International Changes in COVID-19 Clinical Trajectories Across 315 Hospitals and 6 Countries: a 4CE Consortium Study

Griffin M Weber, Harrison G Zhang, Sehi L'Yi, Clara-Lea Bonzel, Chuan Hong, Paul Avillach, Alba Gutiérrez-Sacristán, Nathan P Palmer, Amelia Li Min Tan, Xuan Wang, William Yuan, Nils Gehlenborg, Anna Alloni, Danilo F Amendola, Antonio Bellasi, Riccardo Bellazzi, Michele Beraghi, Mauro Bucalo, Luca Chiovato, Kelly Cho, Arianna Dagliati, Hossein Estiri, Robert W Follett, Noelia García Barrio, David A Hanauer, Darren W Henderson, Yuk-Lam Ho, John H Holmes, Meghan R Hutch, Ramakanth Kavuluru, Katie Kirchoff, Jeffrey G Klann, Ashok K Krishnamurthy, Trang T Le, Molei Liu, Ne Hooi Will Loh, Sara Lozano-Zahonero, Yuan Luo, Sarah Maidlow, Adeline Makoudjou, Alberto Malovini, Marcelo Roberto Martins, Bertrand Moal, Michele Morris, Danielle L Mowery, Shawn N Murphy, Antoine Neuraz, Kee Yuan Ngiam, Marina P Okoshi, Gilbert S Omenn, Lav P Patel, Miguel Pedrera Jiménez, Robson A Prudente, Malarkodi Jebathilagam Samayamuthu, Fernando J Sanz Vidorreta, Emily R Schriver, Petra Schubert, Pablo Serrano Balazote, Byorn WL Tan, Suzana E Tanni, Valentina Tibollo, Shyam Visweswaran, Kavishwar B Wagholikar, Zongqi Xia, Daniela Zöller, The Consortium For Clinical Characterization Of COVID-19 By EHR (4CE), Isaac S Kohane, Tianxi Cai, Andrew M South, Gabriel A Brat

Submitted to Journal of Medical Internet Research on: June 23, 2021

Table of Contents

International Changes in COVID-19 Clinical Trajectories Across 315 Hospitals and 6 Countries: a 4CE Consortium Study

Griffin M Weber[1*] MD, PhD; Harrison G Zhang[1*] HSDG; Sehi L'Yi[1*] PhD; Clara-Lea Bonzel[1] MSc; Chuan Hong[1] PhD; Paul Avillach[1] MD, PhD; Alba Gutiérrez-Sacristán[1] PhD; Nathan P Palmer[1] PhD; Amelia Li Min Tan[1] BSc, PhD; Xuan Wang[1] PhD; William Yuan[1] PhD; Nils Gehlenborg[1] PhD; Anna Alloni[2] PhD; Danilo F Amendola[3] MSc; Antonio Bellasi[4] MD, PhD; Riccardo Bellazzi[5] MS, PhD; Michele Beraghi[6] MS; Mauro Bucalo[2] MS; Luca Chiovato[7] MD, PhD; Kelly Cho[8] PhD; Arianna Dagliati[9] MS, PhD; Hossein Estiri[10] PhD; Robert W Follett[11] BS; Noelia García Barrio[12] MS; David A Hanauer[13] MS, MD; Darren W Henderson[14] BS; Yuk-Lam Ho[8] MPH; John H Holmes[15, 16] MS, PhD; Meghan R Hutch[17] BS; Ramakanth Kavuluru[18] PhD; Katie Kirchoff[19] MSHI; Jeffrey G Klann[10] MEng, PhD; Ashok K Krishnamurthy[20] PhD; Trang T Le[15] PhD; Molei Liu[21] MSc; Ne Hooi Will Loh[22] MBBS; Sara Lozano-Zahonero[23] PhD; Yuan Luo[17] PhD; Sarah Maidlow[24] AA; Adeline Makoudjou[23] MD; Alberto Malovini[25] PhD; Marcelo Roberto Martins[26] MSc; Bertrand Moal[27] MD, PhD; Michele Morris[28] BA; Danielle L Mowery[15] PhD; Shawn N Murphy[29] MD, PhD; Antoine Neuraz[30] MD, PhD; Kee Yuan Ngiam[31] MBBS, MRCS, MMed; Marina P Okoshi[32] MD, PhD; Gilbert S Omenn[33] MD, PhD; Lav P Patel[34] MS; Miguel Pedrera Jiménez[12] MS; Robson A Prudente[32] PhD; Malarkodi Jebathilagam Samayamuthu[28] MD; Fernando J Sanz Vidorreta[11] BS; Emily R Schriver[35] MS; Petra Schubert[8] MPH; Pablo Serrano Balazote[12] MS, MD; Byorn WL Tan[36] MBBS; Suzana E Tanni[32] MD, PhD; Valentina Tibollo[25] MS; Shyam Visweswaran[28] MD, PhD; Kavishwar B Wagholikar[10] MBBS, PhD; Zongqi Xia[37] MD, PhD; Daniela Zöller[23] PhD; The Consortium For Clinical Characterization Of COVID-19 By EHR (4CE)[1]; Isaac S Kohane[1] MD, PhD; Tianxi Cai[1] ScD; Andrew M South[38] MS, MD; Gabriel A Brat[1] MPH, MD

[1]Department of Biomedical Informatics Harvard Medical School Boston US
[2]BIOMERIS (BIOMedical Research Informatics Solutions) Pavia IT
[3]Clinical Research Unit Botucatu Medical School São Paulo State University Botucatu BR
[4]Division of Nephrology Department of Medicine Ente Ospedaliero Cantonale Lugano CH
[5]Department of Electrical, Computer and Biomedical Engineering University of Pavia Pavia IT
[6]Information Technology Department Azienda Socio-Sanitaria Territoriale di Pavia Pavia IT
[7]Unit of Internal Medicine and Endocrinology Istituti Clinici Scientifici Maugeri SpA SB IRCCS Pavia IT
[8]Massachusetts Veterans Epidemiology Research and Information Center Veterans Affairs Boston Healthcare System Boston US
[9]Department of Electrical Computer and Biomedical Engineering University of Pavia Pavia IT
[10]Department of Medicine Massachusetts General Hospital Boston US
[11]Department of Medicine David Geffen School of Medicine University of California, Los Angeles Los Angeles US
[12]Health Informatics Hospital Universitario 12 de Octubre Madrid ES
[13]Department of Learning Health Sciences University of Michigan Medical School Ann Arbor US
[14]Department of Biomedical Informatics University of Kentucky Lexington US
[15]Department of Biostatistics, Epidemiology, and Informatics University of Pennsylvania Perelman School of Medicine Philadelphia US
[16]Institute for Biomedical Informatics University of Pennsylvania Perelman School of Medicine Philadelphia US
[17]Department of Preventive Medicine Northwestern University Chicago US
[18]Institute for Biomedical Informatics University of Kentucky Lexington US
[19]Medical University of South Carolina Charleston US
[20]Department of Computer Science Renaissance Computing Institute University of North Carolina at Chapel Hill Chapel Hill US
[21]Department of Biostatistics Harvard T.H. Chan School of Public Health Boston US
[22]Department of Anaesthesia National University Health System Singapore SG
[23]Institute of Medical Biometry and Statistics Faculty of Medicine and Medical Center University of Freiburg Freiburg DE
[24]Michigan Institute for Clinical & Health Research Informatics University of Michigan Ann Arbor US
[25]Laboratory of Informatics and Systems Engineering for Clinical Research Istituti Clinici Scientifici Maugeri SpA SB IRCCS Pavia IT
[26]Clinical Hospital of Botucatu Medical School São Paulo State University Botucatu BR
[27]Informatique et archivistique médicales unit Bordeaux University Hospital Bordeaux FR
[28]Department of Biomedical Informatics University of Pittsburgh Pittsburgh US
[29]Department of Neurology Massachusetts General Hospital Boston US
[30]Department of Biomedical Informatics Hôpital Necker-Enfants Malade, Assistance Publique Hôpitaux de Paris University of Paris Paris FR
[31]Department of Biomedical Informatics, Institute for Digital Medicine National University Health System Singapore SG
[32]Internal Medicine Department Botucatu Medical School São Paulo State University Botucatu BR
[33]Department of Computational Medicine & Bioinformatics, Internal Medicine, Human Genetics, and Public Health University of Michigan Ann

Arbor US
[34]Division of Medical Informatics Department of Internal Medicine University of Kansas Medical Center Kansas City US
[35]Data Analytics Center University of Pennsylvania Health System Philadelphia US
[36]Department of Medicine National University Health System Singapore SG
[37]Department of Neurology University of Pittsburgh Pittsburgh US
[38]Section of Nephrology Department of Pediatrics Brenner Children's Hospital, Wake Forest School of Medicine Winston Salem US
[*]these authors contributed equally

Corresponding Author:
Tianxi Cai ScD
Department of Biomedical Informatics
Harvard Medical School
10 Shattuck St
Boston
US

Abstract

Background: Many countries have experienced two predominant waves of COVID-19-related hospitalizations. Comparing the clinical trajectories of patients hospitalized in separate waves of the pandemic enables further understanding of the evolving epidemiology, pathophysiology, and healthcare dynamics of the COVID-19 pandemic.

Objective: In this retrospective cohort study, we analyzed electronic health record (EHR) data from patients with SARS-CoV-2 infections hospitalized in participating healthcare systems representing 315 hospitals across six countries. We compared hospitalization rates, severe COVID-19 risk, and mean laboratory values between patients hospitalized during the first and second waves of the pandemic.

Methods: Using a federated approach, each participating healthcare system extracted patient-level clinical data on their first and second wave cohorts and submitted aggregated data to the central site. Data quality control steps were performed at the central site to correct for implausible values and harmonize units. Statistical analyses were performed by computing individual healthcare system effect sizes and synthesizing these using random effects meta-analyses to account for heterogeneity. We focused the laboratory analysis on C-reactive protein (CRP), ferritin, fibrinogen, procalcitonin, D-dimer, and creatinine based on their reported associations with severe COVID-19.

Results: Data were available for 79,487 patients, of which 32,452 were hospitalized in the first wave and 47,035 in the second wave. The prevalence of male patients and patients aged 50–69 decreased significantly between the first and second wave. Patients hospitalized in the second wave had a 9.6% reduction in risk of severe COVID-19 compared to patients hospitalized in the first wave (95% CI: 8.2–11.1%). Demographic subgroup analyses indicated that patients aged 26–49; male and female patients; and Black patients had significantly lower risk for severe disease in the second wave compared to the first wave. At admission, the mean values of CRP were significantly lower in the second wave compared to the first. On the seventh hospital day, mean values of CRP, ferritin, fibrinogen, procalcitonin, and creatinine were significantly lower in the second wave compared to the first. In general, countries exhibited variable changes in laboratory testing rates from the first to the second wave. At admission, there was a significantly higher testing rate for D-dimer in France, Germany, and Spain.

Conclusions: Patients hospitalized in the second wave were at significantly lower risk for severe COVID-19. This corresponded to mean laboratory values in the second wave that were more likely to be in typical physiological ranges on the seventh hospital day compared to the first wave. Our federated approach demonstrated the feasibility and power of harmonizing heterogeneous EHR data from multiple international healthcare systems to rapidly conduct large-scale studies to characterize how COVID-19 clinical trajectories evolve.

(JMIR Preprints 23/06/2021:31400)
DOI: https://doi.org/10.2196/preprints.31400

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

International Changes in COVID-19 Clinical Trajectories Across 315 Hospitals and 6 Countries: a 4CE Consortium Study

Griffin M Weber MD, PhD[1]*, Harrison G Zhang[1]*, Sehi L'Yi PhD[1]*,Clara-Lea Bonzel MSc,[1] Chuan Hong PhD[1], Paul Avillach MD, PhD[1], Alba Gutiérrez-Sacristán PhD[1], Nathan P Palmer, PhD[1], Amelia LM Tan BSc, PhD[1], Xuan Wang PhD[1], William Yuan PhD[1], Nils Gehlenborg PhD[1], Anna Alloni PhD[2], Danilo F Amendola MSc[3], Antonio Bellasi MD, PhD[4], Riccardo Bellazzi MS, PhD[5], Michele Beraghi[6], Mauro Bucalo MS[2], Luca Chiovato MD, PhD[7], Kelly Cho PhD[8], Arianna Dagliati MS, PhD[5], Hossein Estiri PhD[9], Robert W Follett[10], Noelia García-Barrio MS[11], David A Hanauer MD, MS[12], Darren W Henderson[13], Yuk-Lam Ho MPH[8], John H Holmes MS, PhD[14,15], Meghan R Hutch BS[16], Ramakanth Kavuluru PhD[17], Katie Kirchoff MSHI[18], Jeffrey G Klann MEng, PhD[9], Ashok K Krishnamurthy PhD[19], Trang T Le PhD[14], Molei Liu MSc[20], Ne Hooi Will Loh MBBS, FFICM[21], Sara Lozano-Zahonero PhD[22], Yuan Luo PhD[16], Sarah Maidlow[23], Adeline Makoudjou MD[22], Alberto Malovini PhD[24], Bertrand Moal MD, PhD[25], Michele Morris BA[26], Danielle L Mowery PhD[14], Shawn N Murphy MD, PhD[27], Antoine Neuraz MD, PhD[28], Kee Yuan Ngiam MBBS, MRCS, MMed, FRCS[29], Marina P Okoshi MD, PhD[30], Gilbert S Omenn MD, PhD[31], Lav P Patel MS[32], Miguel Pedrera-Jiménez MS[11], Robson A Prudente PhD[30], Malarkodi J Samayamuthu MD[26], Javier Sanz BS[10], Emily R Schriver MS[33], Petra Schubert MPH[8], Pablo Serrano-Balazote MD,MS[11], Byorn WL Tan MBBS[34], Suzana E Tanni MD, PhD[30], Valentina Tibollo MS[24], Shyam Visweswaran MD, PhD[26], Kavishwar B Wagholikar MBBS, PhD[9], Zongqi Xia MD, PhD[35], Daniela Zoeller PhD[22], The Consortium for Clinical Characterization of COVID-19 by EHR (4CE)[1], Isaac S Kohane MD, PhD[1], Tianxi Cai ScD[1†‡], Andrew M South MD, MS[36†], Gabriel A Brat MD, MPH[1†]

