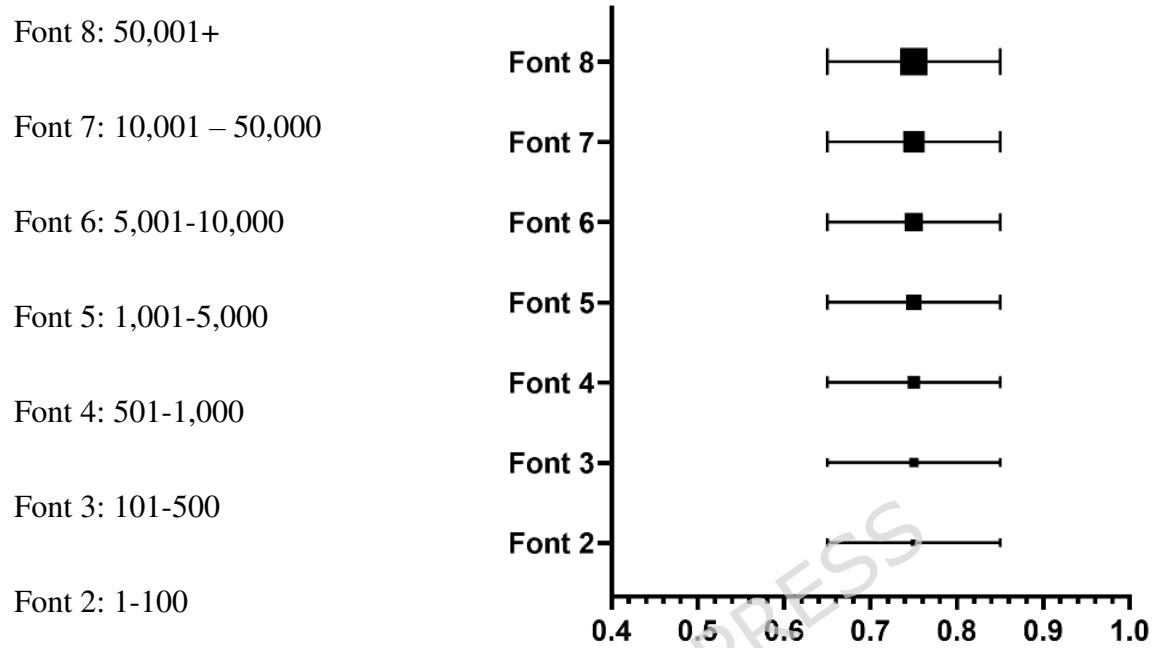
The relevant pooled AUC from the first tranche meta-analyses (Table 1) has been included as a reference (in black).

Squares indicate data from Europe/Northern America and Australia.

Circles indicate data from Asian countries.

*Indicate that the AUCs shown were for severe/critical illness rather than mortality.

The size of each symbol representing a patient cohort corresponds with the size of the cohort population:



Declarations**Ethics approval**

Ethical approval was obtained from the National Health System (UK) Health Research Authority via the Integrated Research Application System (reference 281880) for analysis of the Cambridge (UK) data, in accordance with the Declaration of Helsinki. All the other data has been published and/or is in the public domain.

Consent for publication

Not applicable.

Availability of data and materials

All the root studies are fully cited in the references, supplemental data and at <https://covid19.cimr.cam.ac.uk/>.

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

All authors declare: no support from any commercial organization for the submitted work; no financial relationships with any organizations 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.

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.

The funders did not have a role in the study design, data collection or analyses, or the decision to submit for publication.

Authors Contributions

MERS, NA, FS, WT, and NV collected data, reviewed source manuscripts, and helped analyze data. MS, FS, and SA contributed to the creation of online resources. OF, NA and RJS performed statistical analyses. RT, NWM, SJM, MGS and KB contributed to the data collection, 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.

Acknowledgements

We thank the ISARIC4C Consortium for organising the collection and access to the UK data.

We thank the following clinician-scientists for specifically re-visiting their datasets to provide us with further information regarding vaccination status and viral variants.

Wang J¹, Rzymiski P², Patel A³, Altintop TU⁴, Meletis G⁵, Homen-Fernandez⁶, Jalali A⁷

¹Northern Health, Northern Hospital, 185 Cooper St Epping, Melbourne, VIC, 3076, Australia.

²Department of Environmental Medicine, Poznan University of Medical Sciences, 60-806 Poznań, Poland.

³Infectious Diseases Clinic, Institute of Medical Sciences, Navarangpura, Ahmedabad, 380009, India.

⁴Department of Medical Microbiology, Amasya University Sabuncuoglu Serefeddin Research and Training Hospital, Amasya, Türkiye.

⁵Microbiology Department, AHEPA University Hospital, Thessaloniki, Greece.

⁶Servicio de Enfermedades Infecciosas, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Hospital Clínico San Carlos, Madrid, Spain.

⁷Department of Biology, Faculty of Sciences, Arak University, Postal Code: 38156-8-8349, Arak, Iran.

We thank Ms Natalie Doughty, Mr Chris Davies, Dr Benjamin Dunmore, and Mr Nikita Zubkov who provided public and patient input into accessibility of data, website, and software.

Many thanks to all root study authors, patients and essential workers who have selflessly shared their time and data.