Fluid, pH and electrolyte imbalance associated with COVID-19 mortality

Satu Nahkuri (satu.nahkuri@roche.com)
Roche Innovation Center Zurich
Tim Becker
Roche Innovation Center Munich https://orcid.org/0000-0002-6573-1109
Vitalia Schueller
Pharma Research and Development Roche Innovation Center Munich
Steffen Massberg
University Hospital, LMU Munich https://orcid.org/0000-0001-7387-3986
Anna Bauer-Mehren
Roche Innovation Center Munich

Article

Keywords: COVID-19, mortality, pre-existing comorbidities, pH imbalance, electrolyte imbalance,

Posted Date: March 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-145823/v1

Version of Record: A version of this preprint was published at Communications Medicine on November 25th, 2021. See the published version at https://doi.org/10.1038/s43856-021-00051-x.

Abstract

The threat of COVID-19 has harried the world since early 2020. Risk of death from the infection is associated with age and pre-existing comorbidities such as diabetes, dementia, cancer, and impairment of immunological, hepatic or renal function. It still remains incompletely understood why some patients survive the disease, while others perish. Our univariate and multivariate analyses of real world data from U.S. electronic health records indicate that a priori diagnoses of fluid, pH and electrolyte imbalance are highly and independently associated with COVID-19 mortality. We propose that pre-existing homeostatic aberrations are magnified upon the loss of ACE2, which is a core component of the electrolyte management system as well as the entry point of internalizing SARS-CoV-2 viruses. Moreover, we also suggest such fragility of electrolyte homeostasis may increase the risk of plasma volume disturbances during the infection. Future interventional studies should investigate whether the risk of death can be alleviated by personalized management of the fluid and electrolyte balance of at-risk individuals before and during COVID-19.

Introduction

In 2020, the world ground to a halt owing to the COVID-19 pandemic that still continues on its course. Achieving durable universal sterilizing immunity, may that be through widespread primary infections or population-penetrating and transmission-halting vaccinations, remains a remote and possibly impermanent prospect. Instead, it seems a chilling possibility that a large proportion of the remaining worldwide COVID-19-naive population will contract the virus in the next few months. We anticipate that even slight improvements in fortifying the health of risk group individuals in anticipation of a potential infection, as well as improving the outcomes of patients succumbing to the severe form of COVID-19, may translate to vast savings in loss of life.

The novel SARS-CoV-2 virus that causes the COVID-19 infection enters human cells via the ACE2 receptor. The infection first occurs in upper airways and at later stages may proceed to the lung, gastrointestinal tract, kidney, heart, or brain (Gupta et al. 2020). ACE2 and its homolog ACE are core enzymes of the renin-angiotensin-aldosterone system (RAAS), which regulates electrolyte homeostasis, blood pressure and cardiovascular health (Patel et al. 2016), as well as restores balance upon volume disturbance of extracellular fluid (Fountain and Lappin 2020; Navar 2014). The antagonistic effects of ACE and ACE2 are largely achieved by increasing or decreasing, respectively, the amount of circulating Angiotensin II and consequently, aldosterone, a hormone that is released from the adrenal cortex and regulates fluid and electrolyte excretion or retainment by the kidneys.

The underlying mechanisms of infection and viral spread are not fully understood, and despite the advances in prevention and treatment of severe COVID-19, an unmet need remains to better understand the clinical course and the risk factors of severe disease and death. Here, we present an agnostic and data-driven analysis of real world data from U.S. electronic health records of 122,250 COVID-19 patients to identify factors that predict death during a COVID-19 infection. Our unbiased analyses reveal a group

of novel, interconnected risk factors for COVID-19 mortality, in addition to confirming many that are already known. We present a working hypothesis of a link between these new risk factors and the renin-angiotensin-aldosterone system, and suggest opportunities for better care and/or prevention of severe COVID-19 disease.

Methods

Study design and participants

We extracted COVID-19 cases collected by Optum® with a diagnosis date between 20 February and 1 July 2020. The Optum® de-identified COVID-19 EHR dataset contains patient-level medical and administrative records from hospitals, emergency departments, outpatient centers, and laboratories from across the United States. Data de-identification is performed in compliance with the HIPAA Expert Method and managed according to Optum® customer data use agreements. The COVID-19 EHR dataset sources clinical information from hospital networks that provide data meeting Optum's internal data quality criteria. We confirmed COVID-19 diagnosis either by documented ICD-10 codes (Supplementary Methods) or via positive PCR test result (Supplementary Table S1). Survival time was computed as the number of days between the date of COVID-19 diagnosis and last documented clinical activity (vitals, labs, medication, encounter, collected until 13 July 2020) or documented death.

We analyzed variables that were observed at least a month before the infection. We selected such a conservative buffer time period, since some patients might have had an undiagnosed COVID-19 infection for days before an opportunity to get tested. We also wanted to exclude any physiological changes potentially incurred by the virus during the incubation period, which may be up to 14 days (Lauer et al. 2020). We used the median value captured between 1 and 12 months before the initial COVID-19 diagnosis for vitals and laboratory measurements, and the entire past medical history (median length ~4 years) for variables with a long-term effect, such as diagnoses of chronic indications (Supplementary Methods). Prior inoculations were handled analogously.

We investigated all disease entities that are a part of the Charlson comorbidity index (Quan et al. 2011), the AHRQ (Agency for Healthcare Research and Quality 2018), or the former Elixhauser definition (Elixhauser et al. 1998). A detailed description of assignment of ICD codes to disease entities can be found in Supplementary Table S2.

After quality control (Supplementary Methods), 249 variables were available for primary univariate analysis. For multivariate analysis, patients with an exaggerated proportion of missing data (Supplementary Methods) were removed, leaving 55,757 patients for analysis. Variables available for less than 10,000 patients were mean-imputed, resulting in a remaining missingness rate of 12.4% in total. The remaining missing values were imputed using the missForest R-package (Stekhoven and Buhlmann 2012).

Association analysis and model development

*We originally set out to identify prognostic biomarkers that could identify patients at risk of COVID-19 mortality already before the onset of the disease. As primary analysis, we performed time-to-event analysis using Cox regression (Enderlein 1987). Univariate analysis was conducted using age, sex, ethnicity, race, insurance status, and US region/division as covariate parameters for adjustment. We applied a Bonferroni-correction with the number of variables (m=249) to account for multiple testing and required a significance level of α=0.05/m=2*10^{-4}. The univariate associations were calculated for the entire patient cohort, as well as separately for the age groups <50, 50-70, 70-80 and >80 years (Supplementary Table S3).*

In order to allow comparison of hazard ratios (HRs) between different variables, we report the 2-standard-deviations hazard ratio "HR^{2SD}". It is computed as HR^{2SD} =HR$^{\{2*SD\}}$, where SD is the standard deviation of the respective variable.

*As secondary analysis, multivariable modelling was performed. We pursued two approaches in parallel. First, we performed a backward selection procedure on the Cox regression model of all eligible variables. We iteratively removed the variable with least impact on model performance until all remaining parameters were significant at α_1=0.05/249=2*10^{-4} (Bonferroni-correction). By construction, the procedure controls the family-wise error rate at α=0.05. In parallel, we derived a regularized Lasso model (Simon et al. 2011). We fitted a L1 (Lasso) regularized Cox-Proportional Hazards Model using glmnet version 3.02 (Simon et al. 2011), with the concordance index (C-index) (Steck, Krishnapuram, and Dehing-Oberije 2008) as the performance measure. The regularization parameter λ was optimized using ten-fold cross-validation. We selected λ such that we extracted the most regularized model with a C-index within one standard error of the best performing model.*

More details on model assumption checking, calibrations and performance measures can be found in Supplementary Methods.

