

RESEARCH ARTICLE

Osmolality as a strong predictor of COVID-19 mortality and its possible links to other biomarkers

Sirin Cetin¹, Ayse Ulgen^{2,3*}, Hakan Sivgin⁴, Meryem Cetin⁵, Wentian Li^{6,7*}

1 Department of Biostatistics, Faculty of Medicine, Amasya University, Amasya, Türkiye, **2** Department of Mathematics, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom, **3** Department of Biostatistics, Faculty of Medicine, Girne American University, Karmi, Cyprus, **4** Department of Internal Medicine, Faculty of Medicine, Tokat Gaziosmanpaşa University, Tokat, Türkiye, **5** Department of Medical Microbiology, Faculty of Medicine, Amasya University, Amasya, Türkiye, **6** Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, New York, United States of America, **7** The Robert S. Boas Center for Genomics and Human Genetics, The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, New York, United States of America

* ayshe.ulgen@global.t-bird.edu; wqli2012@gmail.com, wentian.li@stonybrook.edu



OPEN ACCESS

Citation: Cetin S, Ulgen A, Sivgin H, Cetin M, Li W (2025) Osmolality as a strong predictor of COVID-19 mortality and its possible links to other biomarkers. PLoS One 20(9): e0331344. <https://doi.org/10.1371/journal.pone.0331344>

Editor: Youhua Tan, Hong Kong Polytechnic University, CHINA

Received: April 8, 2025

Accepted: August 14, 2025

Published: September 16, 2025

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0331344>

Copyright: © 2025 Cetin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The data used in this paper can be downloaded from: <https://github.com/wlicol/osmolality>.

Abstract

Osmolality, concentration of solute particles, was rarely used for prognosis for COVID-19. By analyzing blood samples of more than 1300 COVID-19 patients from Tokat, Turkey (including 100 surviving and 30 deceased inpatients), we found calculated osmolality to be an excellent prognostic biomarker for mortality and significantly associated with hospitalization, independent from gender and age. Although calculated osmolality is defined as a weighted sum of sodium, glucose, and urea, the three are not necessarily independent. Other blood test biomarkers, ferritin, creatine, and chloride are also correlated with osmolality after conditioning on age. By applying a combination of collider analysis and mediation analysis, we design a pipeline to construct a causal model among all these variables in their relationship to osmolality. We confirm that while glucose and sodium are independent contributors of osmolality, glucose and urea, urea and sodium are correlated. We also conclude that ferritin and creatine are associated with osmolality through urea, and chloride's association to osmolality is through sodium.

Introduction

Many risk factors have been identified for COVID-19 severity/hospitalization and mortality [1–3]. It is of particular interest to use routine blood test at the time of hospital admission to predict the COVID-19 disease prognosis [4,5]. The best established risk factors that are well accepted now for severity and mortality are age [6–9] and co-morbidities [10,11] such as diabetes, hypertension, and obesity [12–17].

The severity and mortality of COVID-19 is associated with a large number of risk factors. In our previous work on COVID-19 patients from Turkey, 10~30 biomarkers are found to be associated with time-to-death or time-to-release data [18]. With a larger number of risk

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

factors, one approach is to use the top factors where the ranking is based on multiple criteria [18], and another approach is to combine multiple factors into a composite score [19].

Osmolality (either measured or calculated [20,21]) or osmolarity are measure of the number of osmotically active molecules dissolved in the blood (per unit weight or per unit volume). Some of these molecules are the by-products (waste) of certain metabolic process, and their level may reflect biochemical abnormality. There are several formulas for calculated osmolality [22,23]. The formula we use here is a simple linear combination of three measurements: sodium, urea (a nitrogenous breakdown product), and glucose [24] plasma osmolality = $2 \times \text{sodium} + \text{urea}/5.99 + \text{glucose}/18$. An alternative of this formula is: plasma osmolality = $2 (\text{Sodium}) + (\text{BUN})/2.8 + (\text{Glucose})/18$, but BUN (blood urea nitrogen) is approximately half ($28/60 = 1/2.14$) of level of urea, because of nitrogen content in urea has a molecular weight of 28 out of the total molecular weight of 60. Other similar formula can be found in [25,26]. In this paper, we will use term osmolality to imply the calculated osmolality. Sodium apparently contributes the most numerically to osmolality (96.5% in our data), followed by urea (1.9%) and glucose (1.5%).

All components of osmolality – sodium, urea, and glucose, are known to be significant biomarkers for COVID-19 severity/mortality [18]. Our previous study to rank factors from time-to-death, time-to-release, and both, result in ranks of #1, #11, #4 for urea, #8, #12, #9 for glucose, and #9, #20, #15 for sodium [18]. Other publications have also pointed out links between COVID-19 disease severity and hypernatremia/hyponatremia, high urea and hyperglycemia (e.g., [27–32]). It is reasonable to assume that osmolality itself could be an important risk factor for COVID-19. Osmolality had been proposed as a prognosis biomarker for other diseases such as stroke [33], heart failure [34–37], pulmonary embolism [38]. However, its application to COVID-19 patients was limited in the literature (see, e.g., [39,40]).

In this paper, we first establish that osmolality is one of the best prognostic biomarkers for COVID-19. We then use osmolality as the focus of our attention to link several other factors that are also associated with COVID-19 severity. At a first glance, as sodium, urea, and glucose are the only components in the definition of calculated osmolality, there could be no “indirect path”, using a jargon from the mediation analysis, to be explained in the Method section, from other factors to be involved in osmolality. In order to untangle the links among a set of correlated variables, we design a pipeline utilizing collider analysis (see, e.g., [41] and Sect 6.2 of [42]) and mediation analysis (see, e.g., [43,44]) (both will be discussed in the Method section), to assign causal arrows for any three-variable at a time. Using the pipeline, we first decide the relationship between the three components for osmolality: sodium, urea, and glucose. Then we decide which of these three components are more likely to be linked to other variables that contribute to COVID-19 severity: ferritin, creatine, and chloride, by another round of collider analysis and mediation analysis as described in the pipeline.

Data and methods

COVID-19 patient data

The COVID-19 patient data was collected from Tokat State Hospital, Tokat, Turkey, from March 30, 2020 to December 1, 2020. The data were accessed from 01/05/2021 to 01/07/2022 for research purposes. There are $n = 1323$ patients, with 364 male and 959 female. The median age is 41. The COVID-19 status was determined by nasopharyngeal swab samples tested by RT-PCR. Inpatients were all admitted to hospital according to the WHO confirmation guideline. Subjects were fully anonymized before we accessed them. Verbal consent was informed from the patients at the medical visit. No minors were recruited in the study. This is a retrospective study and all data are fully anonymized before accessing them.

The following patients were excluded from this dataset: 1. those with hematological malignancies, immunodeficiency states, chronic kidney disease, heart failure and those on renal replacement therapy; 2. those who received osmotic agents like mannitol prior to the first blood draw with RT-PCR confirmed in hospitalized COVID-19 patients from the study; 3. those with age 17 or younger.

As a retrospective study, the patient data were extracted from the medical case records. The blood samples were collected within 24 hours of hospital visit/admission. This study was carried out with the approval of Tokat Gaziosmanpasa University Faculty of Medicine Non-Interventional Clinical Research Ethics Committee (decision No: 21-KAEK-104)

The following blood test factors are used in this paper: urea, ferritin, creatine, calcium, sodium, chloride, potassium, glucose, vitamin D, lactate dehydrogenase (LDH), C-reactive protein (CRP), aspartate aminotransferase (AST), alanine aminotransferase (ALT). There were more blood test factors that are not included here because of a large proportion of missing data, in particular, various immune cell counts (missing rates more than 80%), D-dimer (missing rate 59%), activated partial thromboplastin time (APTT, missing rate 47%), and fibrinogen (37%). In this study, all patients had routine blood samples taken, but approximately 5 = 11% of records in COVID-19 positive patients were missing LDH, CRP, and chloride results, due to unplanned gaps from changes in admission protocols, variation in clinical priorities, or retrospective record limitations. But importantly, test results of the three components of osmolality (sodium, potassium, glucose) – our primary target of analysis, are fully available. Although the SARS-CoV-2 variant was not tested in these samples, the dominating variant in Turkey in 2020 is alpha variant [45].