* Weber, Zhang, and L'Yi contributed equally
† Cai, South, and Brat contributed equally
‡Email: tcai@hsph.harvard.edu

1. Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA
2. BIOMERIS (BIOMedical Research Informatics Solutions), Pavia, Italy
3. Clinical Research Unit of Botucatu Medical School, São Paulo State University, Botucatu, Brazil
4. Department of Medicine, Division of Nephrology, Ente Ospedaliero Cantonale, Lugano, Switzerland
5. Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy
6. IT Department, ASST Pavia
7. Unit of Internal Medicine and Endocrinology, Istituti Clinici Scientifici Maugeri SpA SB IRCCS, Pavia, Italy

8. Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), VA Boston Healthcare System
9. Department of Medicine, Massachusetts General Hospital, Boston, MA, USA
10. Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
11. Health Informatics, Hospital Universitario 12 de Octubre, Madrid, Spain
12. Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, Michigan, USA
13. Department of Biomedical Informatics, University of Kentucky, Lexington, KY, USA
14. Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
15. Institute for Biomedical Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
16. Department of Preventive Medicine, Northwestern University, Chicago, IL, USA
17. Institute for Biomedical Informatics, University of Kentucky, Lexington, KY, USA
18. Medical University of South Carolina, Charleston, SC, USA
19. Renaissance Computing Institute/Department of Computer Science, University of North Carolina Chapel Hill, Chapel Hill, NC, USA
20. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA
21. Department of Anaesthesia, National University Health System Singapore, Singapore
22. Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg, Germany
23. MICHR Informatics, University of Michigan, Ann Arbor, Michigan, USA
24. Laboratory of Informatics and Systems Engineering for Clinical Research, Istituti Clinici Scientifici Maugeri SpA SB IRCCS, Pavia, Italy
25. IAM Unit, Bordeaux University Hospital, Bordeaux, France
26. Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, USA
27. Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
28. Department of Biomedical Informatics, Hôpital Necker-Enfants Malades, Assistance Publique Hôpitaux de Paris (APHP), University of Paris, Paris, France
29. Department of Biomedical informatics, WiSDM, National University Health System Singapore
30. Internal Medicine Department of Botucatu Medical School, São Paulo State University, Botucatu, Brazil
31. Dept of Computational Medicine & Bioinformatics, Internal Med, Human Genetics, and Public Health, University of Michigan, Ann Arbor, MI, USA
32. Department of Internal Medicine, Division of Medical Informatics, University Of Kansas Medical Center, Kansas City, KS, USA
33. Data Analytics Center, University of Pennsylvania Health System, Philadelphia, PA, USA
34. Department of Medicine, National University Hospital, Singapore
35. Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA
36. Department of Pediatrics-Section of Nephrology, Brenner Children's Hospital, Wake Forest School of Medicine, Winston Salem, NC, USA

Abstract

Background: Many countries have experienced two predominant waves of COVID-19-related hospitalizations. Comparing the clinical trajectories of patients hospitalized in separate waves of the pandemic enables further understanding of the evolving epidemiology, pathophysiology, and healthcare dynamics of the COVID-19 pandemic.

Objective: In this retrospective cohort study, we analyzed electronic health record (EHR) data from patients with SARS-CoV-2 infections hospitalized in participating healthcare systems representing 315 hospitals across six countries. We compared hospitalization rates, severe COVID-19 risk, and mean laboratory values between patients hospitalized during the first and second waves of the pandemic.

Methods: Using a federated approach, each participating healthcare system extracted patient-level clinical data on their first and second wave cohorts and submitted aggregated data to the central site. Data quality control steps were performed at the central site to correct for implausible values and harmonize units. Statistical analyses were performed by computing individual healthcare system effect sizes and synthesizing these using random effects meta-analyses to account for heterogeneity. We focused the laboratory analysis on C-reactive protein (CRP), ferritin, fibrinogen, procalcitonin, D-dimer, and creatinine based on their reported associations with severe COVID-19.

Results: Data were available for 79,613 patients, of which 32,467 were hospitalized in the first wave and 47,146 in the second wave. The prevalence of male patients and patients aged 50–69 decreased significantly between the first and second wave. Patients hospitalized in the second wave had a 9.9% reduction in risk of severe COVID-19 compared to patients hospitalized in the first wave (95% CI: 8.5–11.3%). Demographic subgroup analyses indicated that patients aged 26–49 and 50–69; male and female patients; and Black patients had significantly lower risk for severe disease in the second wave compared to the first wave. At admission, the mean values of CRP were significantly lower in the second wave compared to the first. On the seventh hospital day, mean values of CRP, ferritin, fibrinogen, and procalcitonin were significantly lower in the second wave compared to the first. In general, countries exhibited variable changes in laboratory testing rates from the first to the second wave. At admission, there was a significantly higher testing rate for D-dimer in France, Germany, and Spain.

Conclusions: Patients hospitalized in the second wave were at significantly lower risk for severe COVID-19. This corresponded to mean laboratory values in the second wave that were more likely to be in typical physiological ranges on the seventh hospital day compared to the first wave. Our federated approach demonstrated the feasibility and power of harmonizing heterogeneous EHR data from multiple international healthcare systems to rapidly conduct large-scale studies to characterize how COVID-19 clinical trajectories evolve.

Keywords: SARS-CoV-2; electronic health records; federated study; retrospective cohort study; meta-analysis; severe COVID-19; laboratory trajectory

Introduction

From January 2020 to June 2021, the coronavirus disease 2019 (COVID-19) pandemic has resulted in over 170 million confirmed cases of SARS-CoV-2 infection and 3.7 million confirmed deaths worldwide [1]. Similar to previous viral pandemics, the resurgence in SARS-CoV-2 infections and subsequent hospitalizations since the first documented outbreaks have been characterized by a series of "waves." To date, there have been reports of at least two waves in numerous countries: an initial one in the Spring of 2020 and a resurgence of cases in the Summer and Fall of 2020 [1–10]. A limited number of single-center studies have reported differences in laboratory values, demographic composition, and disease management between patients with COVID-19 admitted in the first and second waves [6,8,11,12]. Thus, there is substantial interest in comparing the clinical trajectories of patients with SARS-CoV-2 who were hospitalized across different waves of the pandemic to better understand the rapidly evolving epidemiology, pathophysiology, and healthcare dynamics of the COVID-19 pandemic. This may further inform healthcare workers, policymakers, and public health experts on how to anticipate potential additional waves due to SARS-CoV-2 variants [13].

Single-center studies are limited in scope, power, and generalizability, and there is a need for robust, multicenter analyses using multinational cohorts that compare first and second wave patient characteristics. The goal of this study was to use a federated electronic health record (EHR)-based approach to examine international temporal trends in the clinical trajectories of patients hospitalized with SARS-CoV-2 across 6 countries obtained from contributing healthcare systems in the Consortium for Clinical Characterization of COVID-19 EHR (4CE) [14], an international research collaborative of more than 300 hospitals across seven countries that collects patient-level EHR data to study the epidemiology and clinical course of COVID-19. We collected data from 26 participating international healthcare systems covering 79,613 hospitalized patients with SARS-CoV-2 to study changes in (1) hospitalization rates across calendar time; (2) risk of developing severe COVID-19; and (3) mean laboratory values and laboratory testing rates between the first and second waves. We stratified severity risk analyses by country and demographic subgroups.

Methods

Description of the Federated Approach: Participating Healthcare Systems, Local Data Collection, and Central Data Aggregation

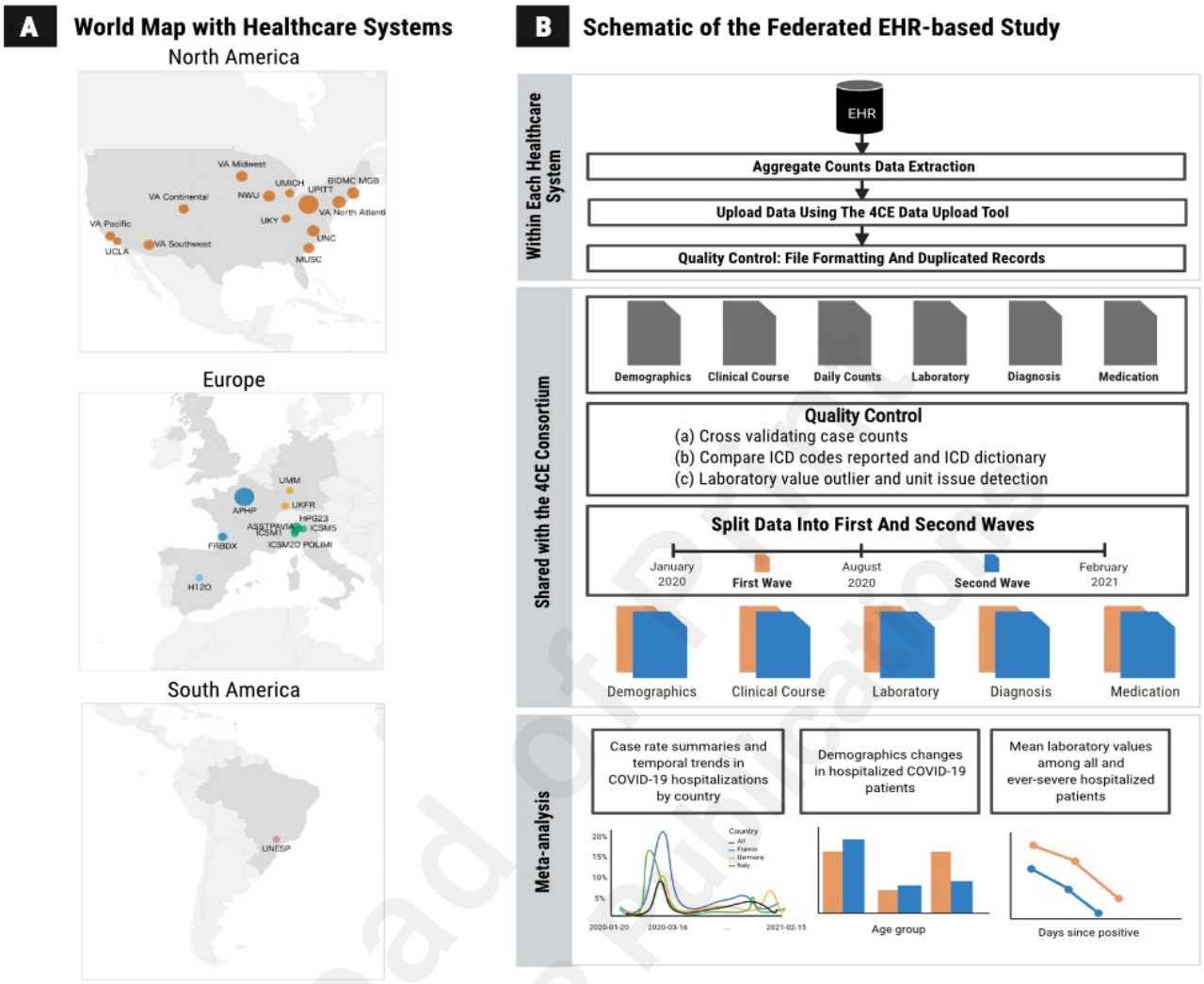
Our analyses were performed on EHR data collected from 315 hospitals (affiliated with 26 regional healthcare systems) across six countries: Brazil, France, Germany, Italy, Spain, and the United States [14-15]. In the United States, we grouped the 170 Veterans Affairs hospitals into 5 regional healthcare systems [16]. See Table 1 for details about participating healthcare systems and Figure 1 for a map with participating healthcare systems.

Similar to our previous 4CE studies, we distributed a SQL database script to each contributing healthcare system, which they ran on their patient-level EHR data to generate aggregate counts and statistics about their patient cohorts after gaining IRB approval [14,15,17]. Healthcare systems then uploaded their

aggregate data via a central 4CE Data Upload Tool. Aggregate data included hospital admission summaries over calendar time; date of positive SARS-CoV-2 reverse transcription polymerase chain reaction test; demographic counts for age, sex, and race groups; and daily trajectories of laboratory test values. Race data were only reported in participating healthcare systems from the USA and only included categories for Black and White patients given the considerable heterogeneity present in race coding systems across healthcare system EHRs. A schematic of our workflow is presented in Figure 1, and further details of collected data are reported in Supplementary Table 1.