Results

Real World Data source

Our real world data set covers the de-identified electronic health records of 122,250 adult COVID-19 patients collected by Optum® from the U.S. healthcare providers and diagnosed between February 20 and July 1, 2020. The overall mortality in this data set is 5.5%, which is well in line with the Case-Fatality Ratio estimate of 5.89% for USA (Hoffmann and Wolf 2020). Full details on the data set are provided in the Methods and Supplementary Methods sections.

Univariate analysis of predisposition

We set out to identify prognostic biomarkers that could identify patients at risk of COVID-19 mortality already before the onset of the disease. We first pursued this by a univariate analysis of 249 clinical variables that were observed at least a month before the infection. We selected a conservative buffer time

period of one month, since we expected that some patients might have an undiagnosed COVID-19 infection for days before an opportunity to get tested. We also wanted to exclude the effect of any early physiological changes incurred by the virus during the incubation period, which may be up to 14 days (Lauer et al. 2020). We used the median value captured between 1 and 12 months before the initial COVID-19 diagnosis for vitals and laboratory measurements, and the entire past medical history (median length ~4 years) for variables with a long-term effect, such as diagnoses of chronic indications. The univariate associations were calculated for the entire data set, as well as for the age groups <50, 50-70, 70-80 and >80 years separately (Supplementary Table S3).

The univariate analysis revealed 127 variables significantly associated (P<0.0002) with mortality (Supplementary Table S3). As expected, age was the strongest prognostic factor, with a per-year hazard-ratio (HR) of 1.08 [1.077;1.084], and a two-standard-deviation hazard-ratio HR^{2SD} of 17.8 [16.3;19.5]. The HR^{2SD} measure can reflect the risk increase between, for instance, patients of age 50 versus 90 (=2SD difference in age). Complementary to previous reports on biomarkers measured during COVID-19, we observed *a priori* decreases of albumin, hemoglobin, calcium, HDL cholesterol, lymphocyte proportion and lymphocyte to leukocyte ratio more frequently in deceased patients. Moreover, non-survivors had a higher likelihood of a history of elevated blood urea nitrogen, creatinine, blood glucose, respiratory rate, red cell distribution width, neutrophil percentage, neutrophil-to-lymphocyte ratio, and total white blood cell count.

An unexpected observation from the univariate analysis was that low, rather than high, measurements of diastolic blood pressure were associated with death (Supplementary Table S3, HR^{2SD}=0.711[0.655;0.771], P=1.43E-16; Figure 1A). This seemed to be in sharp contrast to earlier reports suggesting that hypertension was an important risk factor for COVID-19 (Zuin et al. 2020). Curiously, *a priori* diagnoses of both hypotension (ICD-10 I95.1, I95.9) and hypertension (I.10*, O.100.*, O.109.*) were more common among the deceased patients (HR=1.18 [1.14;1.22], P=8.47E-24 and HR=1.21 [1.14;1.29], P=1.33E-09, respectively). We envisaged that these counterintuitive results might be associated with the natural decrease of DBP with age, or hypotension being a lagging comorbidity of heart failure, together with age and heart failure being risk factors for severe COVID-19 (Williamson et al. 2020). However, after excluding all patients with a prior diagnosis of heart failure and performing an age-group-specific analysis, we still observed an association of low DBP and mortality (Figure 1B). Specifically, patients with a hypertension diagnosis (ICD-10 I.10*) and older than 40 years only showed consistently higher mortality rates, if they had also experienced hypotensive levels of DBP (<60 mmHg) in the past year. In fact, even hypertensive median measurements of DBP (>90 mmHg) were associated with lower mortality than hypotensive ones in the age group 70-79 years (Figure 1C). Indeed, the large epidemiological OpenSAFELY study (Williamson et al. 2020) has also previously reported that the fully adjusted variable of hypertension/elevated blood pressure has a negative rather than positive correlation with COVID-19 mortality, specifically in patients that are 70 years or older.

Another factor that could contribute to the hypotension observation is a decrease in serum albumin levels (Høstmark, Tomten, and Berg 2005). Decreased albumin levels were strongly associated with COVID-19

mortality in our data set, having the strongest effect for any laboratory measurement in the univariate analysis (HR^{2SD}=0.471 [0.442;0.503], P=5.10E-110). We also investigated the association of *a priori* measurements of hypotensive DBP and median albumin levels, and observed that patients with hypotensive DBP typically had low albumin (OR^{2SD}=0.18, P=7.01E-14). We conclude that many COVID-19 victims in our data set have a history of co-occurring hypotension and hypoalbuminemia.

Multivariate analysis findings

To further dissect the predisposition landscape of COVID-19 mortality, we pursued a multivariate time-to-event analysis (Cox regression) of laboratory, vital, comorbidity, immunisation and demographic variables in 55,757 patients, for whom at least ten variables were available (see Supplementary Methods). We performed three multivariate analyses in parallel: one for comorbidities, one for laboratory measurements and vitals, and a combined main model for all variable groups as well as inoculations. Unless otherwise stated, we are discussing the combined multivariate model below.

The combined multivariate Cox model reaches a C-index of 0.853 (SE 0.003), a high level of prognostic power for survival (Supplementary Table S4). Both alternative models also showed good performances, with C-indices 0.841 (0.003) for the comorbidity model, and 0.851 (0.003) for the labs/vitals model. The combined model reveals a plausible combination of factors that are well aligned with previous studies (Williamson et al. 2020) and associated with age, male sex, renal impairment, diabetes, hypoxia, hematological challenge, dementia and cancer (Table 1). As repeatedly reported, the strongest predictors of mortality are age (HR^{2SD}=6.95 [6.11;7.91]) and male sex. We also confirmed independent effects of African ethnicity and insurance status, as previously observed and discussed (Yehia et al. 2020). Furthermore, while increases in red blood cell distribution width (RDW) during COVID-19 had already previously been reported to be associated with mortality (Foy et al. 2020), our predisposition model shows that RDW aberrations precede the infection by at least a month.

The multivariate model indicates an independent protective effect for patients having received a booster vaccination for diphtheria, tetanus and pertussis (NDC code 58160084252). Furthermore, multiple other vaccinations were associated with better survival in the univariate analysis (Supplementary Table S3).

The most striking novel finding of the multivariate model is a comorbidity group of diagnoses associated with fluid, pH and electrolyte imbalance (FPEI) (Table 1). The overall FPEI comorbidity group has not been previously reported to be associated with COVID-19 mortality to our best knowledge. However, given the fundamental role of ACE2 in maintaining electrolyte homeostasis, such a dependency seems intuitively credible.

Electrolyte and pH disturbances of any kind might involve impairment of renal function. That said, while the effect of renal comorbidities was also evaluated for the multivariate analysis, it did not show an independent effect in the combined model. An investigation of the co-dependence of different variables revealed that renal comorbidities were excluded from the model by correlated but stronger independent effects of age, albumin, and blood urea nitrogen measurements, as well as FPEI diagnoses and red blood

cell distribution width (Supplementary Table S5). For comparison, our comorbidity-only multivariate model showed that the effects of FPEI comorbidities was independent to that of renal comorbidities (Supplementary Table S6). Interestingly, FPEI had the highest odds ratio and the lowest P value in the comorbidity multivariate model (HR=1.37 [1.27;1.49], and P=1.71E-14). We conclude that while FPEI disorders in COVID-19 patients may often indicate suboptimal function of the kidneys, they do not always co-occur with a diagnosed renal comorbidity.

A priori levels of DBP did not make it to the multivariate model as an independent variable. However, another dissection of variable co-dependence showed that it was not the history of heart failure that sequestered DBP from the model. Instead, the largest correlating factors that reduced DBP's effects were age, albumin, red cell distribution width and blood urea nitrogen (Supplementary Table S5). Finally, the hypertension comorbidity group did not make it to any of the multivariate models, either. Instead, the most correlated variates that progressed to the model in its stead were age, albumin, blood urea nitrogen, FPEI and Hemoglobin A1c (Supplementary Table S5).