The data used in this paper can be downloaded from <https://github.com/wlicol/osmolality>.

Potential confounding factors not collected in our data

We do not have SARS-CoV-2 variant information for individual patients. The B.1.1.7 (alpha) variant was first reported in the UK in November 2020 and was known to spread fast in mid-December 2020 in UK [46,47] and elsewhere. In an independent investigation, two strains of SARS-CoV-2 seemed to dominate Turkish cases in 2020 [48]: the first was GISAID clade L which contains the original reference genome, and the second was GISAID clade GRY which is the alpha variant [48]. Because alpha variant increased disease severity risk either overall or within certain age range [49–51], a patient's variant status has potential to impact the analysis result.

However, we have only 4 inpatients (out of 130, or 3%) were collected in November–December 2020 period, and one of the 6 ICU patients were from the same period. Although this enrichment from inpatient to ICU may indicate an effect of the possible alpha variant at work, the overall number of samples involved is very small. The implication is that our result is very unlikely to be affected by having the variant information, as most would have the original strain of SARS-CoV-2.

All patients collected in our data, both inpatients and outpatients, were treated with favipravir antiviral medicine. All inpatients were treated by attending physicians according to the current national guidelines prepared by the Scientific Advisory Committee of the Turkish Ministry of Health. There were no biased treatments for inpatients that may impact the disease outcome. Therefore, we do not consider treatment status as a potential confounding factor.

Finally, COVID-19 vaccination has been proven extremely effective in preventing disease severity. Vaccination status is clearly a confounding factor. However, the first administered COVID-19 vaccine shot in Turkey was on January 13, 2021, and for health care workers only

(and not available for the general public). Since all our patients were collected in 2020, none of them were vaccinated.

Collider analysis and conditional analysis

We use partial/conditional correlation to measure correlation between two continuous variables (X and Y) after controlling the effect of the third variable (Z). We may use the notation $\text{cor}(X, Y | Z)$ for partial or conditional correlation. In causal graphs, if two independent causes (X and Y) collide on the outcome variable Z , we expect $\text{cor}(X, Y) = 0$ and $\text{cor}(X, Y | Z) \neq 0$ (often negative). The induced correlation between X and Y by conditioning on Z is the key part of a collider analysis.

Mediation analysis

Mediation analysis deals with the following question: if the independent variable X contributes to outcome variable Y both directly, or independently through a third mediation variable M , how to estimate the proportion of contributions in the two causal paths [44]. The classic mediation analysis assumes all three variables to be continuous, and the relationship between them can be summarized by three linear regressions (two of them are independent) [44]. There is also a more modern “causal mediation analysis” that can handle a variety of variable types [52]. A mediation analysis program would estimate the proportion (p_M) of contribution from X to Y through the indirect path, $X \rightarrow M \rightarrow Y$, the confidence intervals of p_M , and p -value for testing the null hypothesis that $p_M = 0$. The direct path $X \rightarrow Y$, is the one where the impact of X on Y does not go through M , and its proportion is $1 - p_M$.

Logistic regression

Four logistic regressions were carried out for each factor x with either mortality or hospitalization as the binary outcome y : (reg1) raw data $y \sim x$; (reg2) logistic regression using standardized $x' = (x - \text{mean}(x)) / \text{sd}(x)$ (where sd stands for standard deviation): $y \sim x'$; (reg3) conditional on age and gender $y \sim x + \text{age} + \text{gender}$; (reg4) logistic regression using log-transform and standardized $x'' = (\log(x) - \text{mean}(\log(x))) / \text{sd}(\log(x))$: $y \sim x''$.

Logistic regressions model the log-odds as a linear function of x , where odds is defined as $\text{prob}(y = 1) / (1 - \text{prob}(y = 1))$. Odds-ratio (OR) is a ratio of two odds evaluated at one unit apart, i.e., odds at $x_0 + 1$ and at x_0 (for reg1 and reg3). In reg2 model, since the independent variable is standardized, OR is the ratio of odds at $x_0 + \text{sd}(x)$ and at x_0 . In reg4 model, OR is the ratio of odds at $\log(x_0) + \text{sd}(\log(x))$ and at $\log(x_0)$, or equivalently, odds at $e^{\text{sd}(\log(x))} \cdot x_0$ and at x_0 .

ROC analysis

ROC is the abbreviation for receiver operating characteristic) [53,54]. When a continuous factor is used to predict or classify a medical outcome, a cutoff (a threshold) is chosen, partition the samples into a 2-by-2 count table, two correct predictions (true positives and true negatives) and two errors (false positives and false negatives)[55]. In ROC, the x -axis is the false positive rate (false positive counts divided by total number of negatives) and y -axis being true positive rate (true positive counts divided by total number of positives). The point $(x, y) = (0, 0)$ represents the most stringent threshold where all samples are called negative; whereas the point $(x, y) = (1, 1)$ represents the case where the threshold is so low that all samples are called positive. Gradually relaxing the threshold for calling positive results in a curve in a ROC plot (ROC curve). The area under a ROC curve is abbreviated as AUC. If a biomarker

lacks any prognostic ability, the curve is simply the diagonal line and AUC is 0.5. A larger AUC is an indication of a better prognostic marker.

R packages

The R package *ppcor* is used to calculate the partial correlation [56]. Conditional or correlation analysis have been used in other projects of ours in the past [57,58].

The R package *CMAverse* [59] (bs1125.github.io/CMAverse/) is used for causal mediation analysis. The function *cmest* is used and the *pm* (p_M : proportion of indirect impact) value, its confidence interval, and *p*-values, is reported.

The R package *pROC* [60] is used to calculate AUC, area under a ROC curve.

For ordinal regression, we use the *polr* function from MASS package [61]

(cran.r-project.org/web/packages/MASS/), and function *coefest* from the AER package [62] (cran.r-project.org/web/packages/AER/) to get *p*-value from a *polr* object.

Besides the *ppcor*, *CMAverse*, *pROC*, *MASS*, and *AER*, we use R (www.r-project.org) for other standard statistical analyses. *lm* is used for linear regression, *glm*($y \sim x$, *family* = "binomial") for logistic regression (and *glm*($y \sim x + \text{gender} + \text{age}$, *family* = "binomial") for multiple regression with gender and age as covariates), *cor*(x, y , *method* = "spearman") for Spearman correlation coefficient.

A lower *p*-value implies that the data is unlikely to be produced from a null hypothesis (e.g., no correlation, regression coefficient is zero, etc.). How small a *p*-value can we claim the test result to be "significant"? We follow the argument in [63–65] to use the *p*-value = 0.001 or 0.005 as the cutoff threshold.

Results

Osmolality is one of the best prognostic biomarkers for COVID-19 mortality

Of the $n = 1323$ patients, $n = 1193$ were outpatients and $n = 130$ were hospitalized (including 6 in ICU). Then, out of $n = 1323$ patients, $n = 30$ deceased (one of them died in ICU). Because the number of samples in ICU is too small (less power for a statistical analysis), we focused on deceased or hospitalized patients as the contracting group with the rest.

We test a series of logistic regressions for outcome (death, or hospitalization, or both) by using the original value, using the standardized value (which wouldn't affect the *p*-value, but affect the regression coefficient, thus odds-ratio), adding gender and age as covariates (which may make a factor less significant as a risk factor because age may take away some of the risk signal), using log-transformed values (because many factors are better normally distributed only after a log transformation).

Since *reg1*, *reg3*, *reg4* all have different meanings for OR, the definition of OR depends on the scale as well as the distribution of the variable. It would be inappropriate to log transform a factor, and use the corresponding OR, if it actually follows a normal distribution; on the other hand, it is inappropriate to use the raw data if the factor follows a log-normal distribution. Since the factors in our data follow different distributions, it is unclear which OR (from which version of logistic regression) should be used to compare the effect size of different factors. To deal with this problem, we use the AUC of ROC as a measure of effect size.

