

## Mapping of biomarker efficacy in SARS-CoV-2: tracking the impact of viral mutations and vaccinations

ME Rahman Shuvo<sup>\*,1</sup> MBBS, Max Schwiening<sup>\*,2</sup> BA, Nikos Avramidis BSc<sup>3</sup>, Felipe Soares<sup>4,5</sup> PhD, Oliver Feng<sup>6</sup> PhD, Susana Abreu<sup>2</sup> PhD, Niki Veale<sup>1</sup> MBBS, Gao Q<sup>2</sup> BSc, William Thomas<sup>1</sup> MBBS, AA Roger Thompson<sup>4</sup> PhD, Richard J Samworth<sup>6</sup> PhD, Nicholas W Morrell<sup>1</sup> PhD, Kenneth Baillie PhD<sup>3</sup>, Stefan J Marciniak<sup>1,2</sup> PhD, Elaine Soon<sup>1,2</sup> PhD (0000-0002-5744-5014)

<sup>1</sup>Department of Respiratory Medicine, Level 5, Box 157, Cambridge University Hospitals NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK.

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

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

<sup>4</sup>Department of Infection, Immunity and Cardiovascular Disease, The University of Sheffield, Beech Hill Road, Sheffield S10 2RX.

<sup>5</sup> Universidade Federal do Rio Grande do Sul, Av. Paulo Gama, 110 Secretaria de Comunicação Social - 8º andar - Reitoria - Farroupilha, Porto Alegre - RS, 90040-060, Brazil.

<sup>6</sup>Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WB.

## Collaborators

Wang J<sup>1</sup>, Rzymiski P<sup>2</sup>, Patel A<sup>3</sup>, Altintop TU<sup>4</sup>, Meletis G<sup>5</sup>, Homen-Fernandez<sup>6</sup>, Jalali A<sup>7</sup>

<sup>1</sup>Northern Health, Northern Hospital, 185 Cooper St Epping, Melbourne, VIC, 3076, Australia.

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

<sup>3</sup>Infectious Diseases Clinic, Institute of Medical Sciences, Navarangpura, Ahmedabad, 380009, India.

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

<sup>5</sup>Microbiology Department, AHEPA University Hospital, Thessaloniki, Greece.

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

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

## 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](mailto:eels2@cam.ac.uk) , [e\\_soon@hotmail.co.uk](mailto:e_soon@hotmail.co.uk)

### **Contributorship**

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

### **Acknowledgements**

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

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

## **Copyright for publication**

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to (i) publish, reproduce, distribute, display and store the Contribution, (ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, (iii) create any other derivative work(s) based on the Contribution, (iv) to exploit all subsidiary rights in the Contribution, (v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, (vi) licence any third party to do any or all of the above.

## **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**

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. All the other data has been published and is in the public domain.

## 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.

**Data sharing:** The relevant anonymized data will be shared on reasonable request.

***To the Editor:***

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 antiviral treatment, and no vaccine, resulting in a global death toll of >6.9 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. We noticed that 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). All the models examined had significantly worse validation area under curves (AUCs) than the area under curves of 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). We ran these meta-analyses to identify patterns in biomarker efficacy and to determine the effect of vaccination and virus variants on biomarker efficacy.

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 01<sup>st</sup> December 2019 to 30<sup>th</sup> June 2021. The second search covered 01<sup>st</sup> December 2019 to 31<sup>st</sup> 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'

All papers reporting mortality data for hospitalized patients swab-positive for SARS-CoV-2 with a biomarker level within 48h of admission were examined. Ethical approval was obtained from the Integrated Research Application System (reference 281880) for analysis of the Cambridge (UK) data. We excluded reports of patients already admitted to intensive care or restricted to specific groups (e.g. hemodialysis). Mortality (30-day or in-hospital) was used as the endpoint. This study is registered with PROSPERO (CRD42022366893); and the PRISMA flow diagram is shown in Fig.1(A). Further information is available at <https://covid19.cimr.cam.ac.uk/>.

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 the 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, a *random effects model* was fitted(8).

We examined 1,930 articles that were published from the beginning of the SARS-CoV-2 pandemic on 01<sup>st</sup> December 2019 to 30<sup>th</sup> June 2021 for the baseline study and 2,758 papers from 01<sup>st</sup> December 2019 to 31<sup>st</sup> 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 (Fig.1B). For example, admission CRP levels were 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,  $P < 0.0001$ ). This also held true for admission D-dimer and IL-6 levels. This could explain why the prognostic scores proposed for SARS-CoV-2 performed differently in different countries, as the ‘building blocks’ underpinning these prognostic scores have intrinsically different effectiveness in different populations. Interestingly, troponin and urea levels had universally ‘good’ pooled AUCs. 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.