Fluid, pH and electrolyte imbalance (FPEI) co-morbidity group deep dive

To further characterize the FPEI finding, we performed additional analyses involving both the prior medical history and the observations during the COVID-19 infection. First, we inspected the overlap of patient cohorts with prior FPEI and/or renal comorbidities (Table 2, Figure 2). While the overlap was substantial, it was far from complete. Less than half of the FPEI patients in our data set also had a renal comorbidity diagnosis. Interestingly, having a history of both renal and FPEI comorbidities was associated with a very high mortality, 20%. In fact, a quarter of all non-survivors in the entire data set had a combination of FPEI and renal diagnoses. However, an additional 18% of all deceased patients had an FPEI diagnosis but no renal diagnoses prior to the COVID-19 infection.

We repeated the univariate and multivariate analyses by including individual ICD codes in the FPEI comorbidity group (Supplementary Table S2). Two FPEI diagnoses made it to the comorbidity multivariate model: acidosis (ICD-10 E87.2), and mixed disorder of acid-base balance (ICD-10 E87.4) (Supplementary Table S3). Moreover, univariate analyses of all individual FPEI ICD codes showed a significant association with mortality (Supplementary Table S7). Lowest P values were observed for acidosis (E87.2), hyperkalemia (E87.5) and hypo-osmolality/hyponatremia (E87.1).

We also investigated the frequency of disturbances in the levels of electrolytes or pH in the thirty days following the onset of COVID-19. We distinguished incidences of low or high potassium, sodium, chloride, or total CO2 and used thresholds for normal ranges from Healthline ("Healthline Basic Metabolic Panel" 2017). Non-survivors were more likely than survivors to have any of the eight aberrations in the non-FPEI-diagnosed cohort, and more likely to have any aberration bar low potassium and low chloride in the FPEI-diagnosed cohort (Figure 3A-H). Furthermore, high potassium as well as both low and high sodium, chloride, and total CO2 were more frequent in non-survivors with a prior FPEI diagnosis than in non-survivors without one (Figure 3A-H). For comparison, the difference in the electrolyte levels of the sub-cohorts with or without any observed electrolyte imbalance during the 30 days after diagnosis were

significant in all eight categories for the non-FPEI-diagnosed patients, and in all bar low sodium and high total CO2 for the FPEI-diagnosed patients (Figure 3I-P) To summarize, our analysis shows that both elevated and decreased levels of basic electrolytes as well as pH-balance-associated total CO2 are more frequent in COVID-19 non-survivors than survivors.

Finally, we inspected the frequency of FPEI diagnoses assigned to the patients during the course of the COVID-19 infection. We plotted the proportion of patients having received an FPEI diagnosis on day -7… +30 relative to the original COVID-19 diagnosis date. We observed that the proportion of FPEI diagnosed patients throughout the followup period was significantly higher among the non-survivors (Figure 4A). Furthermore, we distinguished FPEI diagnoses to subgroups associated with volume/fluid depletion, fluid overload, or aberrations of pH. Each of these subcategories is more often observed in non-survivors (Figure 4B-D). Hence FPEI diagnoses associated with physiological manifestations of fluid depletion, overload, or pH balance disturbance alike were more frequent during COVID-19 in non-survivors than survivors.

Our observations prompted us to hypothesize that patients with *a priori* aberrances of fluid, pH and electrolyte management have an elevated dependency on the RAAS, and the sequestration of ACE2 by COVID-19 further undermines their overall homeostatic balance. This is compatible with the prior observations of renal comorbidities as a risk factor for COVID-19, but it is important to note that most FPEI-deficient patients in our data set lack an actual renal diagnosis. To summarize, we propose that FPEI is a risk factor for COVID-19 mortality both in absence of and coupled with pre-existing renal comorbidities.

Discussion

Despite all advances in COVID-19 vaccines and treatments, there is still an urgent need to better understand the course of disease and the risk factors of severe COVID-19. We have analysed a rich real world data set of 122,250 COVID-19 patients and characterized the pre-existing clinical factors that put the patients at risk of death. We foresee that our findings could generate novel hypotheses for COVID-19 treatment options and preventive medicine. The concordance of COVID-19 mortality prognostic factors implicated by our multivariate analysis and those from existing literature such as the large epidemiological OpenSAFELY study is high. This demonstrates that analysing real world data to dissect COVID-19 predisposition is a powerful approach, and confirms that risk factors in European and American patients have a considerable overlap. Moreover, some variables such as red blood cell distribution width, for which an association during the infection had previously been reported (Foy et al. 2020), were shown by our study to deviate from normal levels at least a month before the diagnosis (Table 2).

The multivariate and univariate models suggest multiple surprising associations. First, a previous inoculation with Diphtheria-Tetanus-Pertussis booster is independently associated with lower risk of mortality in the combined multivariate model. Curiously, herpes zoster vaccinations also show evidence

of a protective effect in the univariate analysis, whereas a past diagnosis of herpes zoster is associated with lower mortality in the comorbidity multivariate model (Supplementary Tables S1, S5). Generally, it can be hypothesized that patients with active inoculation schedules have a keen interest in maintaining their health, and the financial means to pursue this goal. However, compelling computational evidence has been presented that the DTP vaccine harbors a number of T cell epitopes with potential to induce cross-reactive immunity towards SARS-CoV-2 (Reche 2020). Second, the multivariate model suggests an independent association of low height with death (Table 1). Indeed, prospective observational studies have previously indicated that height has an inverse correlation with mortality from respiratory disease (Smith et al. 2000). Intriguingly, the initial total volume administered to a patient according to the ARDS protocol is a direct derivative of height, with shorter patients receiving a smaller initial volume of supplementary air by default ("NIH NHLBI ARDS Clinical Network Mechanical Ventilation Protocol Summary" 2008). One could also hypothesize that taller patients receive a lower initial viral load when exposed to droplets during the transmission, assuming a vertical position. Further studies will be necessary to identify clinical implications of these associations.

The most prominent novel outcome of our multivariate and univariate analyses is that prior fluid, pH and electrolyte imbalance (FPEI) is an independent predisposing factor for COVID-19 mortality. Out of various FPEI axes, those of metabolic acidosis seem strongest, complementing previous observations of diabetic ketoacidosis during COVID-19 (Li et al. 2020). However, all types of prior FPEI aberrations demonstrate an association with death in our univariate analysis and measurements of electrolyte levels during the infection.

Notably, ACE2 that is the entry point of SARS-CoV-2 into human cells is also an antagonist of ACE in the renin-angiotensin-aldosterone system (RAAS), which regulates electrolyte homeostasis, blood pressure and cardiovascular health (Patel et al. 2016), and restores balance upon volume disturbance of extracellular fluid (Fountain and Lappin 2020; Navar 2014). A large proportion of active ACE2 normally resides in the lung alveoli, while abundant presence is also observed in the heart, gastrointestinal tract, nasopharyngeal region and vascular endothelium, and limited expression in many other tissues including brain, liver and kidney (Hamming et al. 2004; J. Chen et al. 2020). Such widespread expression of ACE2 may also partly explain the variety of COVID-19 clinical symptoms.

The antagonistic effects of ACE and ACE2 are largely achieved by increasing or decreasing the amount of circulating Angiotensin II, respectively. Ang II is a potent secretagogue of aldosterone, an adrenal cortex hormone that stimulates kidney principal cells and alpha intercalated cells to enhance renal reabsorption of sodium and water, excretion of potassium, and the maintenance of acid-base balance (Atlas 2007; Wagner 2014). The homeostatic regulatory function of Ang II is largely endocrine, but autocrine, paracrine and intracrine effects have also been identified (Fountain and Lappin 2020). Notably, the half-life of Ang II in the bloodstream is around 30 seconds (van Kats et al. 1997), so its direct effects are short-lived, whereas aldosterone can function for hours or days.