P-value is a measure of statistical significance and is not comparable between studies with different sample sizes. However, when comparing different factors on the same outcome, *p*-values are comparable because the sample size is the same (except for some disadvantage

for LDH, chloride, and CRP as these have slightly lesser samples). AUC as a measure of prediction performance is comparable among factors. Therefore, we use the *p*-value from reg1 (raw data), from reg4 (log-transformed, plus conditional on age and gender), and from AUC of ROC, to construct three ranking lists of these factors. The three ranks are averaged which is used to organize Table 1 (smaller *p*-values and/or larger AUC are on the top). The top five factors (all have AUC larger than 90%) are urea, calcium, ferritin, osmolality, and LDH. The next group, all have AUC between 70% and high 80%, includes creatine, sodium, glucose, chloride, AST, CRP, and vitamin D. The last two factors, potassium and ALT, are not good prognostic markers for mortality.

Table 1 also shows that for many factors, reg4 is less significant than reg1. One important difference between reg4 and reg1 is that reg4 adjusts the contribution from the age variable. One may treat age as a surrogate to deterioration of body at cellular and systemic level. On the

Table 1. Measuring and Testing a factor's association with mortality by logistic regression and AUC.

factor	n(NA) (rate)	range	reg1 raw OR(95%CI) <i>p</i> -value	reg2 standardize OR (95%CI)	reg3 gender + age OR(95%CI) <i>p</i> -value	reg4 log, sd, gender + age OR(95%CI) <i>p</i> -value	ROC AUC (95%CI)	rank
age		(18 -90)	1.16 (1.11-1.2) 7.8E-14	10.3 (5.59-18.99)	NA	NA	NA	
urea	0	(8.84 -296.7)	1.04(1.03-1.05) 5.2E-18	2.55 (2.06-3.15)	1.02(1.015-1.03) 3.4E-7	3.18(2.21-4.59) 5.7E-10	0.97 (0.94-1)	1
calcium	5 (0.4%)	(5.97 -11.9)	0.06(0.032-0.11) 2.9E-19	0.169 (0.12-0.25)	0.14(0.075-0.26) 6.5E-10	0.34(0.24-0.48) 7.1E-10	0.96 (0.93-0.99)	2
ferritin	26 (2%)	(1 -2000)	1.003(1.003-1.004) 1.2E-19	2.31 (1.93-2.77)	1.003(1.002-1.004) 9.1E-10	7.86(3.91-15.79) 6.9E-9	0.95 (0.91-0.98)	3
osmolality	0	(259.8 -367.3)	1.13(1.1-1.16) 1E-19	3.25 (2.52-4.2)	1.08(1.05-1.11) 1.4E-8	2.16(1.66-2.81) 8.7E-9	0.92 (0.85-0.99)	4
LDH	168 (10.6%)	(98 -2701.2)	1.012(1.009-1.014) 8E-18	4.65 (3.28-6.6)	1.009(1.006-1.014) 2.4E-9	3.55(2.33-5.41) 4.2E-9	0.94 (0.91-0.98)	5
creatine	0	(0.3 -8.7)	5.24(2.43-8.01) 2.1E-14	2.1 (1.74-2.54)	2.64(1.66-4.21) 4.4E-5	2.01(1.55-2.6) 9.8E-8	0.86 (0.75-0.96)	6
sodium	0	(121.9 -165.8)	1.39(1.28-1.51) 1.5E-14	2.93 (2.23-3.86)	1.2(1.11-1.29) 4.5E-6	1.83(1.41-2.38) 5.3E-6	0.77 (0.64-0.90)	7
glucose	0	(35.9 -627.3)	1.01(1.007-1.013) 1.9E-10	1.72 (1.46-2.03)	1.007(1.004-1.011) 8.3E-5	1.73(1.31-2.28) 1.1E-4	0.88 (0.81-0.94)	8
chloride	122 (9.2%)	(87.5 -137.3)	1.3(1.2-1.4) 7.5E-10	2.64 (1.94-3.59)	1.19(1.1-1.29) 8.5E-6	1.98(1.47-2.66) 6.9E-6	0.70 (0.57-0.84)	9
AST	0	(1.4 -504.5)	1.02(1.015-1.03) 7.3E-8	1.57 (1.33-1.85)	1.03(1.02-1.04) 5.3E-5	2(1.47-2.73) 1E-5	0.75 (0.64-0.86)	10
CRP	179 (11.8%)	(0.1 -452)	1.02(1.01-1.03) 4.5E-7	1.71 (1.39-2.11)	1.01(1.004-1.02) 0.0038	2.21(1.46-3.34) 1.4E-4	0.80 (0.71-0.90)	11
vitaminD	0	(2 -93.3)	0.84(0.78-0.91) 2.5E-5	0.15 (0.063-0.36)	0.85(0.78-0.92) 9.2E-5	0.4 (0.26-0.61) 1.8E-5	0.77 (0.69-0.84)	12
potassium	3 (0.2%)	(2.8 -6.27)	0.72(0.31-1.72) 0.47	0.87 (0.6-1.26)	0.66(0.34-1.29) 0.22	0.77(0.58-1.03) 0.079	0.59 (0.43-0.74)	13
ALT	0	(1.4 -504.5)	1.006(1.001-1.01) 0.012	1.23 (1.05-1.44)	1.007(1.001-1.013) 0.032	1.1(0.79-1.56) 0.55	0.54 (0.41-0.67)	14

n(NA): number of samples without the factor's information (NA for "not available"); The missing rate below is n(NA) divided by n = 1323. range: numerical range of the factor's value; reg1: odds-ratio (OR) and 95% confidence interval (CI) and *p*-value in logistic regression model 1 (single-variable without any transformation); reg2: those results for model 2 (variables are standardized to have mean = 0, variance = 1); reg3: results for model 3 (single-variable without transformation, but with gender and age as two co-variables); reg4: results for model 4 (variable is log-transformed, then standardized, plus covariates of gender and age); AUC: area under curve of ROC (receiver operating characteristic); rank: based on the sum of three ranks from *p*-value in reg1, *p*-value in reg4, and AUC. In these three ranking orders, the number 1 rank is given to the factor with the smallest *p*-value, or the largest AUC. A smaller than 1E-3 = 0.001 *p*-value is marked in bold.

<https://doi.org/10.1371/journal.pone.0331344.t001>

other hand, even after an adjustment of age, many factors are still significant, which indicates that the general aging process is not the only reason that some patients have higher risk for mortality.

We also noticed from Table 1 that osmolality is ranked no.4, behind its component urea, but above its components sodium and glucose. Judging from the p -value for reg1/reg2 and reg3, osmolality is even better than urea. Whether osmolality is slightly below or above urea, we argue that osmolality is probably more robust than urea as a prognostic marker because it contains more information.

Osmolality is a good prognostic biomarker for COVID-19 hospitalization

Table 2 shows similar four logistic regressions as the last subsection, but with hospitalization as the outcome variable. The order of the factor is identical to that in Table 1. However, the last column also shows the ranking order with hospitalization (H) as outcome (average over the ranking from reg1, reg4, and AUC), plus reproducing the ranking order in Table 1 (D for death), as well as the average ranking order based on six ordering lists (three for mortality, three for hospitalization) (A for all). The osmolality is ranked no.4 in all three lists. A non-integer rank indicates a tie.

Interestingly, for hospitalization outcome, as well as for overall mortality and hospitalization, osmolality is ranked higher than all of its components (urea, sodium, glucose). Again, we speculate that it is due to the fact that osmolality contains more physiological information than each of its components.

Similar to Table 1, adjusting for gender and age may make an originally significant risk factor to be non-significant (e.g. glucose), or greatly reduce its significance. At the p -value of 0.001 level, only three factors are significant for age-conditioned (plus log-transformed) regression analysis: LDH, CRP, and calcium. Generally speaking, our factors are better prognostic markers for mortality than for hospitalization, which include the mortality as a subset.

Combining mortality and hospitalization by ordinal logistic regression

To combine the results from the previous subsections on mortality and hospitalization, one may also treat them as two level of severity of different degrees. Although it is possible to code outpatients, surviving inpatients and deceased inpatients as 0,1,2, and apply a linear regression, a better option is to use the ordinal regression where 0,1,2 are ordered factor levels, not numerical values.

Table 3 shows the p -value of ordinal logistic regression in three versions: raw data (reg1), raw data with age and gender as covariate (reg3), logarithm-transformed, then standardized factor value with age and gender as covariate (reg4). The second version of ordinal logistic regression: standardized factor value, produces the same (or essentially the same) p -value as reg1.