In order to ensure high-quality EHR data across healthcare systems and countries, extensive data quality control was performed. When sites uploaded the data using the 4CE Data Upload Tool, an initial online quality control verified that all Comma Separated Value (CSV) files were under the standard format, which included verification of the file and column names, column orders, data types, code values and ranges, and the absence of duplicated records. This step was crucial in ensuring proper downstream statistical analysis. At the central site, additional quality control steps were completed on all submitted data. These steps included cross-validating consistency of the total case counts, checking that there are no negative values in patient counts, and verifying that no data types were missing. We also checked for consistency between the 3-digit ICD codes and the ICD dictionary. To assess the general consistency of the laboratory data from each site with data observed from all healthcare systems and to detect outliers, we plotted laboratory values across time with all sites overlaid on each other. Using these plots, we were also able to check if certain laboratory values from specific healthcare systems were systematically lower or higher than for other sites, which usually indicated unit errors. If a healthcare system presented any quality control issues, the central site directly contacted its corresponding informaticians to resolve them.

Figure 1: Schematic of the federated EHR-based study involving healthcare systems from six countries.



Cohort Identification

Our study included all patients hospitalized up to February 28, 2021 at participating 4CE healthcare systems with an admission date between 7 days before to 14 days after the date of their first positive SARS-CoV-2 reverse transcription polymerase chain reaction test result. We chose this time window in an effort to mitigate selection bias by identifying hospitalized patients who may have tested positive for SARS-CoV-2 before or after being admitted to a hospital. We further defined the first admission date within this -7 to +14-day time window as the index date, and “days since admission” and “hospital day” are referenced to this index date. EHR data were available for 79,613 patients.

We partitioned patients into first and second-wave cohorts according to their index date. Although different regions had slightly varying temporal trajectories in COVID-19-related hospitalizations, our data indicated two predominant waves of hospitalizations with which we used to partition patients: a first wave from January 1st to July 31st of 2020 and a second wave from August 1st of 2020 to February 28th of 2021. Relatively few admissions occurred between July and August 2020 across all healthcare systems. Table 1 defines healthcare system-specific hospitalization date criteria for the first and second wave.

We further categorized patients as “ever-severe” using the validated 4CE COVID-19 severity algorithm

that allows us to determine whether patients, at any time during their hospitalization, progressed to severe disease, regardless of their recovery from COVID-19 [18]. The algorithm leverages a set of EHR data elements to define severe COVID-19 including: (1) laboratory tests for partial pressure of carbon dioxide or partial pressure of oxygen; (2) ordered medications for sedatives and anesthetics; (3) diagnosis codes for ARDS or ventilator-associated pneumonia; and (4) procedures such as endotracheal tube insertion or invasive mechanical ventilation [18].

Table 1. Participating healthcare systems, metadata on number of hospitals and beds, and hospitalization date used to define the first and second wave cohorts.

Healthcare system	Country	Hospitals	Beds	Inpatient discharges/year	First wave date range	Second wave date range	First wave sample size	Second sample
Assistance Publique - Hôpitaux de Paris	France	39	20,098	1,375,538	2020-01-29 2020-08-10	2020-08-11 2020-11-06	9,827	4,584
ASST della provincia di Pavia	Italy	7	958	29,103	2020-02-28 2020-04-30	2020-05-01 2021-02-15	945	1,543
Beth Israel Deaconess Medical Center	USA	1	673	40,752	2020-03-23 2020-09-02	2020-09-03 2021-02-22	685	585
Bordeaux University Hospital	France	3	2,676	130,033	2020-01-23 2020-07-31	2020-08-01 2020-11-07	331	439
Hospital Universitario 12 de Octubre	Spain	1	1,256	45,035	2020-03-01 2020-07-20	2020-07-21 2021-02-28	2,369	3,730
ASST Papa Giovanni XXIII Bergamo	Italy	1	1,080	45,000	2020-02-25 2020-05-21	2020-05-22 2020-11-13	1,533	371
ICSM Pavia Hospital	Italy	1	426	8,616	2020-02-29 2020-05-08	2020-05-09 2020-11-16	113	58
ICSM Milano Hospital	Italy	1	200	2,432	2020-02-21 2020-05-08	2020-05-09 2020-11-16	38	119
ICSM Lumezzane/Brescia Hospitals	Italy	1	149	1,296	2020-03-11 2020-05-08	2020-05-09 2020-11-16	111	21
Mass General Brigham (Partners Healthcare)	USA	10	3,418	163,521	2020-03-11 2020-07-31	2020-08-01 2021-02-28	2,736	1,735
Medical University of South Carolina	USA	8	1,600	55,664	2020-03-12 2020-05-25	2020-05-26 2020-11-15	127	1,482
Northwestern University	USA	10	2,234	103,279	2020-03-05 2020-07-31	2020-08-01 2020-12-31	2,313	3,567
Policlinico di Milano	Italy	1	900	40,000	2020-02-25 2020-08-01	2020-08-02 2020-11-13	612	304
Medical Center, University of Freiburg	Germany	1	1,660	71,500	2020-03-13 2020-07-31	2020-08-01 2021-02-28	186	490
University of California, LA	USA	2	786	40,526	2020-03-10 2020-08-03	2020-08-04 2020-11-13	425	151
University of Kentucky	USA	3	881	45,714	2020-03-18 2020-07-07	2020-07-08 2020-11-06	113	352
University of Michigan	USA	3	1,000	49,008	2020-03-09 2020-07-31	2020-08-01 2021-02-28	745	1,619
University Medicine	Germany	1	1,352	50,748	2020-03-18	2020-08-04	81	497

Mannheim					2020-08-03	2021-01-23		
University of North Carolina at Chapel Hill	USA	11	3,095	52,000	2020-03-14 2020-06-05	2020-06-06 2020-10-30	458	1,525
Universidade Estadual Julio de Mesquita Filho	Brazil	1	490	28,167	2020-04-01 2020-07-31	2020-08-01 2021-02-28	171	425
University of Pittsburgh	USA	39	8,085	369,300	2020-03-13 2020-07-31	2020-08-01 2021-02-28	685	5,021
VA North Atlantic	USA	49	3,594	151,075	2020-03-01 2020-07-31	2020-08-01 2021-02-04	1,949	2,984
VA Southwest	USA	29	3,115	156,315	2020-03-01 2020-07-31	2020-08-01 2021-02-04	1,679	4,071
VA Midwest	USA	39	2,686	145,468	2020-03-01 2020-07-31	2020-08-01 2021-02-04	1,544	4,617
VA Continental	USA	24	2,110	113,260	2020-03-01 2020-07-31	2020-08-01 2021-02-04	1,497	3,495
VA Pacific	USA	29	2,296	114,569	2020-03-01 2020-07-31	2020-08-01 2021-02-04	1,194	3,361
	Total	315	66,818	3,427,919			32,467	47,146

Statistical Analysis

Centralized random effects meta-analyses were performed to summarize individual healthcare system effect sizes. To account for heterogeneity between healthcare systems, we harmonized effect sizes using DerSimonian and Laird random effects meta-analysis [19]. Weights assigned to healthcare system effect sizes during meta-analysis were kept constant between corresponding first and second cohort analyses to facilitate effective comparisons between waves. All statistical analyses were performed using R software version 4.0.2.

We estimated the intensity rate of hospitalizations over time within each participating healthcare system and averaged at the country level. Within each healthcare system, the intensity rate for a given calendar date was estimated as the proportion of patients in the cohort who were hospitalized on that date. We further summarized the prevalence of demographic subgroups in the first and second wave. We excluded the VA healthcare systems only when estimating the prevalence of demographic subgroups in our cohort due to their unique demographic profiles [20]. We report the prevalence of demographic subgroups including the VA healthcare systems in the Supplementary Appendix.

We then estimated the absolute risk of severe COVID-19 in the first and second waves and the relative risk of severe COVID-19 in the second wave compared to the first wave. Within each healthcare system and over a set time period of interest, the absolute risk was estimated as the proportion of patients who ever developed severe disease among all patients in the corresponding cohort. We stratified these analyses by country and demographic subgroups. Analyses of absolute risk and relative risk for severe COVID-19 included all participating healthcare systems.

We then compared standardized mean laboratory test values stratified by disease severity at days 0, 1, and

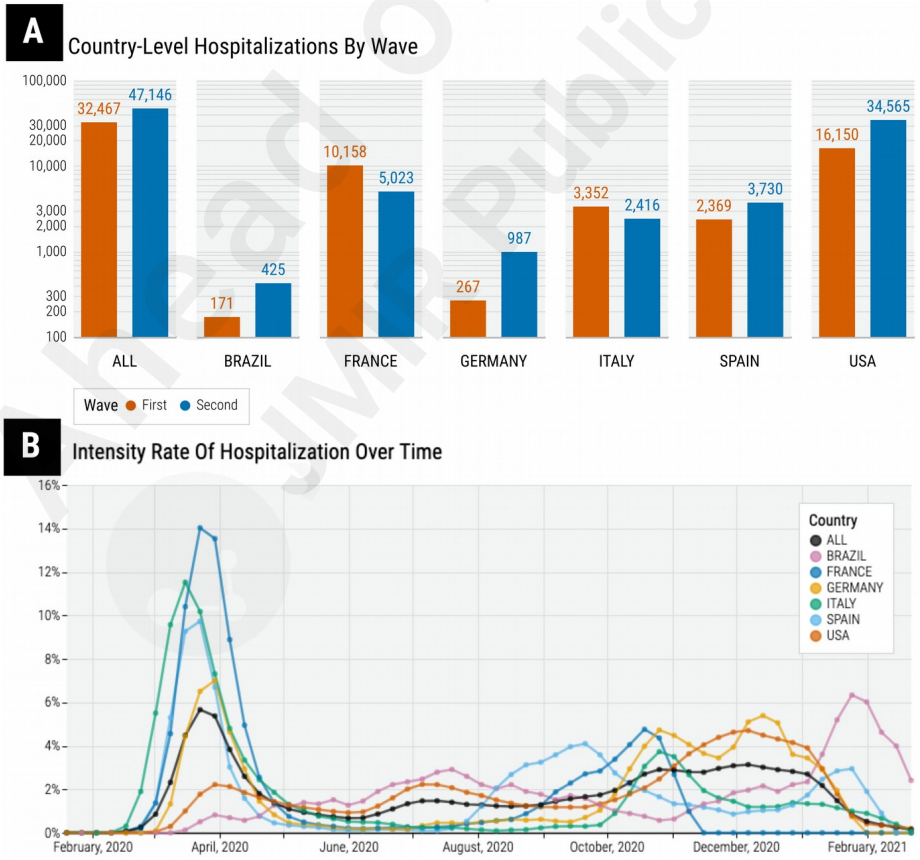
7 to investigate changes in laboratory trajectories between the two waves. We focused on six blood laboratory values associated with worse outcomes and severe disease in patients with COVID-19: C-reactive protein (CRP), ferritin, fibrinogen, procalcitonin, D-dimer, and creatinine [21–28]. To facilitate effective comparisons, we defined standardized laboratory values as relative to each laboratory test's harmonized value on the index date in the first wave. We also summarized the proportion of all and ever-severe patients having each laboratory test at day 0–14 stratified by country to examine any changes in clinical practice regarding laboratory testing.

Results

Characteristics of the Study Population and Trends in Hospitalization

In the study population of 79,613 hospitalized patients with SARS-CoV-2, 32,467 were hospitalized during the first wave and 47,146 were hospitalized during the second wave. In this cohort, the USA represented the country with the most hospitalizations. As seen in Figure 2B, hospitalization rates generally peaked in March–April of 2020 and again in the final months of 2020 across all six countries.

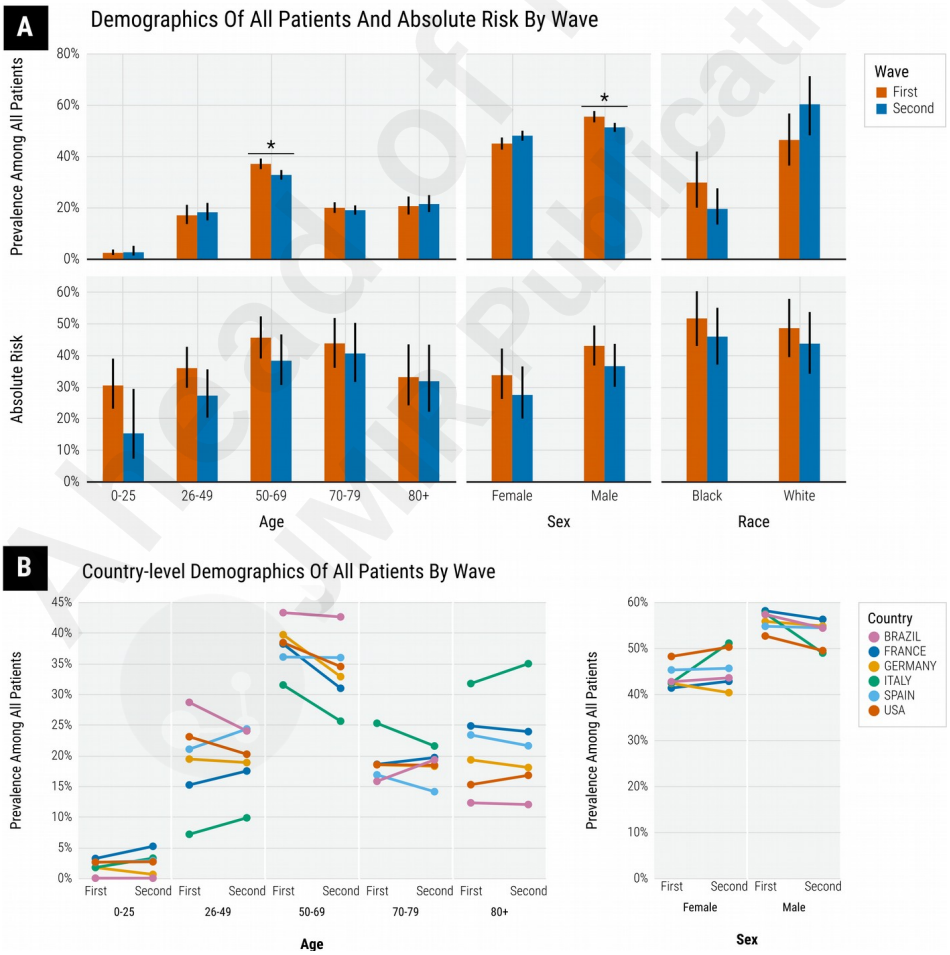
Figure 2: Total hospitalizations in the cohort between first and second waves and intensity rate of hospitalizations over time by country.