The internalization of SARS-CoV-2 viruses via ACE2 restricts the receptor bioavailability on the cell surface, whereas the infected cells also actively downregulate ACE2 expression, as well as initiate cleavage of the membrane-bound ACE2 to produce a soluble form that is released into the circulation (Groß et al. 2020). Interestingly, the soluble form of ACE2 has previously been shown to be associated with heart failure independently of COVID-19 (Epelman et al. 2008). Since COVID-19 patients may have a reduced presence of membrane-bound ACE2 in infected tissues and potentially an abundance of soluble ACE2, it can be envisaged that the balancing effect of ACE2 on Ang II will be locally and systemically undermined by high SARS-CoV-2 viral load, evoking fluctuations in the levels of aldosterone, and consequently, plasma electrolytes. This molecular ambiguity of the collection of factors that we identify as FPEI has a single common denominator, ACE2 disturbance, with a straightforward and plausible connection with the disease.

We believe the role of all-round electrolyte and pH imbalance in COVID-19 could have gone partially unappreciated until now, owing to a focus on the median laboratory measurements (Lippi, South, and Henry 2020; D. Chen et al. 2020), which will obviously average out some of the effects of opposite extremities. However, recent reports have pointed out that both hypernatremia and hyponatremia are frequently present in COVID-19 patients (Hu et al. 2020). Our findings implicate that fundamental instability of the renin-angiotensin-aldosterone system (RAAS), manifested as earlier FPEI diagnoses, and potentially as aberrations of blood pressure levels, may sensitize patients to unfavourable outcomes in COVID-19.

The role of ACE2 receptors in SARS-CoV-2 and SARS-CoV infections has been frequently studied. Curiously, older individuals and Type II diabetics have lower baseline ACE2 levels despite being at higher risk of severe COVID-19 (J. Chen et al. 2020). It is tempting to hypothesize that a lack of ACE2 would manifest as an increase of Angiotensin II and aldosterone, and hence sodium and fluid retention, increasing overall fluid volume. However, our observations cover a comorbidity space much wider than that of hypernatremia. Rather, both elevated and reduced levels of sodium, potassium, chloride and total CO2 are significantly more frequently observed in COVID-19 non-survivors than survivors. The full consequences of the reduction and differential distribution of ACE2 in COVID-19 are yet to be understood. There are physiological feedback mechanisms counterbalancing the effects of RAAS based on the antidiuretic hormone and the natriuretic peptide, which may also become activated upon electrolyte or pH imbalance (Reid et al. 1983; Richards 1996) . A model whereby ACE2 disturbance always introduces an excess of fluids is likely a simplification. Instead, we propose that various pre-existing, and in some cases mutually exclusive, homeostatic aberrations may be magnified, when ACE2 and the RAAS are disordered by COVID-19. We also emphasize that an intact homeostatic regulation system is instrumental in preventing both excess and contracted plasma volumes, both of which are known risk factors for intensive care mortality (Perner 2009; Ogbu, Murphy, and Martin 2015).

Another striking finding of our predisposition models is that of hypoalbuminemia, which is also often associated with abnormally low levels of blood pressure and highly correlated with low diastolic blood pressure in our data set. Previous studies have indicated that a large majority of patients that are

critically ill with COVID-19, and almost all non-survivors, either present with hypoalbuminemia at the start of the infection or develop it during the course of the disease (W. Huang et al. 2020). Low serum albumin is also a predictor of overall mortality in the elderly (Corti 1994). Patients with low albumin will lose colloidal osmotic (oncotic) pressure, and experience an extracellular fluid shift from the intravascular space into the extravascular space (Gounden, Vashisht, and Jialal 2020). Notably, hypoalbuminemia coupled with excess extracellular fluid may result in co-occurring oedema and hypovolemia (Bobkova et al. 2016). Such a mechanism has also been suggested to contribute to COVID-19 (J. Huang et al. 2020), and would be particularly disastrous in the lung. Indeed, compelling suggestions have been presented that volume contraction may be aggravating COVID-19 outcomes (Tsolaki et al. 2020).

While previous discussions on the effect of ACE2 in COVID-19 have often focused on its role in regulating vasodilation and dampening inflammation (Groß et al. 2020), our findings complement this picture by showing that the RAAS-managed balance of fluid, pH and electrolytes is more frequently disturbed in COVID-19 victims than survivors already before the onset of the infection, and associated with an increased risk of death independently of age and renal comorbidities. Future observational and interventional studies may show that careful and personalized correction of fluid and electrolyte homeostasis early in the course of the infection correlates with better outcomes of COVID-19.

Limitations

Our analysis is based on a real world data resource and thus carries the burden of associated difficulties and limitations. In contrast to controlled randomized settings, systematic bias and negative effects by data errors cannot be completely ruled out *a priori*. On the other hand, the large sample size available can help to overcome issues and to detect phenomena, which are otherwise overlooked. Indeed, the majority of risk factors we identify are consistent with previous findings in the literature, thereby giving credibility to our novel findings. Furthermore, results for comorbidities from patient history are consistent with those of biomarkers of the respective diseases, demonstrating internal consistency and plausibility of our data. In summary, while the exact risk effect size might be influenced by the way the data was collected, it is likely that the qualitative conclusions we provide are correct.

Declarations

Author contributions

S.N. defined the research questions, supervised the analysis, and wrote most of the article. T.B. designed and performed the multivariate analysis and wrote the methods sections of the article. V.S. performed the statistical analyses for Figures 1, 2 and 3 and commented on the article. S.M. interpreted the clinical relevance of the findings and commented on the article. A.B.-M. reviewed the statistical findings and wrote parts of the article.

Competing interests

S. N. and A. B.-M. are Roche employees. A. B.-M. holds Roche stock options. T. B. and V. S. are contractors paid by Roche. S. M. is not paid by Roche.

Materials and correspondence

Upon requests, please reach out to satu.nahkuri@roche.com or anna.bauer-mehren@roche.com.

Data availability

The data presented is owned by Optum®, and access to it can be licensed.

Funding

This study was funded by Roche Holding AG. No grant number is applicable.

Code availability

All statistical analyses were implemented in R and are available on request.

References

Agency for Healthcare Research and Quality. 2018. “Elixhauser Comorbidity Software for ICD-10-CM Healthcare Cost and Utilization Project.” Agency for Journal of Statistical Software 23 Healthcare Research and Quality, Rockville, MD, USA. https://www.hcup-us.ahrq.gov/toolssoftware/comorbidityicd10/comorbidity_icd10.jsp.

Atlas, Steven A. 2007. “The Renin-Angiotensin Aldosterone System: Pathophysiological Role and Pharmacologic Inhibition.” *Journal of Managed Care Pharmacy* 13 (8 Supp B): 9–20. https://doi.org/10.18553/jmcp.2007.13.s8-b.9.

Bobkova, Irina, Natalia Chebotareva, Lydiya Kozlovskaya, and Evgenij Shilov. 2016. “Edema in Renal Diseases – Current View on Pathogenesis.” *Nephrology @ Point of Care* 2 (1): pocj.5000204. https://doi.org/10.5301/pocj.5000204.

Chen, Dong, Xiaokun Li, Qifa Song, Chenchan Hu, Feifei Su, Jianyi Dai, Yinghai Ye, Jianping Huang, and Xiaoming Zhang. 2020. “Hypokalemia and Clinical Implications in Patients with Coronavirus Disease 2019 (COVID-19).” Preprint. Infectious Diseases (except HIV/AIDS). https://doi.org/10.1101/2020.02.27.20028530.

Chen, Jiawei, Quanlong Jiang, Xian Xia, Kangping Liu, Zhengqing Yu, Wanyu Tao, Wenxuan Gong, and Jing-Dong J. Han. 2020. “Individual Variation of the SARS-CoV-2 Receptor ACE2 Gene Expression and Regulation.” *Aging Cell* 19 (7). https://doi.org/10.1111/acel.13168.