We expected that multiple rounds of vaccinations and ongoing mutations into different strains would significantly impact biomarker efficacy. The ‘inflammatory’ biomarkers (CRP, D-dimer and IL-6) have generally declined in effectiveness in the vaccinated and variant cohorts (Fig.1C-D). This is particularly illustrated by the Azarfar cohort(8) (Delta variant), who were unvaccinated, with AUCs of 0.576 (0.434-0.713) for CRP and 0.620 (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 results in lower levels of IL-6 production(9, 10). It is also likely that the targeting of inflammation in general (e.g. dexamethasone) and IL-6 pathways in particular (e.g. tocilizumab) has had an effect on biomarker efficacy.

To examine this in greater detail we interrogated data collected from the ISARIC(11), GenOMICC(12) and COG-UK(13) consortiums. Fig.1(E) shows increasing loss of efficacy for CRP from the original meta-analyses to all unvaccinated in the cohort to all vaccinated to intensive care admissions. The AUCs for urea varied less between vaccinated and unvaccinated and had greater prognostic capability in every patient category examined, as did age. We think this shows that co-morbidity is now the leading factor in predicting susceptibility to fatal SARS-CoV-2 in the UK.

We hypothesize that biomarker efficacy in any population at this given time now 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,

and that each country likely has its personalised ‘ingredient list’. From this it follows that prognostic biomarkers/scores should be individualized to particular populations. The simplest and most practical way to achieve this would be to set up information frameworks ahead of time, ideally online. To maximise efficiency, these processes should be automated as much as possible, *e.g.* by using programmes to harvest pre-specified data from medical software systems. Dunning *et al*(14) have proposed a multiple-tier system for this. We propose an even simpler set-up, with basic requirements being age, sex, biomarker levels, date of admission, date of outcome and outcome (death, survival, ITU admission and length of stay). Secondary outcomes (such as imaging data and complications) could be ‘bolt-on’ options. Furthermore, centres could be designated to collect data from specific groups, *e.g.* pregnant and immunocompromised people. 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 meta-analyses. We note it took over a year for data collection for our initial tranche to be completed, by which time SARS-CoV-2 had evolved sufficiently for the original analyses to be outdated. This system would also allow the rapid organization of clinical trials by pinpointing ‘at-risk’ groups who could be targeted early for vaccination or intervention programmes. There are also significant geographical ‘black

boxes' where no published data is available. We would urge large organisations such as the World Health Organisation to identify this as an area requiring further funding.

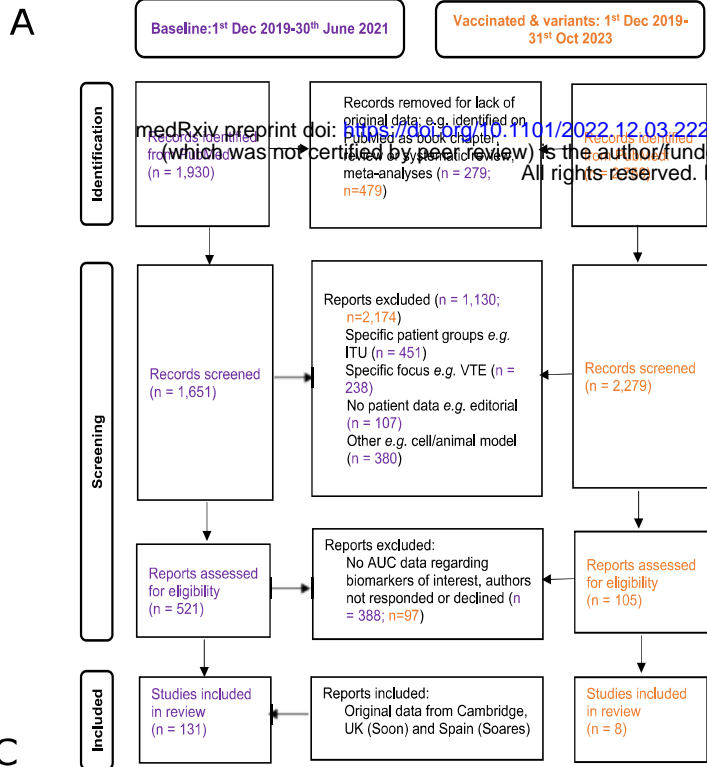
## References (body of manuscript):