The result in Table 3 is consistent with that in Table 1 and Table 2, with calcium, LDH, ferritin, osmolality as the most significant factors for COVID-19 prognosis. CRP is more significant for hospitalization than mortality, whereas urea is more significant for mortality than hospitalization. Interestingly, two common known risk factors, vitamin D and glucose, are near the bottom of list: glucose loses its significance mostly by conditioning on age and vitamin D is only borderline significant.

Table 2. Measuring and testing a factor's association with hospitalization by logistic regression and AUC.

factor	reg1	reg2	reg3	reg4	ROC	rank
	raw	standardize	gender + age	log,sd,gender + age		
	OR(95%CI)	OR(95%CI)	OR(95%CI)	OR(95%CI)	AUC	(D,H,A)
	<i>p</i> -value		<i>p</i> -value	<i>p</i> -value	(95%CI)	
urea	1.028(1.02-1.035) 7.4E-14	1.92(1.62-2.27)	1.009(1.002-1.016) 0.011	1.25(1.02-1.51) 0.028	0.68 (0.63-0.73)	(1,6,5)
calcium	0.27(0.2-0.36) 4.5E-20	0.44(0.37-0.52)	0.53(0.4-0.71) 1.8E-5	0.68(0.57-0.81) 1.5E-5	0.70 (0.65-0.76)	(2,2,1)
ferritin	1.003(1.002-1.003) 1E-16	1.89(1.63-2.2)	1.001(1.001-1.002) 3.8E-6	1.44(1.12-1.85) 0.0041	0.71 (0.66-0.76)	(3,3,3)
osmolality	1.08(1.06-1.10) 4.3E-17	2.06(1.74-2.43)	1.03(1.01-1.05) 0.0018	1.3(1.1-1.54) 0.0017	0.68 (0.63-0.74)	(4,4,4)
LDH	1.008(1.006-1.009) 3.9E-18	2.75(2.19-3.46)	1.004(1.002-1.006) 4.5E-5	1.56(1.29-1.9) 6E-6	0.74 (0.69-0.79)	(5,1,2)
creatinine	3(2.12-4.26) 6E-10	1.64(1.4-1.91)	1.45(1.05-2.02) 0.026	1.21(1.03-1.42) 0.023	0.62 (0.56-0.67)	(6,8,6)
sodium	1.18(1.12-1.24) 3.1E-10	1.71(1.45-2.02)	1.07(1.02-1.12) 0.0083	1.23(1.05-1.45) 0.0098	0.61 (0.55-0.68)	(7,7,7)
glucose	1.006(1.004-1.009) 5.1E-8	1.43(1.26-1.62)	1.001(0.998-1.004) 0.42	1.07(0.9-1.26) 0.46	0.67 (0.62-0.72)	(8,9,5,9)
chloride	1.01(0.96-1.06) 0.7	1.04(0.86-1.24)	1.02(0.98-1.06) 0.39	1.06(0.91-1.25) 0.46	0.53 (0.47-0.6)	(9,12,5,12)
AST	1.016(1.009-1.023) 1.2E-5	1.36(1.19-1.56)	1.01(1-1.02) 0.0069	1.21(1.02-1.43) 0.033	0.6 (0.54-0.66)	(10,9,5,10)
CRP	1.02(1.013-1.026) 2.1E-8	1.66(1.39-1.99)	1.01(1-1.02) 0.0041	1.63(1.31-2.03) 1.2E-5	0.69 (0.64-0.74)	(11,5,8)
vitaminD	0.98(0.96-1) 0.08	0.83(0.67-1.02)	0.98(0.96-1) 0.055	0.81(0.67-0.99) 0.037	0.55 (0.5-0.61)	(12,11,11)
potassium	1.14(0.75-1.72) 0.55	1.06(0.88-1.26)	0.78(0.52-1.16) 0.21	0.88(0.74-1.04) 0.14	0.5 (0.44-0.56)	(13,12,5,13)
ALT	1.003(0.999-1.006) 0.16	1.1(0.96-1.26)	1.002(0.997-1.006) 0.51	0.97(0.8-1.18) 0.76	0.53 (0.47-0.59)	(14,14,14)

Similar to Table 1, but changing the outcome variable from death/mortality to hospitalization status. A *p*-value is in bold if it is smaller than $1E-3 = 0.001$. The order of factors is identical to that in Table 1 (under D of the last column); however, we also list the ranking in terms of association with hospitalization (under H), and ranking from both mortality and hospitalization (under A). The H-ranking is based on the *p*-value from reg1, that from reg4, and AUC of ROC. If three ranking lists lead to a tie for H-ranking, a fractional rank is used (e.g., if two factors are tied at no.9, they are assigned to a rank of 9.5).

<https://doi.org/10.1371/journal.pone.0331344.t002>

Factors correlated with osmolality besides its components, either directly or after conditional on age

Since sodium, urea, and glucose are components of calculated osmolality, it is not surprising that these three quantities are highly correlated with osmolality (Table 4). The reason we also show results using outpatients only in Table 4 is to avoid the potential selection bias, which would be present if both a factor and osmolality are causes of the disease severity. Besides these, ferritin, creatine, ALT, LDH, CRP, AST, chloride are also significantly correlated with osmolality (though chloride loses the significance at 0.001 level if only outpatients are used). On the other hand, potassium, calcium, and vitamin D are not significantly associated with osmolality.

We use Spearman correlation in this calculation so that it is not required that these variables follow a normal distribution (another choice is to use the Pearson correlation on log-transformed variable values).

Table 3. P-values for testing a factor's association with both hospitalization and mortality by ordered/ordinal logistic regression.

factor	p-value		
	reg1	reg3	reg4
	raw	gender + age	log.sd,gender + age
calcium	4.6E-26	1.3E-9	4.6E-10
LDH	6.4E-23	8.1E-9	3.7E-10
ferritin	2.3E-23	1.5E-10	2.5E-5
osmolality	0	9.2E-8	5.1E-8
urea	6.4E-20	2.7E-6	1.5E-5
CRP	3.2E-10	9.5E-5	4.7E-7
sodium	3.6E-13	3.4E-5	5.4E-5
creatinine	8.4E-14	0.00014	6.3E-5
AST	1.9E-7	3.7E-5	0.00089
vitaminD	0.057	0.0095	0.003
chlorine	0.4	0.013	0.021
glucose	6.2E-9	0.091	0.1
potassium	0.57	0.16	0.073
ALT	0.12	0.2	0.96

Similar to Tables 1 and 2, but changing the outcome variable to an ordered categorical variable representing three levels of COVID-19 prognosis (outpatient, surviving inpatient, deceased inpatients). The second regression (reg2) using the standardized factor value either lead to an identical (or similar) *p*-value as the first regression (reg1) using the raw data, thus not included in the table. The factors are ranked by the mean of logarithm of *p*-values of the last two columns, from small (more significant) to large (not significant).

<https://doi.org/10.1371/journal.pone.0331344.t003>

Table 4. Spearman correlation (scc) between individual factors and osmolality.

factor	outpatients		all patients		all (cond. age)	
	scc	p-value	scc	p-value	scc	p-value
sodium (Na)	0.83	4.3E-308	0.83	0	0.85	0
urea	0.44	4.2E-57	0.46	2.4E-70	0.3	2E-28
age	0.42	7.5E-52	0.46	5.2E-69	NA	NA
glucose	0.35	1E-35	0.39	3.9E-48	0.22	8E-16
ferritin	0.3	7.5E-52	0.46	5.2E-69	0.3	4.9E-29
creatinine	0.27	1.2E-21	0.29	9.9E-27	0.17	1E-19
ALT	0.2	5.9E-12	0.18	1.7E-11	0.057	0.04
LDH	0.16	2.4E-7	0.23	1.3E-15	0.064	0.027
CRP	0.11	0.00022	0.15	5.6E-7	0.05	0.12
AST	0.099	0.00063	0.12	1.1E-11	0.014	0.61
chloride	0.094	0.0021	0.12	2.3E-5	0.23	2.2E-15
potassium (K)	0.064	0.028	0.051	0.066	-0.03	0.28
calcium (Ca)	-0.007	0.81	-0.08	0.0038	0.001	0.72
vitaminD	0.05	0.08	0.003	0.91	-0.04	0.13

outpatients: only the outpatients are used; all patients: all patients are used; all(cond. age): using all patients but scc is conditional (partial correlation) on age. The *p*-value is for testing scc = 0, and those *p*-values smaller than 0.001 are marked with bold font.