Corti, Maria-Chiara. 1994. “Serum Albumin Level and Physical Disability as Predictors of Mortality in Older Persons.” *JAMA: The Journal of the American Medical Association* 272 (13): 1036.

https://doi.org/10.1001/jama.1994.03520130074036.

Elixhauser, A, C Steiner, DR Harris, and RM Coffey. 1998. “Comorbidity Measures for Use with Administrative Data.” *Medical Care* 36 (1): 8–27.

Enderlein, G. 1987. “Cox, D. R.; Oakes, D.: Analysis of Survival Data. Chapman and Hall, London – New York 1984, 201 S., £ 12,–.” *Biometrical Journal* 29 (1): 114–114. https://doi.org/10.1002/bimj.4710290119.

Epelman, Slava, W.H. Wilson Tang, Stephen Y. Chen, Frederick Van Lente, Gary S. Francis, and Subha Sen. 2008. “Detection of Soluble Angiotensin-Converting Enzyme 2 in Heart Failure.” *Journal of the American College of Cardiology* 52 (9): 750–54. https://doi.org/10.1016/j.jacc.2008.02.088.

Fountain, John H., and Sarah L. Lappin. 2020. *Physiology, Renin Angiotensin System*. Treasure Island (FL): StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK470410/.

Foy, Brody H., Jonathan C. T. Carlson, Erik Reinertsen, Raimon Padros I. Valls, Roger Pallares Lopez, Eric Palanques-Tost, Christopher Mow, M. Brandon Westover, Aaron D. Aguirre, and John M. Higgins. 2020. “Association of Red Blood Cell Distribution Width With Mortality Risk in Hospitalized Adults With SARS-CoV-2 Infection.” *JAMA Network Open* 3 (9): e2022058. https://doi.org/10.1001/jamanetworkopen.2020.22058.

Gounden, Verena, Rishik Vashisht, and Ishwarlal Jialal. 2020. *Hypoalbuminemia*. Treasure Island (FL): StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK526080/.

Groß, Sonja, Christopher Jahn, Sarah Cushman, Christian Bär, and Thomas Thum. 2020. “SARS-CoV-2 Receptor ACE2-Dependent Implications on the Cardiovascular System: From Basic Science to Clinical Implications.” *Journal of Molecular and Cellular Cardiology* 144 (July): 47–53. https://doi.org/10.1016/j.yjmcc.2020.04.031.

Gupta, Aakriti, Mahesh V. Madhavan, Kartik Sehgal, Nandini Nair, Shiwani Mahajan, Tejasav S. Sehrawat, Behnood Bikdeli, et al. 2020. “Extrapulmonary Manifestations of COVID-19.” *Nature Medicine* 26 (7): 1017–32. https://doi.org/10.1038/s41591-020-0968-3.

Hamming, I, W Timens, Mlc Bulthuis, At Lely, Gj Navis, and H van Goor. 2004. “Tissue Distribution of ACE2 Protein, the Functional Receptor for SARS Coronavirus. A First Step in Understanding SARS Pathogenesis.” *The Journal of Pathology* 203 (2): 631–37. https://doi.org/10.1002/path.1570.

“Healthline Basic Metabolic Panel.” 2017. https://www.healthline.com/health/basic-metabolic-panel#normal-results.

Hoffmann, Christian, and Eva Wolf. 2020. “Older Age Groups and Country-Specific Case Fatality Rates of COVID-19 in Europe, USA and Canada.” *Infection*, October. https://doi.org/10.1007/s15010-020-01538-w.

Høstmark, Arne T, Sissel E Tomten, and John E Berg. 2005. “Serum Albumin and Blood Pressure: A Population-Based, Cross-Sectional Study:” *Journal of Hypertension* 23 (4): 725–30. https://doi.org/10.1097/01.hjh.0000163139.44094.1d.

Hu, Weihua, Xinke lv, Chang Li, Yang Xu, Yiding Qi, Zhuheng Zhang, Mingxuan Li, et al. 2020. “Disorders of Sodium Balance and Its Clinical Implications in COVID-19 Patients: A Multicenter Retrospective Study.” *Internal and Emergency Medicine*, October. https://doi.org/10.1007/s11739-020-02515-9.

Huang, Jiaofeng, Aiguo Cheng, Rahul Kumar, Yingying Fang, Gongping Chen, Yueyong Zhu, and Su Lin. 2020. “Hypoalbuminemia Predicts the Outcome of COVID-19 Independent of Age and Co-morbidity.” *Journal of Medical Virology* 92 (10): 2152–58. https://doi.org/10.1002/jmv.26003.

Huang, Wei, Chenze Li, Zhiquan Wang, Hong Wang, Ning Zhou, Jiangang Jiang, Li Ni, Xin A. Zhang, and Dao-Wen Wang. 2020. “Decreased Serum Albumin Level Indicates Poor Prognosis of COVID-19 Patients: Hepatic Injury Analysis from 2,623 Hospitalized Cases.” *Science China Life Sciences* 63 (11): 1678–87. https://doi.org/10.1007/s11427-020-1733-4.

Kats, Jorge P. van, Larissa M. de Lannoy, A. H. Jan Danser, Jan R. van Meegen, Pieter D. Verdouw, and Maarten A. D. H. Schalekamp. 1997. “Angiotensin II Type 1 (AT1) Receptor–Mediated Accumulation of Angiotensin II in Tissues and Its Intracellular Half-Life In Vivo.” *Hypertension* 30 (1): 42–49. https://doi.org/10.1161/01.HYP.30.1.42.

Lauer, Stephen A., Kyra H. Grantz, Qifang Bi, Forrest K. Jones, Qulu Zheng, Hannah R. Meredith, Andrew S. Azman, Nicholas G. Reich, and Justin Lessler. 2020. “The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application.” *Annals of Internal Medicine* 172 (9): 577–82. https://doi.org/10.7326/M20-0504.

Li, Juyi, Xiufang Wang, Jian Chen, Xiuran Zuo, Hongmei Zhang, and Aiping Deng. 2020. “COVID-19 Infection May Cause Ketosis and Ketoacidosis.” *Diabetes, Obesity and Metabolism* 22 (10): 1935–41. https://doi.org/10.1111/dom.14057.

Lippi, Giuseppe, Andrew M South, and Brandon Michael Henry. 2020. “Electrolyte Imbalances in Patients with Severe Coronavirus Disease 2019 (COVID-19).” *Annals of Clinical Biochemistry: International Journal of Laboratory Medicine* 57 (3): 262–65. https://doi.org/10.1177/0004563220922255.

Navar, L. Gabriel. 2014. “Physiology: Hemodynamics, Endothelial Function, Renin–Angiotensin–Aldosterone System, Sympathetic Nervous System.” *Journal of the American Society of Hypertension* 8 (7): 519–24. https://doi.org/10.1016/j.jash.2014.05.014.

“NIH NHLBI ARDS Clinical Network Mechanical Ventilation Protocol Summary.” 2008. http://www.ardsnet.org/files/ventilator_protocol_2008-07.pdf.

Ogbu, Ogbonna C., David J. Murphy, and Greg S. Martin. 2015. “How to Avoid Fluid Overload:” *Current Opinion in Critical Care* 21 (4): 315–21. https://doi.org/10.1097/MCC.0000000000000211.

Patel, Vaibhav B., Jiu-Chang Zhong, Maria B. Grant, and Gavin Y. Oudit. 2016. “Role of the ACE2/Angiotensin 1–7 Axis of the Renin–Angiotensin System in Heart Failure.” *Circulation Research* 118 (8): 1313–26. https://doi.org/10.1161/CIRCRESAHA.116.307708.

Perner, Anders. 2009. “Diagnosing Hypovolemia in the Critically Ill*:” *Critical Care Medicine* 37 (9): 2674–75. https://doi.org/10.1097/CCM.0b013e3181ad77d8.