We report the prevalence of demographic subgroups in Figure 3. Overall, there was a higher prevalence of male and older age patients in both waves. The prevalence of patients aged 50–69 decreased significantly from the first wave (37.1%, 95% CI: 35.0–39.3%) to the second wave (32.3%, 95% CI: 30.5–34.2%). The prevalence of male patients also decreased from the first wave (55.3%, 95% CI: 53.1%–57.5%) to the

second wave (50.9%, 95% CI: 49.1%–52.6%). There were no statistically significant changes in the prevalence of other age or sex groups for the entire cohort. At the country level, we observed that in Spain, the prevalence of patients aged 26–49 increased significantly from 21.0% (95% CI: 19.4–22.7%) in the first wave to 24.3% (95% CI: 23.0–25.7%) in the second wave, while the prevalence of patients aged 70–79 decreased significantly from 16.8% (95% CI: 15.4–18.4%) in the first wave to 14.1% (95% CI: 13.0–15.3%) in the second wave. In the USA, the prevalence of White patients increased (first wave: 46.3%, 95% CI: 36.3–56.6% vs. second wave: 60.2%, 95% CI: 48.1–71.1%), while the prevalence of Black patients decreased (first wave: 29.7%, 95% CI: 20.0–41.8% vs. second wave: 19.4%, 13.3%–27.5%), although both results did not reach statistical significance.

Figure 3: Prevalence of demographic subgroups in the first and second waves for the entire cohort and by country, and estimated absolute risk for severe COVID-19 by demographic subgroup. Error bars indicate 95% confidence intervals. *Indicates a significant difference by non-overlapping 95% confidence intervals.



Risk of Severe Disease in the First and Second Waves

We report the absolute risk and relative risk for severe COVID-19 stratified by country in Figure 4. The absolute risk of severe disease was 0.40 (95% CI: 0.34–0.48) in the first wave and 0.33 (95% CI: 0.25–0.43) in the second wave. The absolute risk for severe disease varied significantly across countries in both waves [first wave vs. second wave]: (1) Brazil: 30.1% vs. 8.7%; (2) France: 66.7% vs. 60.1%; (3)

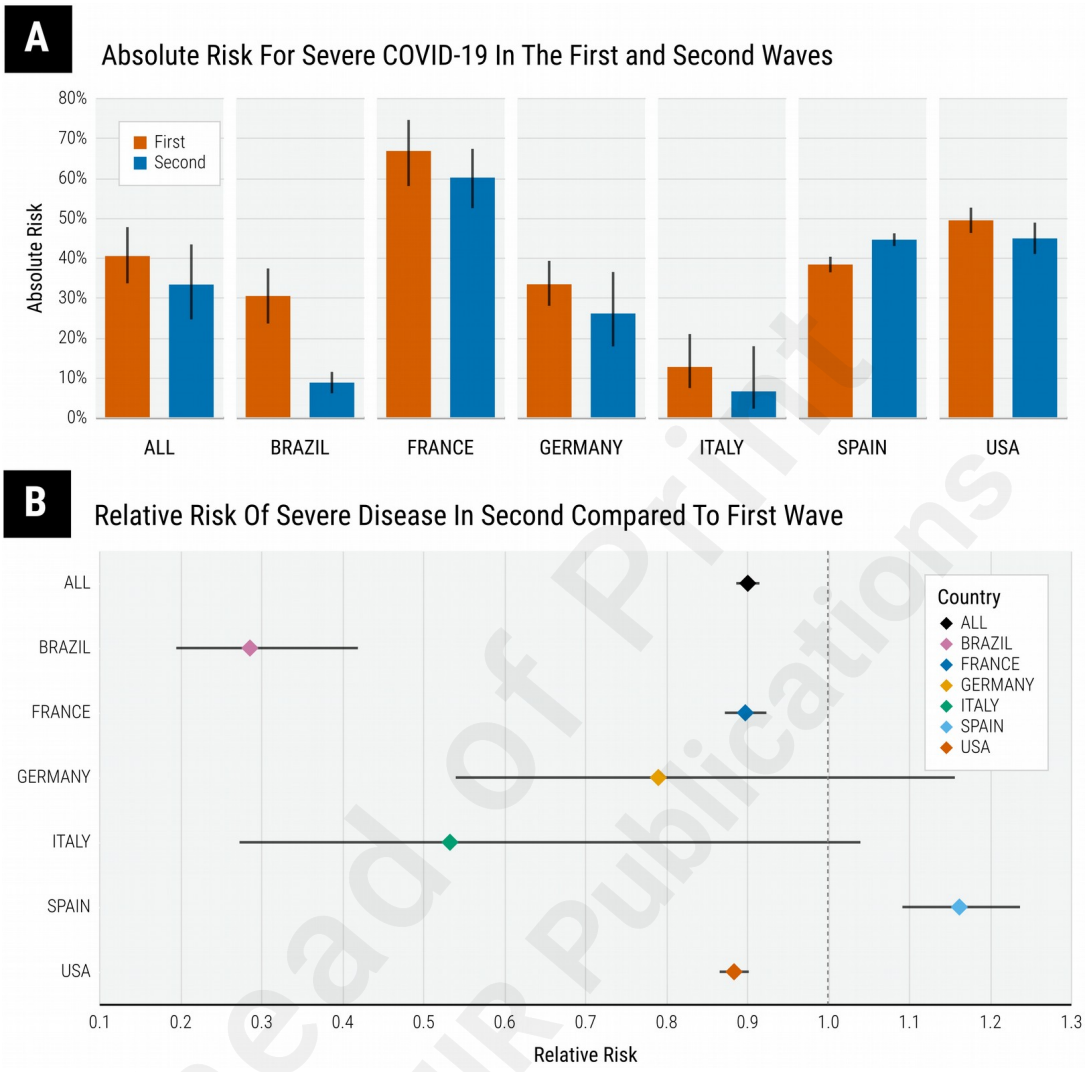
Germany: 33.3% vs. 25.8%; (4) Italy: 12.6% vs. 6.5%; (5) Spain: 38.3% vs. 44.5%; (6) USA: 49.4% vs. 44.8%. The relative risk of severe disease in the second wave compared to the first was more comparable across Brazil (RR: 0.29, 95% CI: 0.20–0.42), France (RR: 0.90, 95% CI: 0.87–0.92), and the USA (RR: 0.88, 95% CI: 0.87–0.90). The observed RR was 0.78 (95% CI: 0.52–1.17) in Germany and 0.53 (95% CI: 0.27–1.04) in Italy, but the reduction in risk was not statistically significant. In contrast, patients in Spain (RR: 1.16, 95% CI: 1.09–1.24) had a slightly higher risk of severe COVID-19 in the second wave compared to the first.

We report the absolute risk and relative risk for severe COVID-19 stratified by demographic subgroups in Figure 3A and Table 2 respectively. Across demographic subgroups in the second wave, there were significant reductions in risk among patients aged 26–49 and 50–69; male and female patients; as well as Black patients. Relative risk effect sizes were generally comparable between sexes and between races. The reduction in risk in the second wave was slightly greater for younger age groups than for older age groups.

Table 2: Relative risk of severe disease in the second wave compared to the first wave stratified by demographic subgroups and by country (95% CI). NA values denote no patients reported in specific demographic subgroups for certain countries. *Denotes statistical significance.

Demographic group		**All countries**	**Brazil**	**France**	**Germany**	**Italy**	**Spain**	**USA**
Years	**00 to 25**	0.75 (0.56, 1.02)	NA	1.50 (1.08, 2.09)*	NA	NA	NA	0.80 (0.59, 1.09)
	26 to 49	0.77 (0.63, 0.94)*	0.31 (0.16, 0.61)*	0.86 (0.57, 1.31)	1.1 (0.51, 2.45)	0.24 (0.07, 0.78)*	1.08 (0.90, 1.29)	0.81 (0.69, 0.94)*
	50 to 69	0.84 (0.72, 0.97)*	0.23 (0.12, 0.43)*	0.95 (0.89, 1.01)	0.88 (0.56, 1.39)	0.50 (0.17, 1.46)	1.12 (1.02, 1.23)*	0.88 (0.79, 0.98)*
	70 to 79	0.91 (0.80, 1.02)	0.26 (0.12, 0.60)*	1.00 (0.95, 1.06)	0.84 (0.53, 1.34)	1.38 (0.97, 1.97)	1.16 (1.03, 1.32)*	0.87 (0.76, 0.99)*
	>=80	1.01 (0.87, 1.17)	0.62 (0.11, 3.34)	0.93 (0.77, 1.12)	0.87 (0.37, 1.96)	1.13 (0.56, 2.28)	1.49 (1.31, 1.69)*	0.97 (0.83, 1.15)
Sex	**Female**	0.84 (0.73, 0.96)*	0.22 (0.12, 0.42)*	0.88 (0.84, 0.92)*	0.59 (0.13, 2.71)	0.87 (0.49, 1.53)	1.13 (1.02, 1.25)*	0.86 (0.76, 0.98)*
	Male	0.85 (0.76, 0.95)*	0.32 (0.19, 0.53)	0.93 (0.90, 0.96)*	0.58 (0.25, 1.34)	0.61 (0.29, 1.25)	1.18 (1.10, 1.28)*	0.89 (0.81, 0.98)*
Race	**Black**	0.89 (0.81, 0.98)*	NA	NA	NA	NA	NA	0.89 (0.81, 0.98)*
	White	0.91 (0.80, 1.03)	NA	NA	NA	NA	NA	0.91 (0.80, 1.03)

Figure 4: Absolute risk for severe COVID-19 in the first and second waves and relative risk of severe COVID-19 in the second compared to the first wave stratified by country. Error bars indicate 95% confidence intervals.



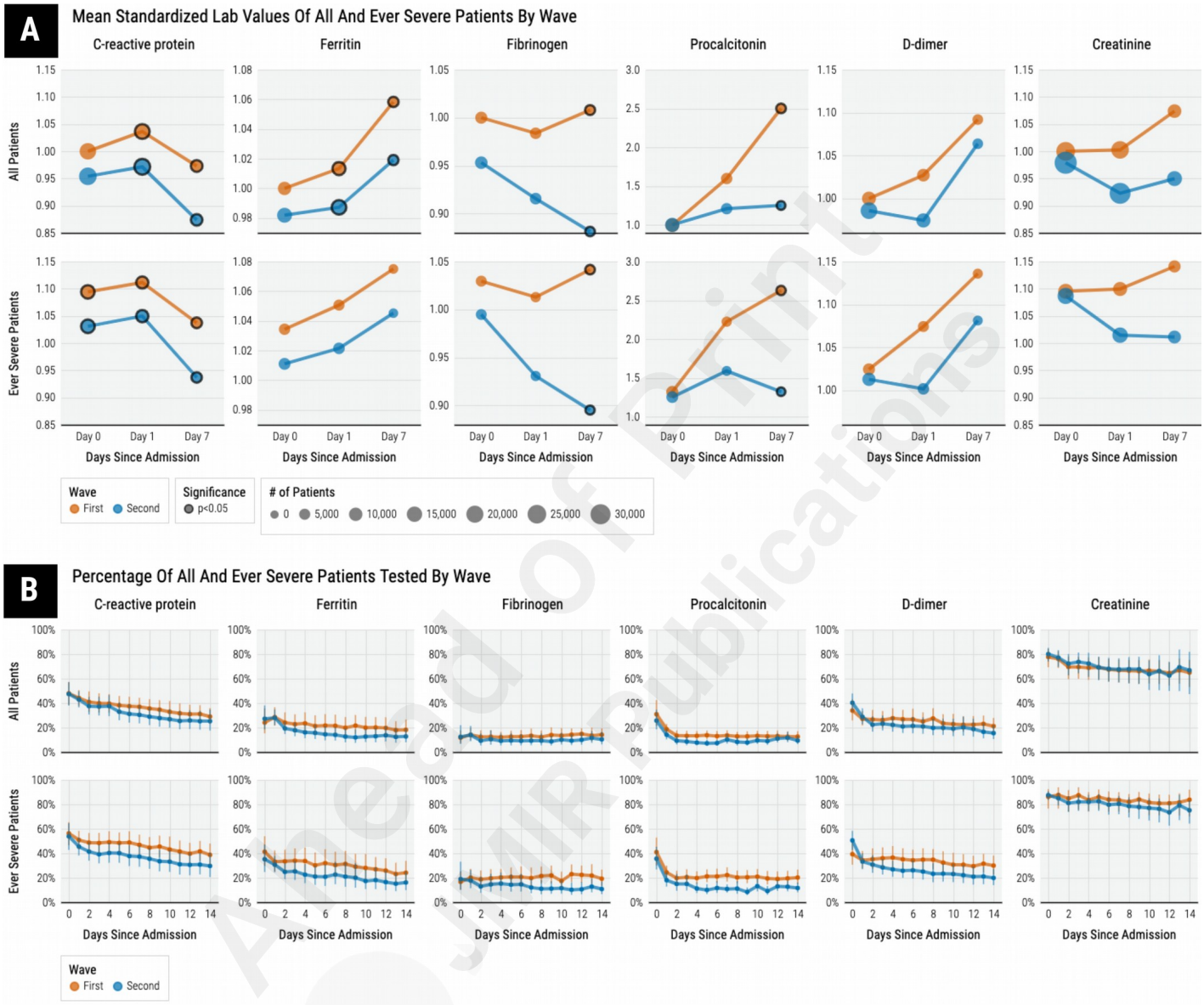
Change in Mean Laboratory Values and Laboratory Testing Rates

We report standardized mean laboratory values in the first and the second wave at day 0, 1, and 7 since the index date of admission for CRP, ferritin, fibrinogen, procalcitonin, D-dimer, and creatinine in Figure 5. Among all patients, we observed significantly lower mean CRP values throughout the first week of hospitalization on day 0, 1, and 7 in the second wave compared to the first wave. All other mean laboratory values on day 0 were not significantly different between the first and second waves. At day 7, we further observed that the mean values of ferritin, fibrinogen, and procalcitonin were all significantly lower during the second wave compared to the first wave.