<https://doi.org/10.1371/journal.pone.0331344.t004>

The correlation between osmolality and other factors can be viewed by scatter plot (Fig 1). Deceased patients are labeled with red, hospitalized patients are labeled with blue, and the outpatients are in grey. From Fig 1, it can be seen directly that osmolality is high in many deceased patients (roughly 20 out of 30).

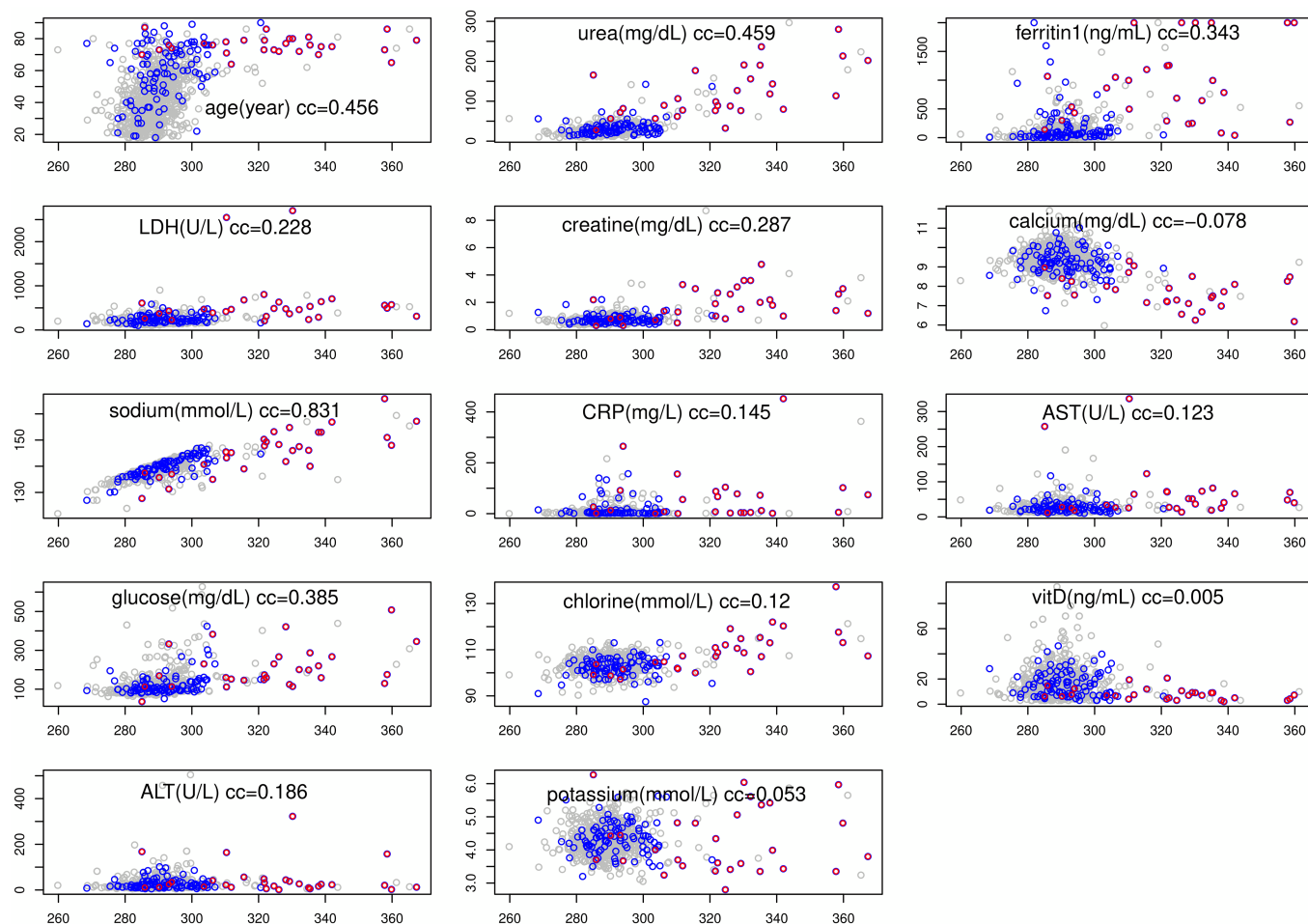


Fig 1. Scatter plots of a factor (y-axis) versus osmolality (x-axis). Outpatients, surviving inpatients, and deceased patients are marked with grey, blue, and red color. The cc is the Spearman correlation coefficient between the factor and osmolality calculated from all patients (see Table 4). The unit of osmolality is mOsm/kg or mOsm/L, and units for other factors are given in the plots.

<https://doi.org/10.1371/journal.pone.0331344.g001>

Since a correlation between two variables might be due to the fact that a third variable (confounder) is a cause of both variables. To exclude this possibility, we re-calculate the correlation between a factor and osmolality conditional on the age variable. For conditional correlation, only ferritin, creatine, and chloride are significant, whereas ALT, LDH, CRP, and AST lose their significance.

Pipeline for constructing a causal model for osmolality components and other associated test variables

Fig 2 shows our plan in proposing a causal model between osmolality and its components, as well as other associated variables. In the first step, all three components (glucose, urea, sodium) should have an arrow pointing to osmolality, simply by definition. However, it is unclear if there should be arrows among these components. Towards that, we pick two of the three components as X and Z , osmolality as Y , and use the collider analysis

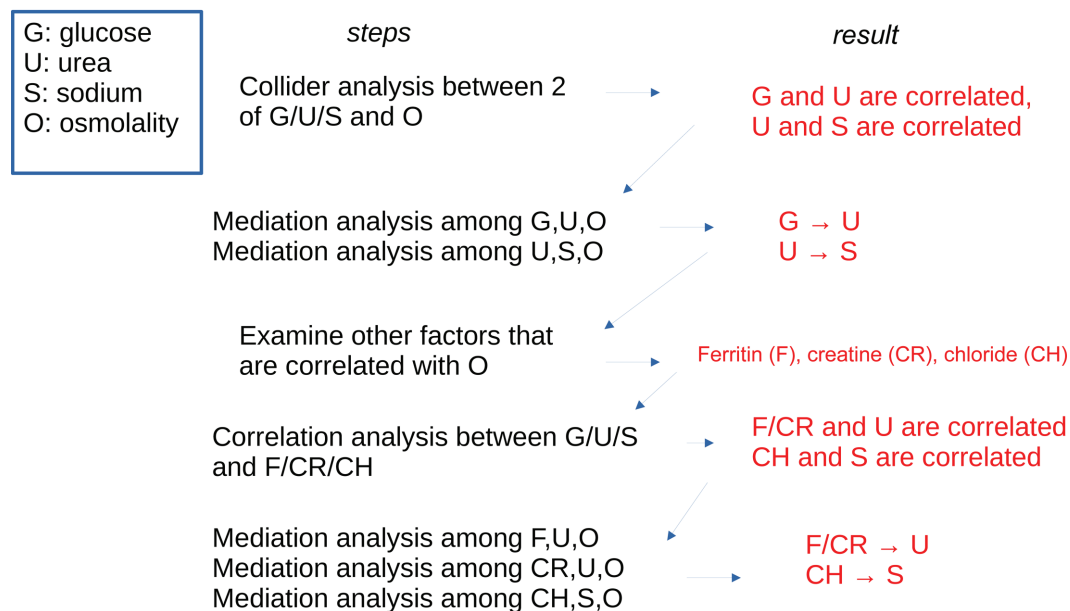


Fig 2. Pipeline for constructing a possible causal model between osmolality and its components as well as other associated variables.

<https://doi.org/10.1371/journal.pone.0331344.g002>

(correlation between X and Z , with or without conditional on Y [66]) to determine if the correlation between X and Z are completely due to their contribution to Y . If yes, there is no arrow between X and Z .