Quan, H., B. Li, C. M. Couris, K. Fushimi, P. Graham, P. Hider, J.-M. Januel, and V. Sundararajan. 2011. “Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries.” *American Journal of Epidemiology* 173 (6): 676–82. https://doi.org/10.1093/aje/kwq433.

Reche, Pedro A. 2020. “Potential Cross-Reactive Immunity to SARS-CoV-2 From Common Human Pathogens and Vaccines.” *Frontiers in Immunology* 11 (October): 586984. https://doi.org/10.3389/fimmu.2020.586984.

Reid, I.A., J. Schwartz, L. Ben, J. Maselli, and L.C. Keil. 1983. “Interactions Between Vasopressin and the Renin–Angiotensin System.” In *Progress in Brain Research*, 60:475–91. Elsevier. https://doi.org/10.1016/S0079-6123(08)64414-3.

Richards, A M. 1996. “The Renin-Angiotensin-Aldosterone System and the Cardiac Natriuretic Peptides.” *Heart* 76 (3 Suppl 3): 36–44. https://doi.org/10.1136/hrt.76.3_Suppl_3.36.

Simon, Noah, Jerome Friedman, Trevor Hastie, and Rob Tibshirani. 2011. “Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent.” *Journal of Statistical Software* 39 (5): 1–13. https://doi.org/10.18637/jss.v039.i05.

Smith, George Davey, Carole Hart, Mark Upton, David Hole, Charles Gillis, Graham Watt, and Victor Hawthorne. 2000. “Height and Risk of Death among Men and Women: Aetiological Implications of Associations with Cardiorespiratory Disease and Cancer Mortality.” *Journal of Epidemiology and Community Health (1979-)* 54 (2): 97–103.

Steck, H, B Krishnapuram, and C Dehing-Oberije. 2008. “On Ranking in Survival Analysis: Bounds on the Concordance Index.” In Platt JC, Koller D, Singer Y, Roweis ST (eds): Advances in Neural Information Processing Systems 20, Curran Associates, Inc.

Stekhoven, D. J., and P. Buhlmann. 2012. “MissForest–Non-Parametric Missing Value Imputation for Mixed-Type Data.” *Bioinformatics* 28 (1): 112–18. https://doi.org/10.1093/bioinformatics/btr597.

Tsolaki, Vasiliki, George E Zakynthinos, Konstantinos Mantzarlis, and Dimosthenis Makris. 2020. “Increased Mortality among Hypertensive COVID-19 Patients: Pay a Closer Look on Diuretics in

Mechanically Ventilated Patients." *Heart & Lung* 49 (6): 894–95. https://doi.org/10.1016/j.hrtlng.2020.06.009.

Wagner, Carsten A. 2014. "Effect of Mineralocorticoids on Acid-Base Balance." *Nephron Physiology* 128 (1–2): 26–34. https://doi.org/10.1159/000368266.

Williamson, Elizabeth J., Alex J. Walker, Krishnan Bhaskaran, Seb Bacon, Chris Bates, Caroline E. Morton, Helen J. Curtis, et al. 2020. "Factors Associated with COVID-19-Related Death Using OpenSAFELY." *Nature* 584 (7821): 430–36. https://doi.org/10.1038/s41586-020-2521-4.

Yehia, Baligh R., Angela Winegar, Richard Fogel, Mohamad Fakih, Allison Ottenbacher, Christine Jesser, Angelo Bufalino, Ren-Huai Huang, and Joseph Cacchione. 2020. "Association of Race With Mortality Among Patients Hospitalized With Coronavirus Disease 2019 (COVID-19) at 92 US Hospitals." *JAMA Network Open* 3 (8): e2018039. https://doi.org/10.1001/jamanetworkopen.2020.18039.

Zuin, Marco, Gianluca Rigatelli, Giovanni Zuliani, Alberto Rigatelli, Alberto Mazza, and Loris Roncon. 2020. "Arterial Hypertension and Risk of Death in Patients with COVID-19 Infection: Systematic Review and Meta-Analysis." *Journal of Infection* 81 (1): e84–86. https://doi.org/10.1016/j.jinf.2020.03.059.

Tables

Table 1: Combined multivariate model of *a priori* prognostic factors for COVID-19 mortality (n=55,757), complemented with univariate results and potential clinical associations.

Variable	Multivariate model, P	HR 2SD	LCL 2SD	UCL 2SD	Univariate model, P	HR 2SD	LCL 2SD	UCL 2SD	N	Possible clinical associations
Age	2.54E-191	6.95	6.11	7.91	0.00E+00	17.8	16.3	19.5	122250	complex
Male gender	7.31E-41	1.82	1.66	1.98	2.39E-100	1.76	1.67	1.86	122250	complex
Albumin	1.56E-18	0.731	0.682	0.784	5.10E-112	0.471	0.442	0.503	46191	renal
Red cell distribution width	7.29E-17	1.34	1.25	1.44	4.86E-82	1.84	1.73	1.96	50295	hematological challenge
Insurance status: Uninsured	1.14E-16	1.26	1.19	1.33	2.69E-17	1.21	1.16	1.26	122250	NA
Blood urea nitrogen	5.21E-14	1.33	1.23	1.43	2.58E-67	1.9	1.77	2.04	53846	renal
Respiratory rate	6.43E-13	1.46	1.32	1.62	5.00E-43	2.21	1.97	2.47	51169	hypoxia
Fluid, pH and electrolyte disorders (FPEI)	9.67E-12	1.3	1.21	1.41	3.97E-76	1.51	1.45	1.58	122250	homeostasis
African-American race	1.30E-11	1.29	1.2	1.38	3.27E-14	1.22	1.16	1.29	122250	complex
Dementia	1.23E-08	1.15	1.09	1.2	8.84E-22	1.17	1.13	1.21	122250	neurology
Hemoglobin A1C	1.44E-08	1.23	1.15	1.33	4.28E-15	1.48	1.35	1.64	27169	diabetes
Metastatic carcinoma	1.49E-08	1.15	1.09	1.2	5.02E-11	1.11	1.08	1.15	122250	cancer
Insurance status: Medicare	2.96E-08	1.23	1.14	1.32	4.14E-10	1.16	1.11	1.22	122250	NA
Oxygen saturation	1.41E-07	0.838	0.785	0.895	7.99E-10	0.818	0.767	0.872	64154	hypoxia
Height	2.09E-07	0.79	0.723	0.864	6.89E-04	0.832	0.748	0.925	61694	pulmonary disadvantage
Moderate or severe liver disease	1.46E-06	1.11	1.07	1.16	7.77E-27	1.17	1.14	1.21	122250	hepatic
Triglycerides	8.04E-06	1.21	1.11	1.31	4.78E-09	1.38	1.24	1.53	31566	diabetes
Carbon dioxide total	8.70E-06	0.867	0.814	0.923	5.43E-15	0.763	0.713	0.817	51259	homeostasis
Congestive heart failure	1.08E-05	1.14	1.08	1.21	1.97E-58	1.34	1.29	1.38	122250	cardiovascular
Boostrix DTP vaccine	1.77E-05	0.843	0.78	0.912	8.91E-15	0.784	0.737	0.833	122250	cross-reactive immunity

Table 2: Mortality rates of patient cohorts with or without a prior history of FPEI and/or renal comorbidities.