Among patients with severe disease, we observed similar results with regards to comparing mean laboratory values between the first and second waves. Mean CRP values during the first week of hospitalization on day 0, 1, and 7 were significantly lower in the second wave compared to the first. Mean fibrinogen and procalcitonin values were significantly lower on day 7 in the second compared to the first wave.

Figure 5: Standardized mean laboratory values and corresponding laboratory testing rates among all

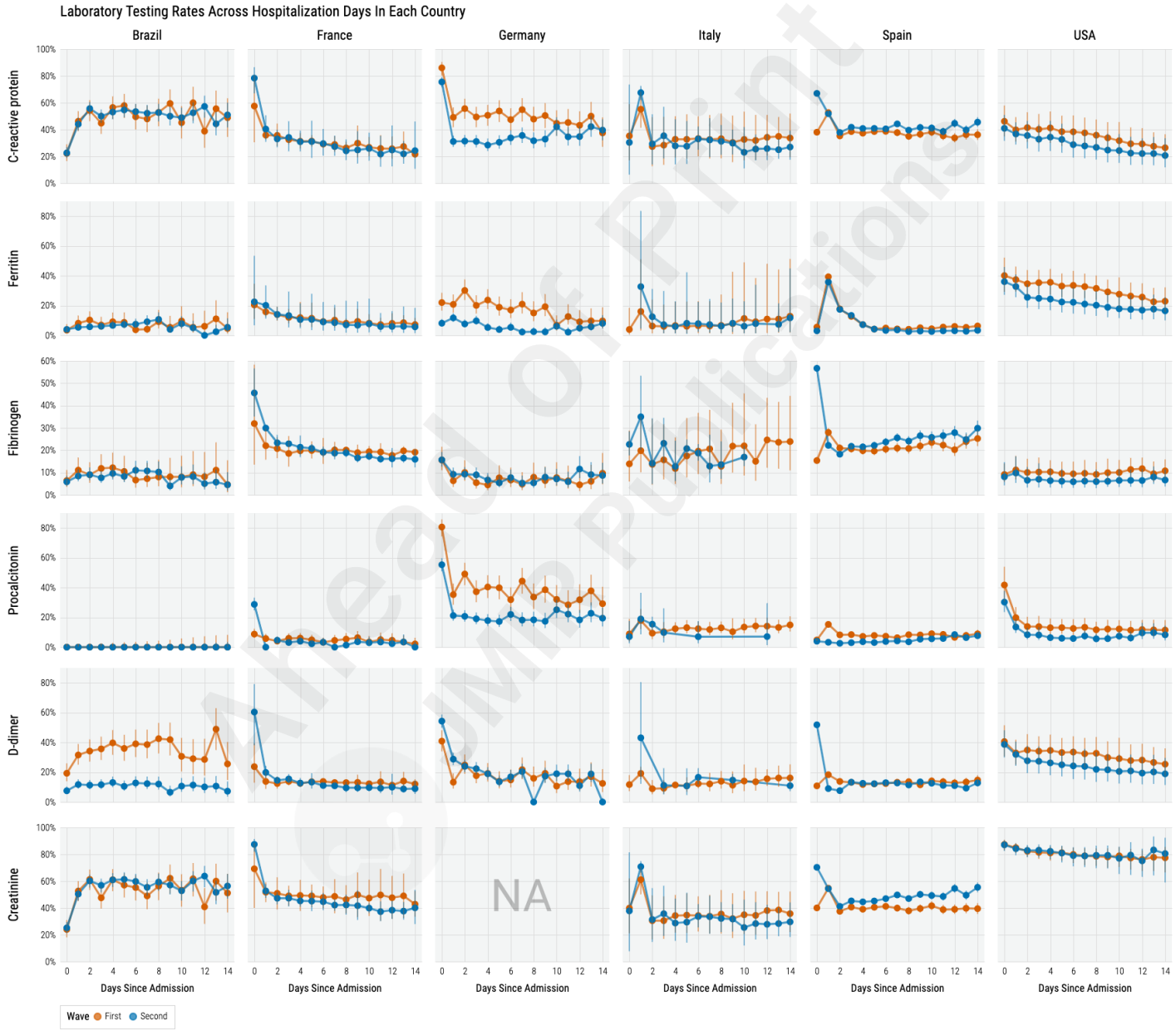
patients and those with severe disease by first and second wave. Error bars indicate 95% confidence intervals.



When comparing the overall laboratory testing rates during hospitalization, as reported in Figure 5, we observed that overall testing rates among all patients for procalcitonin were significantly lower across hospitalization days in the second wave compared to the first wave. Overall procalcitonin testing rates among patients with severe disease were similarly significantly lower across hospitalization days in the second wave compared to the first wave. There were no other significant changes in overall laboratory testing rates. We report laboratory testing rates within each country in Figure 6; creatinine laboratory data were not available from participating healthcare systems in Germany. Laboratory testing rates among countries varied significantly between the first and second wave. European countries exhibited the most changes in testing rates at admission (day 0): in the second wave, there was a significant increase in the D-dimer testing rate in France, Germany, and Spain and in the CRP, creatinine, and fibrinogen testing rate in Spain. In Germany, there were significant decreases in the testing rates for CRP, ferritin, and procalcitonin at admission. In Brazil, there was a significant decrease in testing rate for D-dimer at

admission. In contrast, the USA did not have any significant changes in laboratory testing rates at admission. In all countries except Brazil and Spain, laboratory testing rates in the second wave were generally lower during the second week of hospitalization. In Spain, second hospital week testing rates for CRP, fibrinogen, and creatinine were higher in the second wave compared to the first wave.

Figure 6: Laboratory testing rates across hospitalization days in each country. Error bars indicate 95% confidence intervals. Laboratory data for creatinine were unavailable for healthcare systems in Germany.



Discussion

Principal Results

In this large EHR-based study, we employed a federated approach to rapidly aggregate and harmonize clinical data across 315 international hospitals from six countries that included 79,613 hospitalized

patients with SARS-CoV-2 to offer insights on the evolving clinical trajectory of COVID-19 across the first and second waves. We found that patients hospitalized in the second wave were at significantly lower risk for severe COVID-19, corresponding to lower mean laboratory values for several inflammatory markers during the first week of hospitalization in the second wave compared to the first.

In this study, we capitalized on the availability of real-world EHR data from participating international healthcare systems within the 4CE Consortium to capture pertinent clinical characteristics effectively and accurately. Despite the high heterogeneity in the health systems, we were able to rigorously perform quality checks across all healthcare centers using a multi-disciplinary team approach that engages statisticians, informaticians and clinicians. Additionally, the multinational nature of our data allowed us to identify country-level variations in temporal trends as well as distinguish different clinical phenotypes and trajectories in the second wave compared to the first wave of the pandemic.

Notably, our federated approach demonstrated several advantages over methods where hospitals transfer patient-level data to a central repository [29]. By keeping data local within hospitals, we avoided privacy concerns and regulatory barriers that often delay multi-site research studies; by comparing the results from different hospitals, rather than treating all the data as a single combined dataset, we could identify outliers that suggested data quality problems; finally, by having local data experts at each site running the database queries, we could leverage their help in addressing these data quality concerns [30-31]. Although more sophisticated analyses such as machine learning models or robust multivariable models that can adjust for multiple sources of various types of bias are harder to implement in a federated architecture than in a central repository, this study demonstrated that our federated approach enabled us to obtain early clinical insights into the evolving pandemic and helps us gain confidence in the data.

Our data demonstrated a near-uniform peak in country-specific hospitalizations in the first wave and a variable peak in country-specific hospitalizations in the second wave, reflecting country-specific patterns in resurgence of COVID-19 related hospitalizations that were consistent with international tracking sites [1,9,32]. In all countries except Brazil and the USA, the second wave peak was characterized by lower intensity compared to the first wave peak, possibly reflecting the effects of successful COVID-19 mitigation measures implemented after the first wave [33-34].

Consistent with previous single-country studies, we observed that patients hospitalized in the second wave had an overall lower risk for severe COVID-19 than patients hospitalized in the first wave [6,12,35-36]. When further stratifying our analyses by country, we were able to observe that patients in Spain instead had a significantly higher risk for severe disease in the second wave, contrary to what we observed in France and the USA. Although the reasons for increased risk of severe COVID-19 during Spain's second wave are unclear and likely complex, it is consistent with international tracking sites that indicate increased mortality rate and healthcare resource use in Spain's second wave, and it reflects the importance in being able to identify country-specific variations in our data [37]. We further note that data from Spain originated from one hospital and was likely subject to some forms of bias. Even when stratifying by demographic subgroups, we observed similar patterns indicating reduced risk for severe COVID-19 in the

second wave, particularly among patients aged 26–49 and among Black patients. However, while the risk of severe disease for the entire population was lower in the second wave for patients aged 26–49, country-specific results demonstrated possible collider bias and small-sample bias in that the individual estimates for Brazil, France, Germany, and Spain were not statistically significant and were imprecise with wide confidence intervals. This is likely due in large part to smaller sample sizes of this age group in each country and reflects the need for multi-center studies to improve power. While one might expect to see less severe disease later in the course of the pandemic due to changes in patient populations over time, improved clinical care, and greater utilization of healthcare resources compared to the beginning of the pandemic, it is unclear why we did not observe similar relationships for other patient groups. These possible discrepancies are likely due to a variety of different reasons, as noted above, as well as sources of bias in our data. Further investigation into these country-specific demographic differences in the change in severe disease risk over the course of the pandemic is warranted and is ongoing in the 4CE Consortium.

Our observations comparing laboratory values between the first and second waves support our finding that patients hospitalized during the second wave compared to the first had a lower risk for severe disease. We found that mean laboratory values in the second wave exhibited considerable improvement towards typical physiological values compared to the first wave—especially inflammatory markers. In particular, the mean values of the positive acute phase reactant CRP were lower across the first week of hospitalization in the second wave compared to the first, while the mean values of the positive acute phase reactants ferritin, fibrinogen, and procalcitonin were lower at day 7 [38]. This indicates that, on average, patients hospitalized during the second wave may have had less overall systemic inflammation at admission and had improved their inflammatory states during the first week of hospitalization in comparison to patients admitted in the first wave [39–41]. Considering that there were no new major effective pharmacologic therapies for patients with COVID-19 introduced between the first and second waves, these general patterns may be reflective of a less vulnerable patient population in the second wave as well as improved general clinical management strategies of COVID-19 in the second wave [42–50]. Ongoing 4CE analyses are further investigating these findings.

We further observed variations in laboratory testing rates among countries between the first and second wave. These changes in laboratory testing rates at admission may be reflective of greater understanding of COVID-19 pathophysiology and clinical trajectories leading to changes in clinical protocols. For example, there was a significant increase in the second wave in the testing rate for D-dimer at admission in France, Germany, and Spain. This particular change in clinical practice may be driven in part by the growing literature supporting the association of high D-dimer values with worse outcomes in COVID-19 and the possibility of using D-dimer to clinically classify and evaluate the prognosis of COVID-19 patients [51–55]. Further, although there were no significant changes in laboratory testing at admission in the USA, we observed that testing rates across hospitalization days were generally higher than other European countries regardless of wave. Future investigations are warranted to infer why we observed these patterns.

Study Limitations

We acknowledge several limitations for this EHR-based observational cohort study. This study was

limited to patients who were admitted to a hospital, either because they experienced more severe illness or because they had other possibly biasing conditions; as with many EHR-based studies, we were unable to ascertain the precise reasons for admission. Similar to other EHR-based studies, we were not able to validate if patients were hospitalized due to COVID-19 or happened to have a positive test when admitted for an unrelated medical condition. Thus, we could not completely mitigate selection bias or misclassification bias in our cohort identification. Due to the limited scope of the extracted aggregate data, we could not effectively control for patient-level potentially confounding variables such as comorbidities, medication use (both prior to and during hospitalization), and other societal and environmental factors, all of which can induce many types of bias [56]. Data pertaining to certain countries, most notably Brazil and Germany, may have been subject to small sample bias. Furthermore, mean laboratory values at later days of hospitalization were subject to censoring (transfer, discharge, death) and thus dropout bias, so we could not effectively compare mean laboratory values within a single wave at different timepoints. However, we believe that facilitating comparisons at identical timepoints between different waves is not subject to as much dropout bias. In an effort to provide information regarding the nature of censoring that existed in the data, we report in Supplementary Figure 2 the proportion of patients who were alive and remained in the hospital across hospitalization days for each country by wave. Further, considering the aforementioned limitations, we took special caution to make conclusions that were mostly descriptive in nature. In the future, we hope to disaggregate EHR data to the patient level in order to adjust for many of these biases if it is possible under institutional IRBs.