If two of the three components of osmolality are not independent, i.e., there is a causal arrow between the two, there is still a question of the direction of the arrow. For that we use mediation analysis. If the indirect path $X \rightarrow Z \rightarrow Y$ has a stronger evidence than the indirect path in another model, $Z \rightarrow X \rightarrow Y$, then we assign an arrow from X to Z .

Next, we bring in new variables to the causal model. These are the factors/variables that strongly contribute to the COVID-19 severity, but not part of the official components of osmolality. Let's denote one of these new variables Z^* . We check which one of three components of osmolality is most correlated with a Z^* : that component should be paired with that Z^* .

With the result from the above collider analysis, we assign the Y^* to an osmolality component X if these are not independent. The last step would be the determination of an arrow direction between X and Y^* . This can in principle be decided by a mediation analysis between X , Y^* , and $Y = \text{osmolality}$. But as we will see later, the result is obvious without the need for analysis.

Possible relationship between the three components of osmolality: sodium, urea, and glucose

Table 5 shows the conditional analysis (collider analysis) of any two components of osmolality and osmolality itself. Glucose and sodium are uncorrelated, but become correlated after conditioning on osmolality. This is consistent with the causal model that glucose and sodium both have a causal arrow pointing to osmolality, but there is no arrows between glucose and sodium. A negative correlation after conditioning on the collider is a common result: if the outcome Y receives a contribution from X , the contribution from Z becomes unnecessary,

Table 5. Collider analysis of the three components of osmolality.

pair	cor (Spearman)	conditional cor (Spearman)
glucose and urea	0.269 (pv = 7E-44)	0.11 (pv = 5E-5)
urea and sodium	0.136 (pv = 7E-7)	-0.5 (pv = 3E-84)
glucose and sodium	0.05 (pv = 0.07)	-0.53 (pv = 7E-95)

cor(Spearman): Spearman correlation between a pair of components of osmolality; conditional cor(Spearman): same Spearman correlation between two components of osmolality conditional on osmolality.

<https://doi.org/10.1371/journal.pone.0331344.t005>

and vice versa. Note that since our continuous variables tend to be not normally distributed, a calculation of correlation is done by the non-parametric version of Spearman correlation.

The glucose-urea pair (and urea-sodium pair to a lesser extent) is not independent. It indicates there should be an arrow between glucose and urea (and between urea and sodium). The following mediation analysis is the attempt to resolve this issue (results are shown in Table 6 from the CMAverse program [59]). From Table 6, the evidence for glucose → urea in one causal model is stronger than the urea → glucose in another model. Therefore, we decide that there should be an arrow pointing to urea from glucose. Similar from Table 6, the evidence for urea → sodium is stronger than sodium → urea. These results are incorporated in the causal model in Fig 3.

Bringing in ferritin, creatine, and chloride into the causal model

From Table 4, there are only three variables besides the components of osmolality that are correlated with osmolality after conditional on age: ferritin, creatine, and chloride. If we consider sodium, urea, and glucose to be closer to osmolality, whereas ferritin, creatine and chloride

Table 6. Mediation analysis with osmolality as the outcome variable.

X	M	Y	p_M (CMAverse) (CI)	p-value
log(glucose)	log(urea)	log(osmo)	0.22 (0.15 -0.29)	8E-10
log(urea)	log(glucose)	log(osmo)	0.075 (0.047-0.11)	2E-7
log(urea)	log(sodium)	log(osmo)	0.22 (0.16-0.28)	2E-12
log(sodium)	log(urea)	log(osmo)	0.078 (0.05-0.1)	2E-9

p_M : Proportion of the contribution in mediation path (indirect path) from X to Y through M, the 95% confidence interval; p-value: the p-value for testing $p_M = 0$.

<https://doi.org/10.1371/journal.pone.0331344.t006>

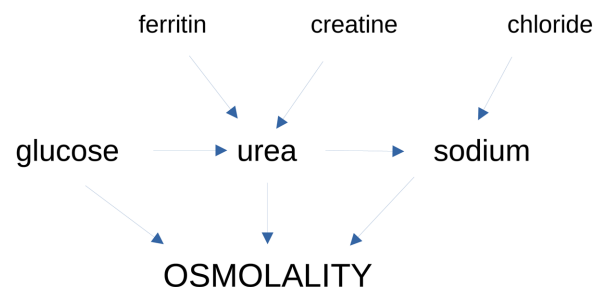


Fig 3. Proposed model of relationship between osmolality and its three components (sodium, urea, glucose), as well as other associated quantities (chloride, creatine, ferritin). Confounder-1 is a proposed confounder for creatine and urea (and probably ferritin). A candidate for confounder-1 is kidney dysfunction. Confounder-2 is a proposed confounder for sodium and chloride. A candidate for confounder-1 is salt level.

<https://doi.org/10.1371/journal.pone.0331344.g003>

to be further from osmolality, as described in the pipeline in Fig 2, we continue to use correlation and mediation analysis to decide where and how to put these new variables into the causal model.

Table 7 shows the Spearman correlation and p -values between ferritin, creatine, and chloride with the three components of osmolality. If we select the component with the highest correlation to be paired with the new variable, then ferritin is paired with urea, creatine is paired with urea also, and chloride is paired with sodium.

Take ferritin and urea pair for example, one could use them and osmolality to do a mediation analysis, with two causal models. Of these two causal models, we can determine the causal arrow between ferritin and urea (the two already have causal arrows towards osmolality). The mediation analysis indicates that there is an arrow from ferritin to urea. This result is actually obvious as urea is a component of osmolality and should be closer to osmolality by definition. All our results are summarized in Fig 3.

Discussions

Our dataset is imbalanced in gender, where the female to male ratio is almost 3 to 1. However, gender is not shown to be correlated with either mortality or hospitalization in this data. The Fisher test p -value for 2-by-2 count table formed by gender and mortality is 0.53. Similarly, the Fisher test p -value for gender and hospitalization is 0.92. As a result, the gender imbalance is not believed to cause any bias in our results. Furthermore, since our model reg4 adjusts both age and gender, even if there is a differential mortality/hospitalization risk in gender, it will be corrected in the reg4 model.

We have not collected a whole set of co-morbidities in this dataset, only the diabetes status. There are 71 patients (5.4%) labeled as diabetes, a relatively small sample size. Diabetic patients are significantly more likely to die ($12/71 = 17\%$ vs $18/1234 = 1.4\%$ for non-diabetic patients, p -value = $9E-9$), or admitted to hospital ($20/71 = 28\%$ vs $110/1142 = 8.8\%$ for non-diabetic patients, p -value = $5E-6$). How diabetes fit into the framework of high-osmolality causing COVID-19 severity will be a topic for future studies.

The non-traditional data analysis methods – collider analysis and mediation analysis – usually require three correlated variables. As long as there are multiple variables that are all pairwise correlated, these methods can be applied. We are not aware of previous attempts to apply these methods in the osmolality context, and our results should need independent check from other approaches. For collider analysis, it is assumed that confounding factors, if exist, do not impact the conclusion. Since it is difficult to assess this assumption, it is a limitation of the method. For mediation analysis, we only compare the relative (not absolute) strength of evidence for two causal models (in Table 6).

Table 7. Correlation between one component of osmolality and one new variable.

pair	cor(Spearman)
glucose and ferritin	0.32 (pv = $8E-32$)
urea and ferritin	0.41 (pv = $7.7E-54$)
sodium and ferritin	0.15 (pv = $3.2E-8$)
glucose and creatine	0.2 (pv = $1.5E-13$)
urea and creatine	0.54 (pv = $8.5E-100$)
sodium and creatine	0.1 (pv = $1.8E-4$)
glucose and chloride	-0.13 (pv = $5.8E-6$)
urea and chloride	-0.068 (pv = 0.02)
sodium and chloride	0.29 (pv = $3.1E-24$)

<https://doi.org/10.1371/journal.pone.0331344.t007>

Osmolality is not usually used as a prognostic marker. However, there have been isolated reports to support the idea of using osmolality to predict medical outcomes [37,67–70]. Using calculated osmolality as a prognostic biomarker for COVID-19 patients is even rare [39]. Even though each of the three components for calculated osmolality urea, glucose, and sodium, have been shown to be useful for predicting mortality from COVID-19 [71–76], the calculated osmolality has not been a focus of attention in hospital management of COVID-19 patients. The purpose of this paper is to show that osmolality itself is an excellent measure for COVID-19 prognosis purpose. In a sense, the combination of three good biomarkers should also be a good biomarker, but we point out the possibility that osmolality could be an even better biomarker than its individual components. Osmolality measure is particularly useful in distinguishing patients who are high in an individual component but not in other component, nevertheless, tend to not have metabolic problems.