	Neither FPEI nor renal history	Both FPEI and renal	Renal only	FPEI only	Total
Survivors	91903 (79.5%)	6967 (6.0%)	3603 (3.1%)	13059 (11.3%)	115532
Non-survivors	3117 (46.4%)	1745 (26.0%)	546 (8.1%)	1310 (19.5%)	6718
Mortality	3.4%	25.0%	15.2%	10.0%	5.8%

Figures

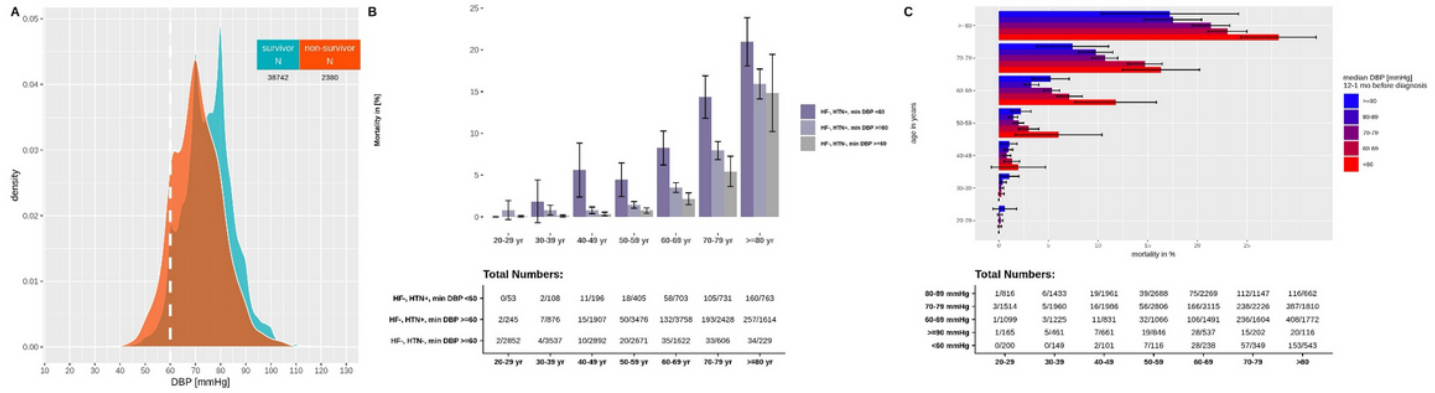

Figure 1

(A) Distribution of the median diastolic blood pressure measurement in COVID-19 survivors versus non-survivors during days -365…-31 before the COVID-19 diagnosis date. (B) Mortality rates of patients without a recorded history of heart failure (HF-) and either with or without a history of hypertension (HTN+ and HTN-, respectively). Cohorts had lowest diastolic blood pressure measurement either at <60 mmHg or >=60 mmHg on days -365…-31 before the COVID-19 diagnosis. Patients are distinguished by age groups, and 95% confidence intervals provided. (C) Mortality rates of patients as distinguished by age group and median diastolic blood pressure measurement levels on days -365...-31 before the COVID-19 diagnosis. 95% confidence intervals are provided.

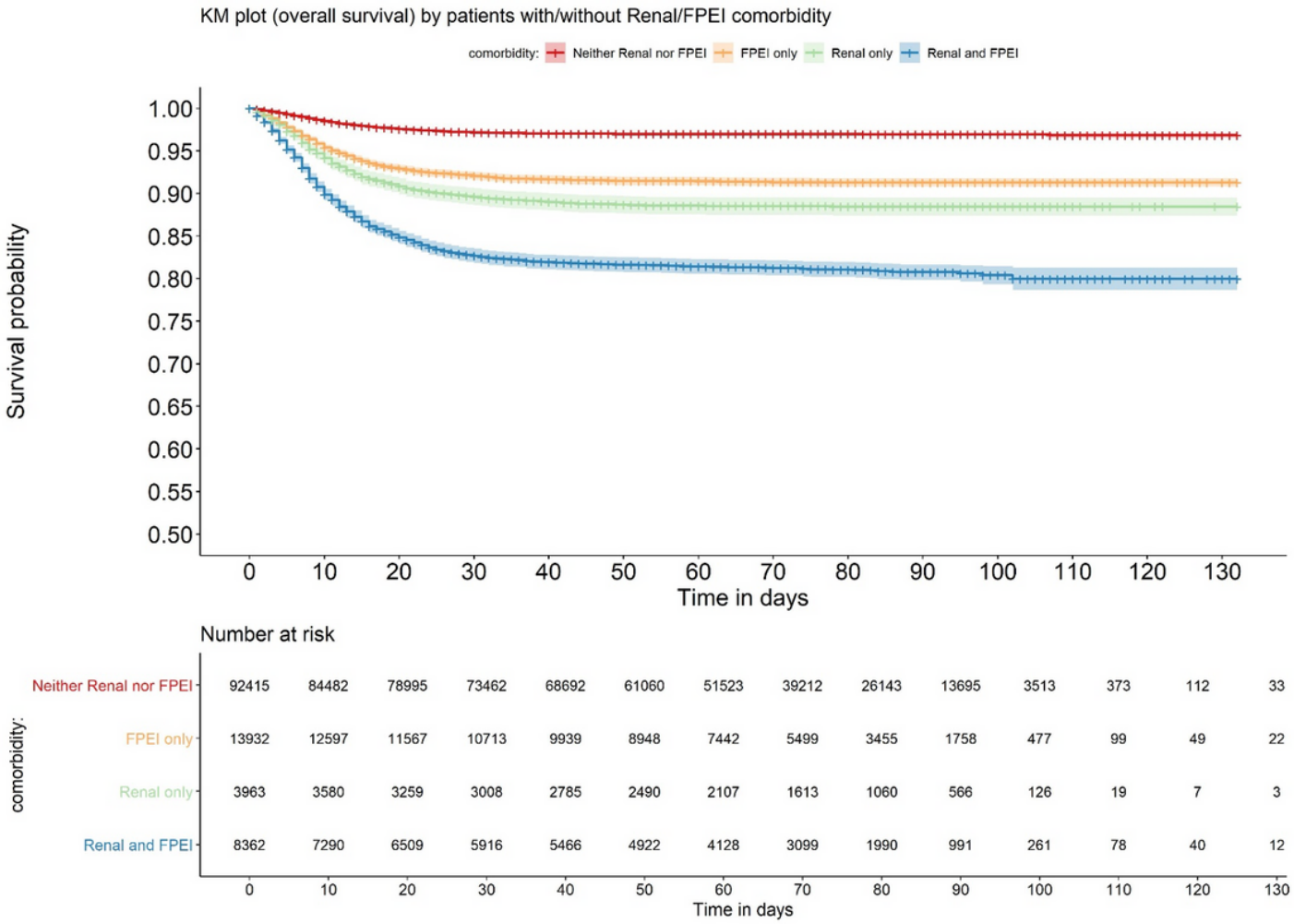

Figure 2

Mortality of patients with a priori comorbidities associated with renal disorders and/or fluid, pH and electrolyte imbalance (FPEI).