Conclusions

To study the evolving epidemiology, pathophysiology, and healthcare dynamics of the COVID-19 pandemic, we leveraged EHR data in a large international cohort of hospitalized patients with SARS-CoV-2 to rapidly characterize the clinical course of patients admitted to hospital during the first two major waves of the pandemic. We were able to characterize changes in hospitalization rates, demographic characteristics, severity risk, and mean laboratory values using data from 79,613 patients across 315 healthcare systems in 6 countries. Our study's federated approach demonstrates the feasibility and power of leveraging real-world EHR data from multiple countries to support our understanding of evolving pandemics such as COVID-19.

Conflicts of Interest

There are no competing interests to report.

Ethics Statement

All study sites were responsible for and obtained ethics approval, as needed, from the appropriate ethics committee at their institution.

The lead authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned have been explained.

Data Sharing Statement

De-identified aggregate data was provided by sites for this study. The 4CE consortium provides samples of de-identified data collected by the consortium and some corresponding visualizations on the consortium website: https://covidclinical.net.

Patients and Public Involvement Statement

Patients and the public were not involved in the design, conduct, or reporting, or dissemination plans of the research.

Funding Statement

GMW is supported by NIH/NCATS UL1TR002541 and UL1TR000005, NIH/NLM R01LM013345, and NIH/NHGRI 3U01HG008685-05S2. NG is supported by NIH/NLM T15 LM007092. DAH is supported by NCATS UL1TR002240. RK is supported by CTSA award UL1TR001998. KK is supported by NIH/NCATS UL1 TR001450. DLM is supported by NIH/NCATS UL1-TR001878. SNM is supported by NIH/NCATS 5UL1TR001857-05 and NIH/NHGRI 5R01HG009174-04. GSO is supported by NIH grants P30ES017885 and U24CA210967. LPP is supported by CTSA Award #UL1TR002366. JS is supported by NIH/NCATS UL1TR001881. SV is supported by NIH/NLM R01LM012095 and NIH/NCATS UL1TR001857. ZX is supported by NINDS R01NS098023. AMS is supported by NIH/NHLBI K23HL148394, L40HL148910 and NIH/NCATS UL1TR001420.

Author Contributions

All authors made contributions to: conception and design; acquisition, analysis, and interpretation of data; drafting the article or revising it critically for important intellectual content; and final approval of the version to be published.

References

1. COVID-19 map - Johns Hopkins Coronavirus resource Center [Internet]. [cited 2021 May 3]. Available from: https://coronavirus.jhu.edu/map.html

2. Devi S. COVID-19 resurgence in Iran. Lancet 2020 Jun 20;395(10241):1896. PMID:32563363

3. Looi M-K. Covid-19: Is a second wave hitting Europe? BMJ 2020 Oct 28;371:m4113. PMID:33115704

4. Shimizu K, Wharton G, Sakamoto H, Mossialos E. Resurgence of covid-19 in Japan. BMJ 2020 Aug 18;370:m3221. PMID:32816821

5. Sabino EC, Buss LF, Carvalho MPS, Prete CA Jr, Crispim MAE, Fraiji NA, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet 2021 Feb 6;397(10273):452–455. PMID:33515491

6. Saito S, Asai Y, Matsunaga N, Hayakawa K, Terada M, Ohtsu H, et al. First and second COVID-19 waves in Japan: A comparison of disease severity and characteristics. J Infect. 2021. p. 84–123. PMID:33152376

7. Win A. Rapid rise of COVID-19 second wave in Myanmar and implications for the Western Pacific region. QJM 2020 Dec 1;113(12):856–857. PMID:33095232

8. Vahidy FS, Drews AL, Masud FN, Schwartz RL, Askary BB, Boom ML, et al. Characteristics and Outcomes of COVID-19 Patients During Initial Peak and Resurgence in the Houston Metropolitan Area. JAMA 2020 Sep 8;324(10):998–1000. PMID:32789492

9. IHME [Internet]. [cited 2021 Jun 17]. Available from: https://covid19.healthdata.org/united-states-of-america?view=cumulative-deaths&tab=trend

10. Coronavirus (COVID-19) Deaths [Internet]. [cited 2021 Jun 17]. Available from: https://ourworldindata.org/covid-deaths

11. Mollinedo-Gajate I, Villar-Álvarez F, de los Ángeles Zambrano-Chacón M, Núñez-García L, de la Dueña-Muñoz L, López-Chang C, et al. First and Second Waves of Coronavirus Disease 2019 in Madrid, Spain: Clinical Characteristics and Hematological Risk Factors Associated With Critical/Fatal Illness [Internet]. Critical Care Explorations. 2021. p. e0346. [doi: 10.1097/cce.0000000000000346]

12. Contou D, Fraissé M, Pajot O, Tirolien J-A, Mentec H, Plantefève G. Comparison between first and second wave among critically ill COVID-19 patients admitted to a French ICU: no prognostic improvement during the second wave? Crit Care 2021 Jan 4;25(1):3. PMID:33397421

13. Long SW, Olsen RJ, Christensen PA, Bernard DW, Davis JJ, Shukla M, et al. Molecular Architecture of Early Dissemination and Massive Second Wave of the SARS-CoV-2 Virus in a Major Metropolitan Area. MBio [Internet] 2020 Oct 30;11(6). PMID:33127862

14. Brat GA, Weber GM, Gehlenborg N, Avillach P, Palmer NP, Chiovato L, et al. International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium. NPJ Digit Med 2020 Aug 19;3:109. PMID:32864472

15. Weber GM, Hong C, Palmer NP, Avillach P, Murphy SN, Gutiérrez-Sacristán A, et al. International

22

Comparisons of Harmonized Laboratory Value Trajectories to Predict Severe COVID-19: Leveraging the 4CE Collaborative Across 342 Hospitals and 6 Countries: A Retrospective Cohort Study. medRxiv [Internet] 2021 Feb 5; PMID:33564777

16. Jones AL, Pettey WBP, Carter ME, Brignone E, Redd A, Suo Y, et al. Regional Variations in Documentation of Sexual Trauma Concepts in Electronic Medical Records in the United States Veterans Health Administration. AMIA Annu Symp Proc 2019;2019:514–522. PMID:32308845

17. Le TT, Gutiérrez-Sacristán A, Son J, Hong C, South AM, Beaulieu-Jones BK, et al. Multinational Prevalence of Neurological Phenotypes in Patients Hospitalized with COVID-19. medRxiv [Internet] 2021 Jan 29; PMID:33655281

18. Klann JG, Weber GM, Estiri H, Moal B, Avillach P, Hong C, et al. Validation of an Internationally Derived Patient Severity Phenotype to Support COVID-19 Analytics from Electronic Health Record Data. J Am Med Inform Assoc [Internet] 2021 Feb 10; PMID:33566082

19. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986 Sep;7(3):177–188. PMID:3802833

20. Office of Policy, Planning. National center for veterans analysis and statistics [Internet]. [cited 2021 Jun 7]. Available from: https://www.va.gov/vetdata/veteran_population.asp

21. Huang I, Pranata R, Lim MA, Oehadian A, Alisjahbana B. C-reactive protein, procalcitonin, D-dimer, and ferritin in severe coronavirus disease-2019: a meta-analysis. Ther Adv Respir Dis 2020 Jan;14:1753466620937175. PMID:32615866

22. Carubbi F, Salvati L, Alunno A, Maggi F, Borghi E, Mariani R, et al. Ferritin is associated with the severity of lung involvement but not with worse prognosis in patients with COVID-19: data from two Italian COVID-19 units. Sci Rep 2021 Mar 1;11(1):4863. PMID:33649408

23. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020 Feb 15;395(10223):497–506. PMID:31986264

24. Smilowitz NR, Kunichoff D, Garshick M, Shah B, Pillinger M, Hochman JS, et al. C-reactive protein and clinical outcomes in patients with COVID-19. Eur Heart J [Internet] 2021 Jan 15; PMID:33448289

25. Shi S, Nie B, Chen X, Cai Q, Lin C, Zhao G, et al. Clinical and laboratory characteristics of severe and non-severe patients with COVID-19: A retrospective cohort study in China. J Clin Lab Anal 2021 Jan;35(1):e23692. PMID:33389777

26. Bi X, Su Z, Yan H, Du J, Wang J, Chen L, et al. Prediction of severe illness due to COVID-19 based on an analysis of initial Fibrinogen to Albumin Ratio and Platelet count. Platelets 2020 Jul 3;31(5):674–679. PMID:32367765

27. Statsenko Y, Al Zahmi F, Habuza T, Gorkom KN-V, Zaki N. Prediction of COVID-19 severity using laboratory findings on admission: informative values, thresholds, ML model performance. BMJ Open 2021 Feb 26;11(2):e044500. PMID:33637550

28. Wang F, Hou H, Wang T, Luo Y, Tang G, Wu S, et al. Establishing a model for predicting the outcome of COVID-19 based on combination of laboratory tests. Travel Med Infect Dis 2020 Jul;36:101782. PMID:32526372

23

29. Haendel MA, Chute CG, Bennett TD, Eichmann DA, Guinney J, Kibbe WA, et al. The National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment. J Am Med Inform Assoc 2021 Mar 1;28(3):427–443. PMID:32805036

30. Wallace S, Myles PS. Solving the challenges of large multicenter trials in anesthesia. HSR Proc Intensive Care Cardiovasc Anesth 2009;1(3):46–53. PMID:23439979

31. de Jonge JC, Reinink H, Colam B, Alpers I, Ciccone A, Csiba L, et al. Regulatory delays in a multinational clinical stroke trial. European Stroke Journal SAGE Publications; 2021 Mar 30;23969873211004845.

32. Coronavirus (COVID-19) Deaths [Internet]. [cited 2021 Jun 17]. Available from: https://ourworldindata.org/covid-deaths

33. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nat Hum Behav 2020 Dec;4(12):1303–1312. PMID:33199859

34. Tam K-M, Walker N, Moreno J. Effect of mitigation measures on the spreading of COVID-19 in hard-hit states in the U.S. PLoS One 2020 Nov 3;15(11):e0240877. PMID:33141823

35. Docherty AB, Mulholland RH, Lone NI, Cheyne CP, De Angelis D, Diaz-Ordaz K, et al. Changes in in-hospital mortality in the first wave of COVID-19: a multicentre prospective observational cohort study using the WHO Clinical Characterisation Protocol UK. Lancet Respir Med [Internet] 2021 May 14; PMID:34000238

36. Karagiannidis C, Windisch W, McAuley DF, Welte T, Busse R. Major differences in ICU admissions during the first and second COVID-19 wave in Germany. Lancet Respir Med 2021 May;9(5):e47–e48. PMID:33684356

37. IHME [Internet]. [cited 2021 Jun 17]. Available from: https://covid19.healthdata.org/spain?view=infections-testing&tab=trend&test=infections

38. Gulhar R, Ashraf MA, Jialal I. Physiology, Acute Phase Reactants. StatPearls Publishing; 2021.

39. Nehring SM, Goyal A, Bansal P, Patel BC. C Reactive Protein. StatPearls Publishing; 2021.

40. Knovich MA, Storey JA, Coffman LG, Torti SV, Torti FM. Ferritin for the clinician. Blood Rev 2009 May;23(3):95–104. PMID:18835072

41. Kell DB, Pretorius E. Serum ferritin is an important inflammatory disease marker, as it is mainly a leakage product from damaged cells. Metallomics 2014 Apr;6(4):748–773. PMID:24549403

42. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB. Pharmacologic Treatments for Coronavirus Disease 2019 (COVID-19): A Review. JAMA 2020 May 12;323(18):1824–1836. PMID:32282022

43. Sheahan TP, Sims AC, Leist SR, Schäfer A, Won J, Brown AJ, et al. Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. Nat Commun 2020 Jan 10;11(1):222. PMID:31924756

44. Gautret P, Lagier J-C, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents 2020 Jul;56(1):105949. PMID:32205204

45. Spinner CD, Gottlieb RL, Criner GJ, Arribas López JR, Cattelan AM, Soriano Viladomiu A, et al. Effect of Remdesivir vs Standard Care on Clinical Status at 11 Days in Patients With Moderate COVID-19: A Randomized Clinical Trial. JAMA 2020 Sep 15;324(11):1048–1057. PMID:32821939

46. Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. Lancet 2020 May 16;395(10236):1569–1578. PMID:32423584

47. Boulware DR, Pullen MF, Bangdiwala AS, Pastick KA, Lofgren SM, Okafor EC, et al. A Randomized Trial of Hydroxychloroquine as Postexposure Prophylaxis for Covid-19. N Engl J Med 2020 Aug 6;383(6):517–525. PMID:32492293

48. Mitjà O, Corbacho-Monné M, Ubals M, Alemany A, Suñer C, Tebé C, et al. A Cluster-Randomized Trial of Hydroxychloroquine for Prevention of Covid-19. N Engl J Med 2021 Feb 4;384(5):417–427. PMID:33289973