The relevance of osmolality to disease severity might also be viewed from a physiological perspective. Osmolality combines status information from renal function, endocrine control, hydration balance, and metabolism. All these statuses will be disturbed during a multiple organ failure and systemic dysfunction. Although we only measured the osmolality at one time point, it could be interesting to monitor the trajectory of osmolality level. In some recent studies, the osmolality trajectories are determined for sepsis patients [77] or sepsis-associated encephalopathy patients [78], and those with increasing osmolality in time have the worst prognosis.

Although sodium, urea and glucose are three components of the calculated osmolality, it is usually ignored that these are not independent. By using mediation analysis, we ask the question of whether one of the components contribute to the calculated osmolality through other components indirectly. Our analysis shows that roughly 18% of the contribution from glucose is through urea, 22% of the contribution from urea is through sodium, and a small proportion, 7.8%, of the contribution from sodium is through urea.

To check whether there are other blood test measures to be associated with the calculated osmolality, even though these are not part of the official components, we calculate the non-parametric correlation (Spearman) conditional on age. The reason to condition on age is because age can be a common factor (confounder) to both a blood test measure and calculated osmolality [79] (see also Fig 1). This analysis led to only three other measurements: chloride, creatine, and ferritin.

Even though chloride is not part of the calculated osmolality in the formula we used, it is part of the measured osmolality [24]. Therefore, its association is not surprising. Our association analysis naturally paired chloride with sodium. It is not impossible that this association is due to the compound salt with both of them as components. Similarly, the creatine-urea could be reasonable as the two are both waste products of metabolic processes. Some kidney dysfunction could lead to a high level of both. Meanwhile, the link between ferritin and urea might be due to the role played by ferritin in regulating iron level and keeping kidney healthy.

A more severe disease leads to more damaged tissues, triggers more inflammation response. Some viruses infect lymphocytes and by increasing their destruction, it causes lymphopenia and a decrease in serum albumin, which is a negative acute-phase reactant, while it causes an increase in the levels of CRP and ferritin, which are positive acute-phase reactants. Published associations between ferritin and COVID-19 severity do not address the issue of causal direction [80,81].

We hope to compare our findings concerning osmolality-COVID19-severity association with other future data collected from other SARS-CoV-2 variants, as our data was mostly limited to the alpha-variant [45]. We hope our causal model for osmolality related variables in Fig 3 will be further examined by other lines of investigation in the future. We also hope that

future investigation of causal relation between osmolality and COVID-19 disease severity, either from a statistical causal inference or from biological studies, will provide hints for medical treatment. Furthermore, our work can be further improved by having more sample sizes and by collecting more potentially relevant confounders such as comorbidities.

Author contributions

Conceptualization: Sirin Cetin, Ayse Ulgen, Hakan Sivgin, Wentian Li.

Data curation: Sirin Cetin.

Formal analysis: Ayse Ulgen, Meryem Cetin, Wentian Li.

Investigation: Ayse Ulgen, Hakan Sivgin, Meryem Cetin, Wentian Li.

Methodology: Sirin Cetin, Ayse Ulgen, Wentian Li.

Project administration: Hakan Sivgin.

Resources: Sirin Cetin, Hakan Sivgin.

Software: Ayse Ulgen, Wentian Li.

Supervision: Meryem Cetin, Wentian Li.

Visualization: Meryem Cetin.

Writing – original draft: Sirin Cetin, Ayse Ulgen, Wentian Li.

Writing – review & editing: Sirin Cetin, Ayse Ulgen, Hakan Sivgin, Meryem Cetin, Wentian Li.

References

1. Li X, Xu S, Yu M, Wang K, Tao Y, Zhou Y, et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J Allergy Clin Immunol*. 2020;146(1):110–8. <https://doi.org/10.1016/j.jaci.2020.04.006> PMID: 32294485
2. Zhang J, Wang X, Jia X, Li J, Hu K, Chen G, et al. Risk factors for disease severity, unimprovement, and mortality in COVID-19 patients in Wuhan, China. *Clin Microbiol Infect*. 2020;26(6):767–72. <https://doi.org/10.1016/j.cmi.2020.04.012> PMID: 32304745
3. Chen R, Liang W, Jiang M, Guan W, Zhan C, Wang T, et al. Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in China. *Chest*. 2020;158:97–105.
4. Ferrari D, Motta A, Strollo M, Banfi G, Locatelli M. Routine blood tests as a potential diagnostic tool for COVID-19. *Clin Chem Lab Med*. 2020;58:1095–9.
5. Zeng F, Li L, Zeng J, Deng Y, Huang H, Chen B, et al. Can we predict the severity of coronavirus disease 2019 with a routine blood test?. *Pol Arch Intern Med*. 2020;130(5):400–6. <https://doi.org/10.20452/pamw.15331> PMID: 32356642
6. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*. 2020;20(6):669–77. [https://doi.org/10.1016/s1473-3099\(20\)30243-7](https://doi.org/10.1016/s1473-3099(20)30243-7)
7. Bonanad C, García-Blas S, Tarazona-Santabalbina F, Sanchis J, Bertomeu-González V, Fácila L. The effect of age on mortality in patients with COVID-19: a meta-analysis with 611,583 subjects. *J Am Med Dir Assoc*. 2020;21:915–8.
8. Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc Natl Acad Sci U S A*. 2020;117(36):22035–41. <https://doi.org/10.1073/pnas.2006392117> PMID: 32820077
9. Kang SJ, Jung SI. Age-related morbidity and mortality among patients with COVID-19. *Infect Chemother*. 2020;52(2):154–64. <https://doi.org/10.3947/ic.2020.52.2.154> PMID: 32537961