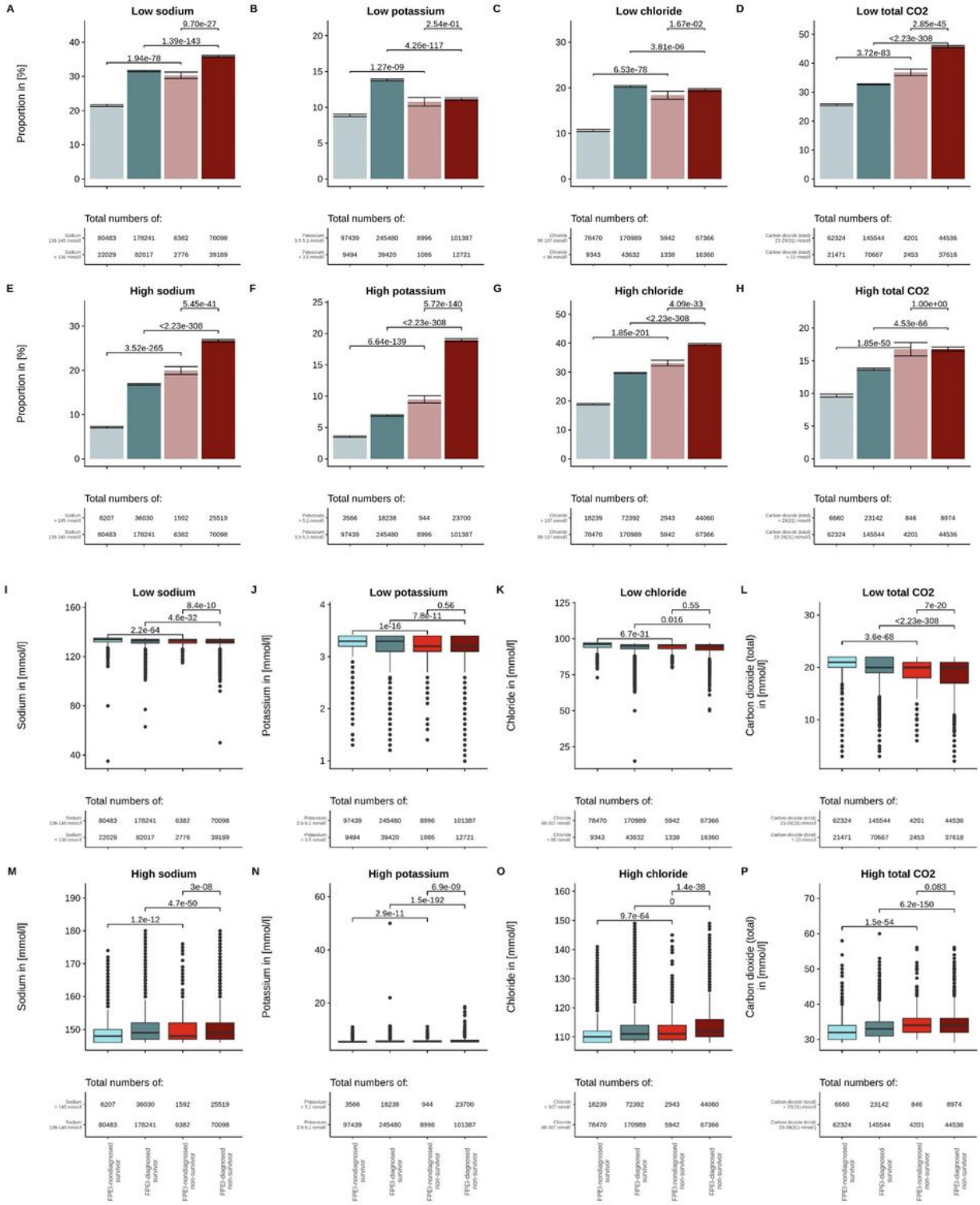

Figure 3

(A)-(H) Proportions of patients featuring abnormally high or low levels of potassium, sodium, chloride, or total CO2 in the thirty days following the COVID-19 diagnosis date. Fisher's Exact test used to compare categories, Bonferroni correction applied. (I)-(P) Boxplots of electrolyte levels among patients featuring abnormally high or low levels potassium, sodium, chloride, or total CO2 in the thirty days following the COVID-19 diagnosis date. Wilcoxon-tested comparison of location shift, Bonferroni correction applied.

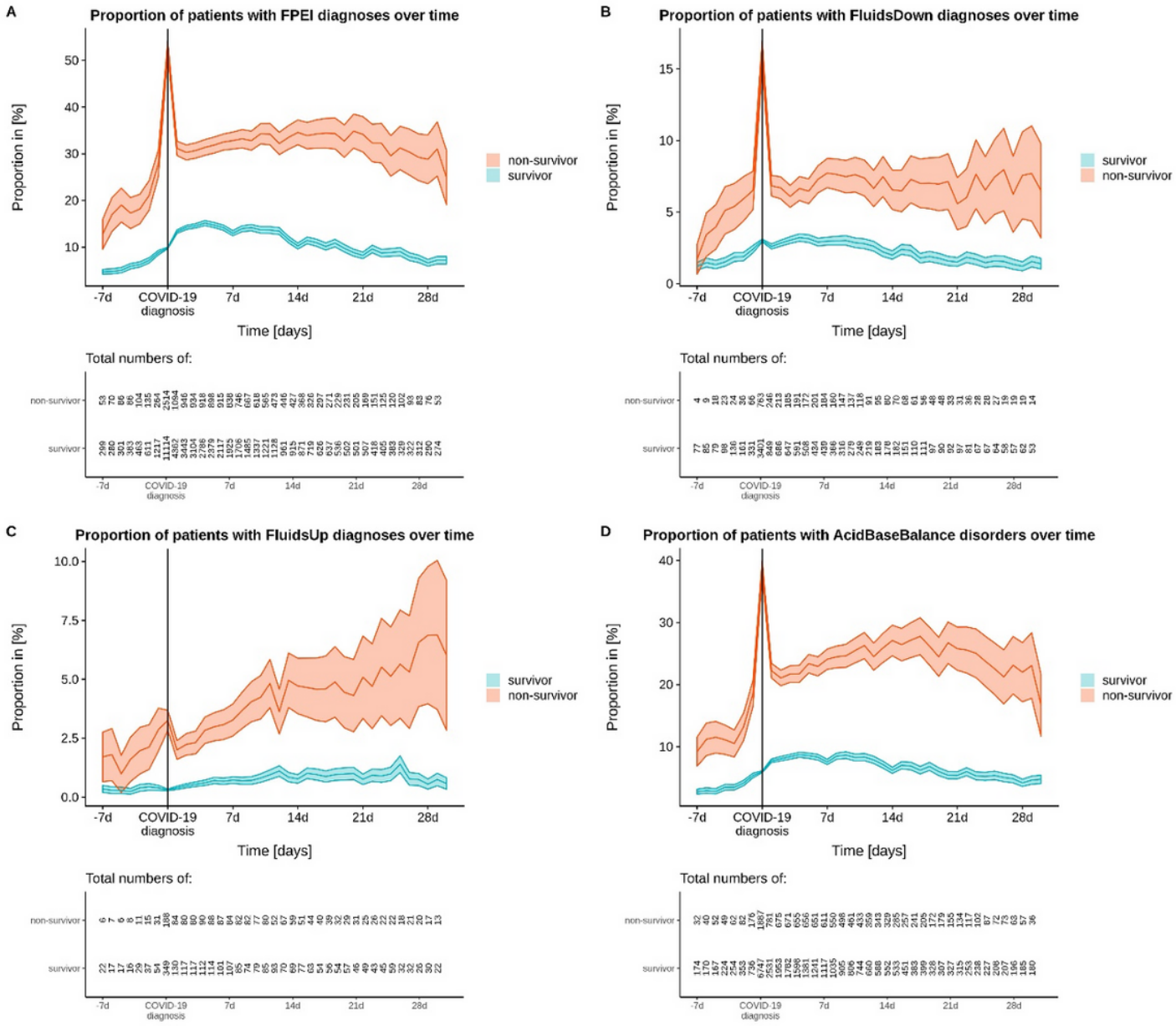

Figure 4

Time series of the proportion of survivors versus non-survivors that were assigned diagnoses associated with (A) any FPEI (B) fluid/volume deficit (C) fluid/volume overload or (D) acid/base imbalance. The ICD-10 codes that were used were (B) E86.0, E86.1 or E86.9, (C) E87.70 or E87.79, and (D) E87.2, E87.3, E87.4, E87.5 or E87.6. Plotted over days -7…+30 with respect to the COVID-19 diagnosis date. Non-survivors with estimated death date before day +7 were excluded.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NahkurietalSupplementaryinformation.docx

- NahkuriEtAlSupplementaryTable1.xlsx
- NahkuriEtAlSupplementaryTable2.xlsx
- NahkuriEtAlSupplementaryTable3.xlsx
- NahkuriEtAlSupplementaryTable4.xlsx
- NahkuriEtAlSupplementaryTable5.xlsx
- NahkuriEtAlSupplementaryTable6.xlsx
- NahkuriEtAlSupplementaryTable7.xlsx