49. Self WH, Semler MW, Leither LM, Casey JD, Angus DC, Brower RG, et al. Effect of Hydroxychloroquine on Clinical Status at 14 Days in Hospitalized Patients With COVID-19: A Randomized Clinical Trial. JAMA 2020 Dec 1;324(21):2165–2176. PMID:33165621

50. WHO Solidarity Trial Consortium, Pan H, Peto R, Henao-Restrepo A-M, Preziosi M-P, Sathiyamoorthy V, et al. Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results. N Engl J Med 2021 Feb 11;384(6):497–511. PMID:33264556

51. Yao Y, Cao J, Wang Q, Shi Q, Liu K, Luo Z, et al. D-dimer as a biomarker for disease severity and mortality in COVID-19 patients: a case control study. J Intensive Care Med 2020 Jul 10;8:49. PMID:32665858

52. Düz ME, Balcı A, Menekşe E. D-dimer levels and COVID-19 severity: Systematic Review and Meta-Analysis. Tuberk Toraks 2020 Dec;68(4):353–360. PMID:33448732

53. Yu H-H, Qin C, Chen M, Wang W, Tian D-S. D-dimer level is associated with the severity of COVID-19. Thromb Res 2020 Nov;195:219–225. PMID:32777639

54. Berger JS, Kunichoff D, Adhikari S, Ahuja T, Amoroso N, Aphinyanaphongs Y, et al. Prevalence and Outcomes of D-Dimer Elevation in Hospitalized Patients With COVID-19. Arterioscler Thromb Vasc Biol 2020 Oct;40(10):2539–2547. PMID:32840379

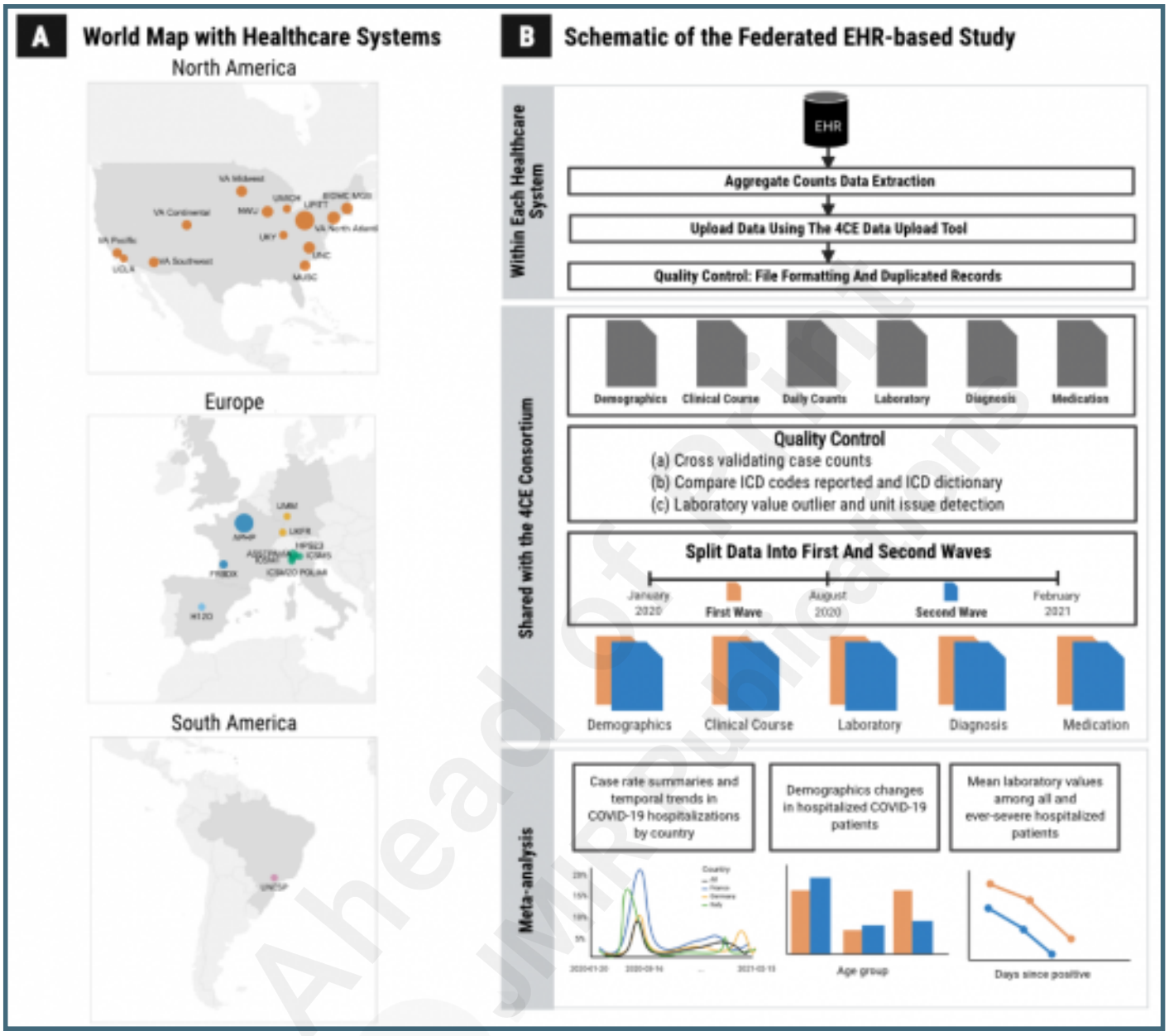
55. He X, Yao F, Chen J, Wang Y, Fang X, Lin X, et al. The poor prognosis and influencing factors of high D-dimer levels for COVID-19 patients. Sci Rep 2021 Jan 19;11(1):1830. PMID:33469072

56. Cohen JB, D'Agostino McGowan L, Jensen ET, Rigdon J, South AM. Evaluating sources of bias in observational studies of angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker use during COVID-19: beyond confounding. J Hypertens 2021 Apr 1;39(4):795–805. PMID:33186321

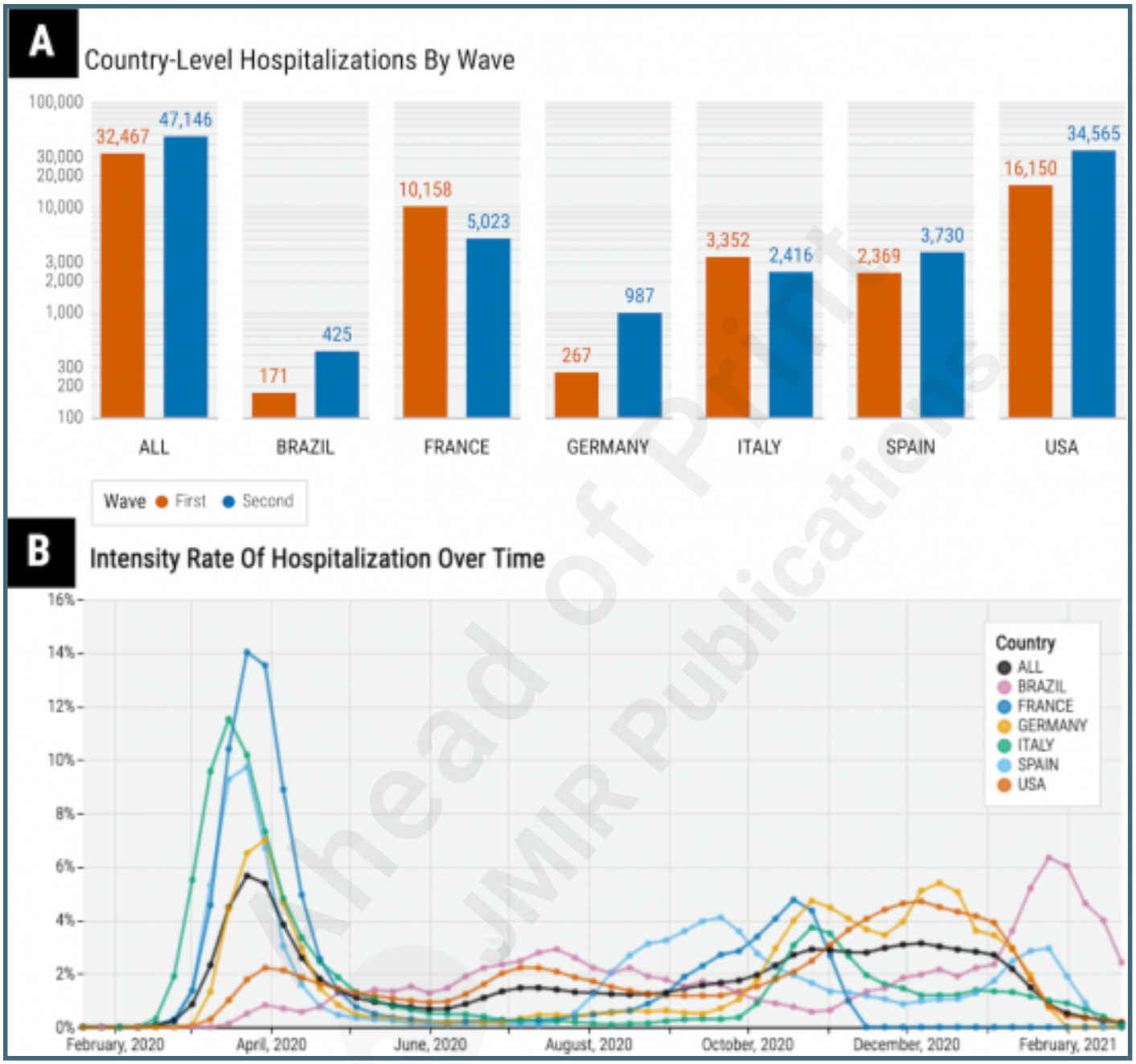
Supplementary Files

Figures

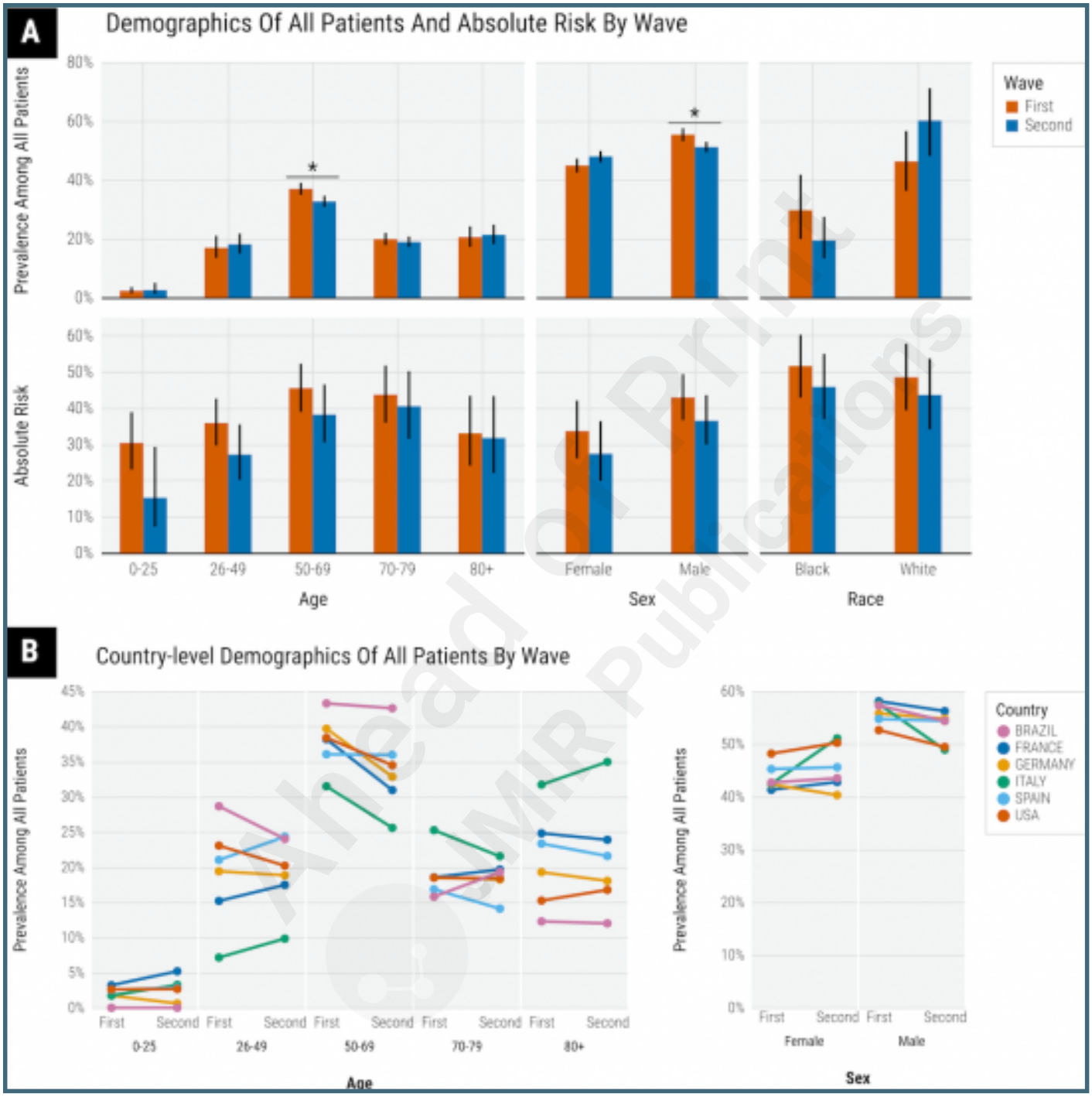
Schematic of the federated EHR-based study involving healthcare systems from six countries.