1. Yanes-Lane M, Winters N, Fregonese F, Bastos M, Perlman-Arrow S, Campbell JR, Menzies D. Proportion of asymptomatic infection among COVID-19 positive persons and their transmission potential: A systematic review and meta-analysis. *PLoS One*. 2020;15(11):e0241536.
2. Bak A, Muggleston MA, Ratnaraja NV, Wilson JA, Rivett L, Stoneham SM, 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;114:79-103.
3. El-Solh AA, Lawson Y, Carter M, El-Solh DA, Mergenhagen KA. Comparison of in-hospital mortality risk prediction models from COVID-19. *PLoS One*. 2020;15(12):e0244629.
4. Chen R, Liang W, Jiang M, Guan W, Zhan C, Wang T, et al. Risk Factors of Fatal Outcome in Hospitalized Subjects With Coronavirus Disease 2019 From a Nationwide Analysis in China. *Chest*. 2020;158(1):97-105.
5. Shang Y, Liu T, Wei Y, Li J, Shao L, Liu M, et al. Scoring systems for predicting mortality for severe patients with COVID-19. *EClinicalMedicine*. 2020;24:100426.
6. Yu C, Lei Q, Li W, Wang X, Liu W, Fan X, Li W. Clinical Characteristics, Associated Factors, and Predicting COVID-19 Mortality Risk: A Retrospective Study in Wuhan, China. *Am J Prev Med*. 2020;59(2):168-75.
7. Wang K, Zuo P, Liu Y, Zhang M, Zhao X, Xie S, 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):2079-88.

8. Azarfar F, Abbasi B, Jalali A, Abbasian MH. Investigation of the relationship between monocyte chemoattractant protein 1 rs1024611 variant and severity of COVID-19. *Cytokine*. 2023;171:156367.
9. Barh D, Tiwari S, Rodrigues Gomes LG, Ramalho Pinto CH, Andrade BS, Ahmad S, 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):297-312.
10. Korobova ZR, Arsentieva NA, Liubimova NE, Batsunov OK, Dedkov VG, Gladkikh AS, et al. Cytokine Profiling in Different SARS-CoV-2 Genetic Variants. *Int J Mol Sci*. 2022;23(22).
11. Garcia-Gallo E, Merson L, Kennon K, Kelly S, Citarella BW, Fryer DV, et al. ISARIC-COVID-19 dataset: A Prospective, Standardized, Global Dataset of Patients Hospitalized with COVID-19. *Sci Data*. 2022;9(1):454.
12. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. *Nature*. 2021;591(7848):92-8.
13. Wright DW, Harvey WT, Hughes J, Cox M, Peacock TP, Colquhoun R, et al. Tracking SARS-CoV-2 mutations and variants through the COG-UK-Mutation Explorer. *Virus Evol*. 2022;8(1):veac023.
14. Dunning JW, Merson L, Rohde GGU, Gao Z, Semple MG, Tran D, et al. Open source clinical science for emerging infections. *Lancet Infect Dis*. 2014;14(1):8-9.

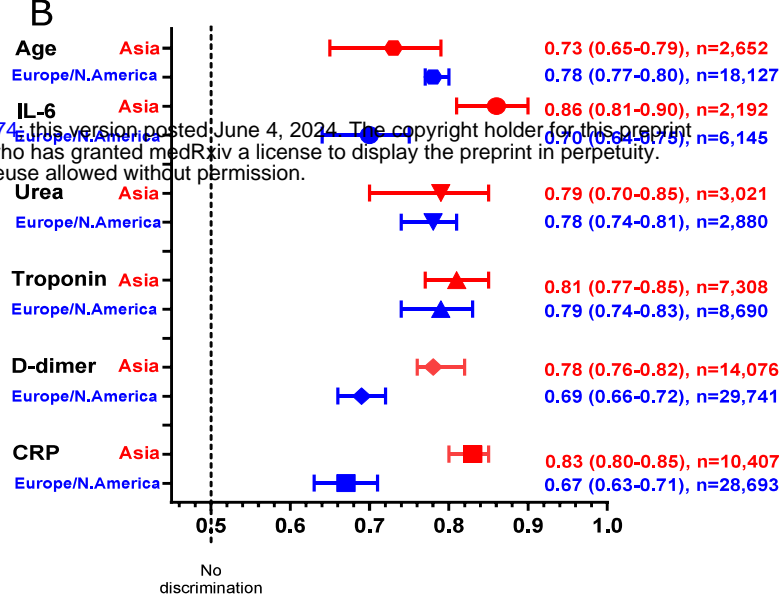
## Figure Legends

Figure 1: Panel (A) shows 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). Panel (B) 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-analysed (CRP, D-dimer, troponin, urea, and IL-6) and age. Panel (C) shows results from variant cohorts, with the relevant pooled AUC from (A) as a reference. \*indicate that the AUCs shown were for severe/critical illness rather than mortality. Panel (D) shows results from vaccinated cohorts, with the relevant pooled AUC from (B) as a reference. Panel (E) shows results from the ISARIC, GenOMICC and COG-UK consortiums; 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/North American countries are shown in blue.