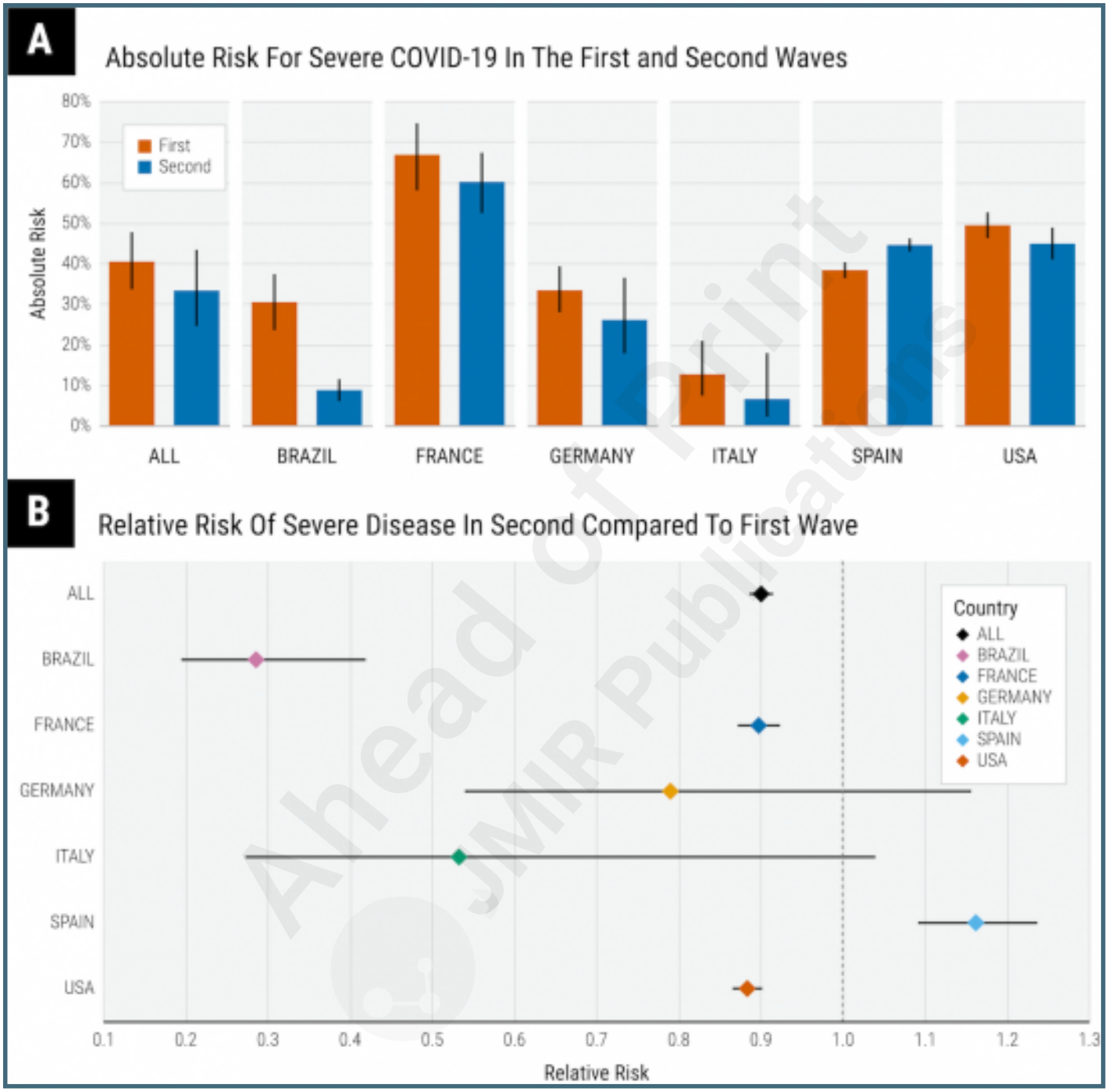
Total hospitalizations in the cohort between first and second waves and intensity rate of hospitalizations over time by country.



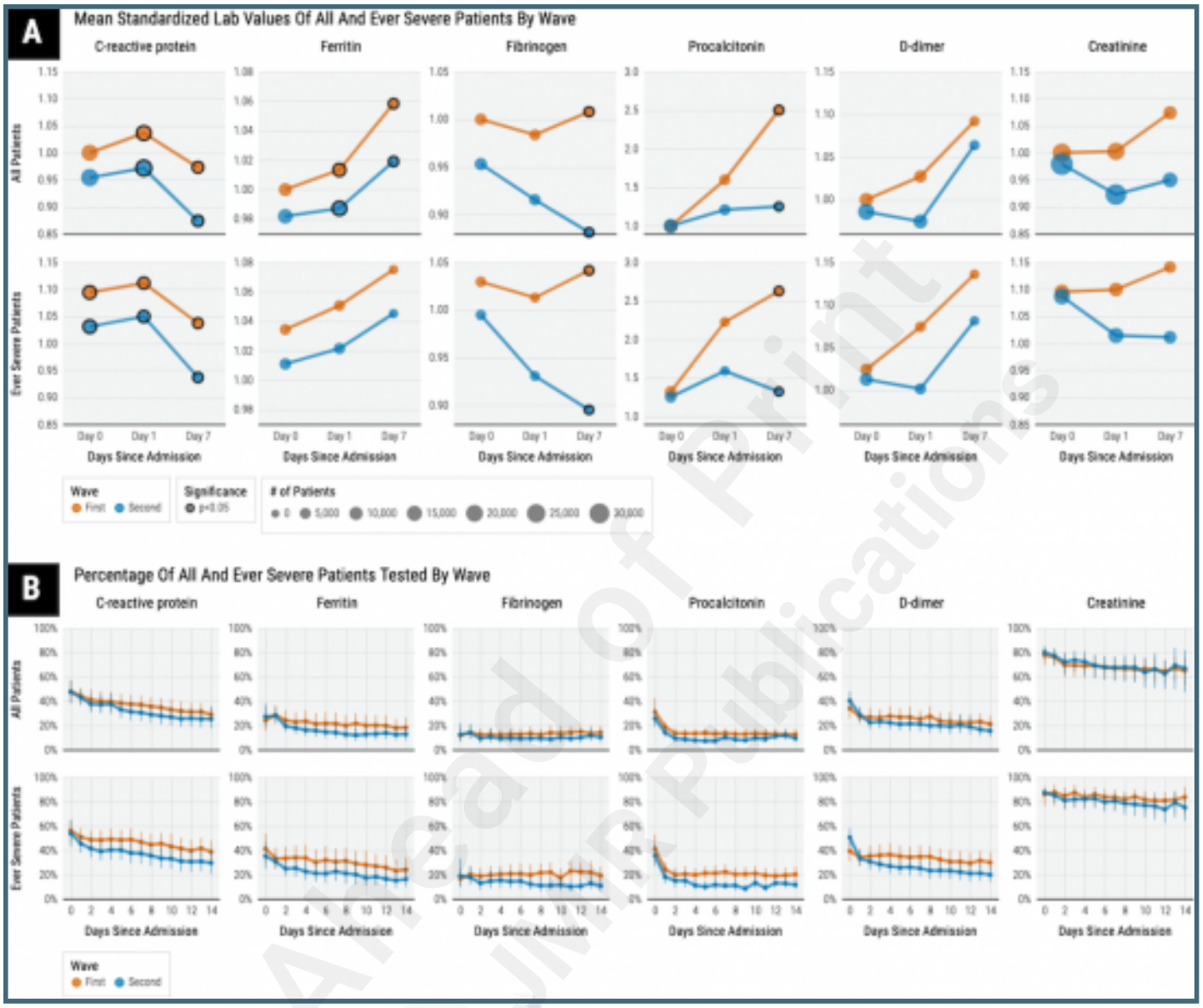
Prevalence of demographic subgroups in the first and second waves for the entire cohort and by country, and estimated absolute risk for severe COVID-19 by demographic subgroup. Error bars indicate 95% confidence intervals. *Indicates a significant difference by non-overlapping 95% confidence intervals.



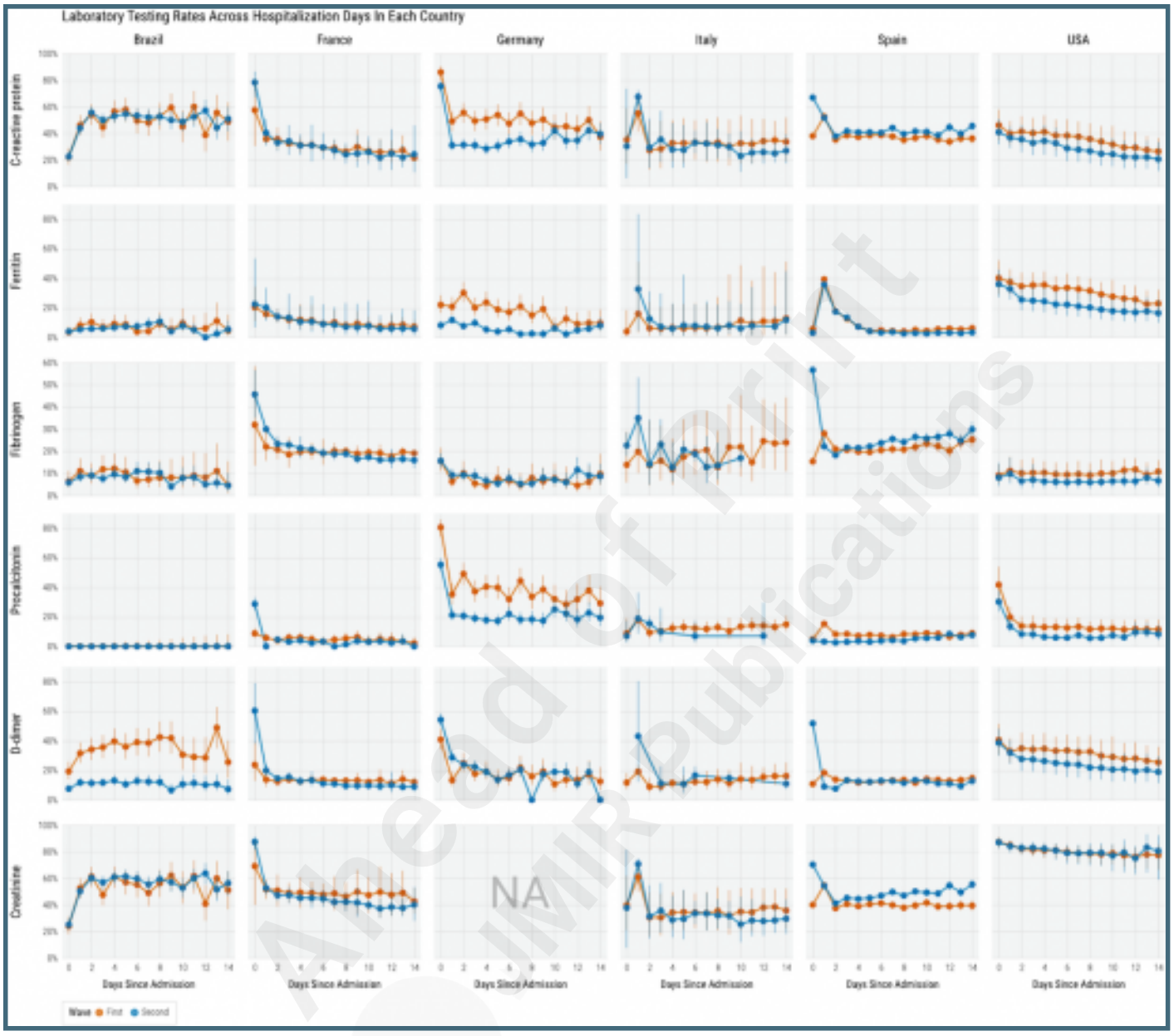
Absolute risk for severe COVID-19 in the first and second waves and relative risk of severe COVID-19 in the second compared to the first wave stratified by country. Error bars indicate 95% confidence intervals.



Standardized mean laboratory values and corresponding laboratory testing rates among all patients and those with severe disease by first and second wave. Error bars indicate 95% confidence intervals.



Laboratory testing rates across hospitalization days in each country. Error bars indicate 95% confidence intervals. Laboratory data for creatinine were unavailable for healthcare systems in Germany.



Multimedia Appendixes

Prevalence of demographic subgroups in the first and second waves for the entire cohort and by country and estimated absolute risk for severe COVID-19 by demographic subgroups inclusive of 5 VA healthcare systems. Error bars indicate 95% confidence intervals. *Indicates a significant difference by non-overlapping 95% confidence intervals.
URL: http://asset.jmir.pub/assets/a6b7ffd0d41feb8eec71b57ddabbfd1f.png

Proportion of patients who are alive and still in the hospital across hospitalization days.
URL: http://asset.jmir.pub/assets/3f3440f9f1ccebc9ecd01fd0ffff369f.png

Descriptions of CSV files generated at participating healthcare systems.
URL: http://asset.jmir.pub/assets/fe268a903a4cada42d5b2be79fd26e1a.docx

Consortium for Clinical Characterization of COVID-19 by EHR (4CE Consortium) Member List.
URL: http://asset.jmir.pub/assets/4fbdb1344d47c08fdcddcfab7ff62b3a.docx