**C**

Source Country	Parameter	2020-mid-2021				
		Pooled AUC Reference	2020	2021	2022	
ISARIC4C UK	AUC CRP	Wild-type	0.67	0.61	0.60	
		non-vaccinated	(0.63-0.71)	(0.61, 0.62)	(0.59, 0.61)	
	AUC CRP	N/A	0.51	0.59		
		fully vaccinated		(0.39,0.63)	(0.58, 0.61)	
	AUC urea	0.78	0.73	0.75		
		non-vaccinated	(0.74-0.81)	(0.72, 0.73)	(0.74,0.76)	
AUC urea	N/A	0.74	0.73			
	fully vaccinated		(0.64,0.84)	(0.72, 0.74)		
Wang Australia	AUC CRP	Wild-type	0.67	0.6	0.55	0.61
		all-cause mortality	(0.63-0.71)*	(0.49, 0.70)	(0.46, 0.64)	(0.51, 0.70)
	AUC CRP	N/A	N/A	0.58	0.61	
		non-vaccinated		(0.49, 0.68)	(0.45, 0.78)	
	AUC CRP	N/A	N/A	N/A	0.57	
		fully vaccinated			(0.44, 0.69)	
	AUC D-dimer	0.69	0.69	0.69	0.63	
		all-cause mortality	(0.66-0.72)*	(0.56, 0.82)	(0.59, 0.78)	(0.52, 0.73)
	AUC D-dimer	N/A	N/A	0.69	0.66	
		non-vaccinated		(0.58, 0.79)	(0.47, 0.85)	
AUC D-dimer	N/A	N/A	0.58	0.57		
	fully vaccinated		(0.14, 1.00)	(0.43, 0.72)		
Rzymiski Poland	AUC CRP	Wild-type	0.67	0.675		
		fully vaccinated	(0.63-0.71)	(0.640 - 0.709)		
	AUC D-dimer	0.69	0.649			
		fully vaccinated	(0.66-0.72)	(0.613 - 0.685)		
AUC IL-6	0.7	0.738				
	fully vaccinated	(0.64-0.75)	(0.695 - 0.778)			
Patel India	AUC CRP	Wild-type	0.83	0.627	0.627	
		non-vaccinated	(0.80-0.85)	(0.389-0.865)	(0.389-0.865)	
	AUC CRP	N/A	0.771			
		fully vaccinated		(0.653-0.889)		
	AUC D-dimer	0.78	0.868			
		non-vaccinated	(0.76-0.82)	(0.715-1.02)		
AUC D-dimer	N/A	0.667				
	fully vaccinated		(0.471-0.868)			
Altintop Turkey	AUC CRP	Wild-type	0.83	0.668		
		non-vaccinated	(0.80-0.85)	(0.592-0.739)		



**D**

Source Country	Parameter	2020-mid-2021			
		Pooled AUC Reference	Alpha	Delta	Omicron
Meletis Greece	AUC CRP	0.67 (0.63-0.71)	0.764	0.767	0.964
	AUC D-dimer	0.69 (0.66-0.72)	0.729	0.800	0.638
	AUC urea	0.77 (0.72-0.82)	0.729	0.638	0.89
	AUC IL-6	0.7 (0.64-0.75)	0.809	0.818	0.883
Homen-Fernandez Spain	AUC CRP	0.67 (0.63-0.71)	0.521	0.586	0.631
	AUC D-dimer	0.69 (0.66-0.72)	0.650	0.596	0.651
	AUC trop	0.81 (0.77-0.85)	0.94	0.788	0.883
	AUC age	0.78 (0.77-0.80)	0.886	0.877	0.749
Jalali Iran	AUC CRP*	0.83 (0.80-0.85)	N/A	0.576	N/A
	AUC D-dimer*	0.78 (0.76-0.82)	N/A	0.620	N/A
Liu China	AUC CRP*	0.83 (0.80-0.85)	N/A	N/A	0.737
	AUC urea*	0.79 (0.70-0.85)	N/A	N/A	0.574