10. Sanyaolu A, Okorie C, Marinkovic A, Patidar R, Younis K, Desai P, et al. Comorbidity and its impact on patients with COVID-19. *SN Compr Clin Med*. 2020;2(8):1069–76. <https://doi.org/10.1007/s42399-020-00363-4> PMID: 32838147
11. Ejaz H, Alsrhani A, Zafar A, Javed H, Junaid K, Abdalla AE, et al. COVID-19 and comorbidities: deleterious impact on infected patients. *J Infect Public Health*. 2020;13(12):1833–9. <https://doi.org/10.1016/j.jiph.2020.07.014> PMID: 32788073
12. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584(7821):430–6. <https://doi.org/10.1038/s41586-020-2521-4> PMID: 32640463
13. Cuschiari S, Grech S. COVID-19 and diabetes: the why, the what and the how. *Journal of Diabetes and its Complications*. 2020;34(9):107637. <https://doi.org/10.1016/j.jdiacomp.2020.107637>
14. Schiffrin EL, Flack JM, Ito S, Muntner P, Webb RC. Hypertension and COVID-19. *Am J Hypertension*. 2020;33:373–4.
15. Alberca RW, de Mendonca Oliveira L, Branco ACCC, Pereira NZ, Sato MN. Obesity as a risk factor for COVID-19: an overview. *Crit Rev Food Sci Nutrition*. 2020;61:2262–76.
16. Landstra CP, de Koning EJP. COVID-19 and diabetes: understanding the interrelationship and risks for a severe course. *Front Endocrinol (Lausanne)*. 2021;12:649525. <https://doi.org/10.3389/fendo.2021.649525> PMID: 34220706
17. Buscemi S, Corleo D, Randazzo C. Risk factors for COVID-19: diabetes, hypertension, and obesity. *Adv Exp Med Biol*. 2021;1353:115–29. https://doi.org/10.1007/978-3-030-85113-2_7 PMID: 35137371
18. Ulgen A, Cetin S, Cetin M, Sivgin H, Li W. A composite ranking of risk factors for COVID-19 time-to-event data from a Turkish cohort. *Comput Biol Chem*. 2022;98:107681. <https://doi.org/10.1016/j.compbiolchem.2022.107681> PMID: 35487152
19. Schoning V, Liakoni E, Baumgartner C, Exadaktylos AK, Hautz WE, Atkinson A, et al. Development and validation of a prognostic COVID-19 severity assessment (COSA) score and machine learning models for patient triage at a tertiary hospital. *J Transl Med*. 2021;19:56.
20. Bhagat CI, Garcia-Webb P, Fletcher E, Beilby JP. Calculated vs measured plasma osmolalities revisited. *Clin Chem*. 1984;30(10):1703–5. <https://doi.org/10.1093/clinchem/30.10.1703> PMID: 6537784
21. Kar E, Kocatürk E, Küskü Kiraz Z, Demiryürek B, Alataş İÖ. Comparison of measured and calculated osmolality levels. *Clin Exp Nephrol*. 2020;24(5):444–9. <https://doi.org/10.1007/s10157-020-01848-1> PMID: 31983046
22. Worthley LI, Guerin M, Pain RW. For calculating osmolality, the simplest formula is the best. *Anaesth Intensive Care*. 1987;15(2):199–202. <https://doi.org/10.1177/0310057X8701500214> PMID: 3605570
23. Khajuria A, Krahm J. Osmolality revisited—deriving and validating the best formula for calculated osmolality. *Clin Biochem*. 2005;38(6):514–9. <https://doi.org/10.1016/j.clinbiochem.2005.03.001> PMID: 15885229
24. Rasouli M. Basic concepts and practical equations on osmolality: biochemical approach. *Clin Biochem*. 2016;49(12):936–41. <https://doi.org/10.1016/j.clinbiochem.2016.06.001> PMID: 27343561
25. Martín-Calderón JL, Bustos F, Tuesta-Reina LR, Varona JM, Caballero L, Solano F. Choice of the best equation for plasma osmolality calculation: comparison of fourteen formulae. *Clin Biochem*. 2015;48(7–8):529–33. <https://doi.org/10.1016/j.clinbiochem.2015.03.005> PMID: 25794428
26. Ebonwu EO, Nagel SE, Repsold L, Pillay TS. Critical evaluation of equations for serum osmolality: proposals for effective clinical utility. *Clin Chim Acta*. 2020;510:79–87. <https://doi.org/10.1016/j.cca.2020.06.043> PMID: 32622969
27. Voets PJGM, Frölke SC, Vogtländer NPJ, Kaasjager KAH. COVID-19 and dysnatremia: a comparison between COVID-19 and non-COVID-19 respiratory illness. *SAGE Open Med*. 2021;9:2050312121102778.
28. Hirsch JS, Uppal NN, Sharma P, Khanin Y, Shah HH, Malieckal DA, et al. Prevalence and outcomes of hyponatremia and hypernatremia in patients hospitalized with COVID-19. *Nephrology Dialysis Transplantation*. 2021;36(6):1135–8. <https://doi.org/10.1093/ndt/gfab067>
29. Shrestha AB, Sapkota UH, Shrestha S, Aryal M, Chand S, Thapa S, et al. Association of hypernatremia with outcomes of COVID-19 patients: a protocol for systematic review and meta-analysis. *Medicine (Baltimore)*. 2022;101(51):e32535. <https://doi.org/10.1097/MD.00000000000032535> PMID: 36595846
30. Yin J, Wang Y, Jiang H, Wu C, Sang Z, Sun W, et al. Blood urea nitrogen and clinical prognosis in patients with COVID-19: a retrospective study. *Medicine (Baltimore)*. 2024;103(8):e37299. <https://doi.org/10.1097/MD.00000000000037299> PMID: 38394490

31. Khunti K, Del Prato S, Mathieu C, Kahn SE, Gabbay RA, Buse JB. COVID-19, hyperglycemia, and new-onset diabetes. *Diabetes Care*. 2021;44:2645–55.
32. Yang Y, Cai Z, Zhang J. Hyperglycemia at admission is a strong predictor of mortality and severe/critical complications in COVID-19 patients: a meta-analysis. *Biosci Rep*. 2021;41:BSR20203584.
33. Bhalla A, Sankaralingam S, Dundas R, Swaminathan R, Wolfe CD, Rudd AG. Influence of raised plasma osmolality on clinical outcome after acute stroke. *Stroke*. 2000;31(9):2043–8. <https://doi.org/10.1161/01.str.31.9.2043> PMID: 10978027
34. Rasouli M, Kiasari AM, Arab S. Indicators of dehydration and haemoconcentration are associated with the prevalence and severity of coronary artery disease. *Clin Exp Pharmacol Physiol*. 2008;35(8):889–94. <https://doi.org/10.1111/j.1440-1681.2008.04932.x> PMID: 18430066
35. Vaduganathan M, Marti CN, Mentz RJ, Greene SJ, Ambrosy AP, Subacius HP, et al. Serum osmolality and postdischarge outcomes after hospitalization for heart failure. *Am J Cardiol*. 2016;117:1144–50.
36. Kaya H, Yucel O, Ege MR, Zorlu A, Yucel H, Gunes H, et al. Plasma osmolality predicts mortality in patients with heart failure with reduced ejection fraction. *Kardiologia Polska*. 2017;75:316–22.
37. Wang S, Zhao D, Yang T, Deng B, Sun J, Gu L, et al. Association of serum osmolality with all-cause and cardiovascular mortality in US adults: a prospective cohort study. *Nutr Metab Cardiovasc Dis*. 2023;33(4):844–52. <https://doi.org/10.1016/j.numecd.2023.01.007> PMID: 36710117
38. Öz A, Çınar T, Hayiroğlu Mİ, Avşar S, Keskin M, Orhan AL. The predictive value of plasma osmolality for in-hospital mortality in patients with acute pulmonary embolism. *Clin Respir J*. 2019;13:174–83.
39. Ramesh J, Rajesh M, Varghese J, Reddy SLS. Calculated plasma osmolality at hospital admission correlates well with eGFR and D-Dimer, a simple outcome predictor and guiding tool for management of severe COVID-19 patients. *Diabetes Metab Syndr*. 2021;15(5):102240. <https://doi.org/10.1016/j.dsx.2021.102240> PMID: 34403950
40. Lucijanac M, Krecak I, Cicic D, Milosevic M, Vukoja D, Kovacevic I, et al. Hypoosmolar and hyperosmolar COVID-19 patients are predisposed to dismal clinical outcomes. *Scand J Clin Lab Invest*. 2023;83(6):397–402. <https://doi.org/10.1080/00365513.2023.2241368> PMID: 37529905
41. Pearl J. Causality: models, reasoning, and inference. Cambridge University Press; 2009.
42. Hernán MA, Robins JM. Causal Inference: what if. Boca Raton: Chapman & Hall/CRC; 2020.
43. VanderWeele TJ. Mediation analysis: a practitioner's guide. *Annu Rev Public Health*. 2016;37:17–32. <https://doi.org/10.1146/annurev-publhealth-032315-021402> PMID: 26653405
44. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173–82. <https://doi.org/10.1037//0022-3514.51.6.1173> PMID: 3806354
45. Cetin M, Balci PO, Sivgin H, Cetin S, Ulgen A, Demir HD, et al. Alpha variant (B.1.1.7) of SARS-CoV-2 increases fatality-rate for patients under age of 70 years and hospitalization risk overall. *Acta Microbiol Immuno Hungarica*. 2021;68:153–61.
46. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, et al. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: insights from linking epidemiological and genetic data. *MedRxiv*. 2020. <https://doi.org/10.1101/2020.12.30.20249034>
47. Kraemer MU, Hill V, Ruis C, Dellicour S, Bajaj S, McCrone JT, et al. Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science*. 2021;373:889–95.
48. Hatirnaz Ng O, Akyoney S, Sahin I, Soykam HO, Bayram Akcapinar G, Ozdemir O, et al. Mutational landscape of SARS-CoV-2 genome in Turkey and impact of mutations on spike protein structure. *PLoS One*. 2021;16(12):e0260438. <https://doi.org/10.1371/journal.pone.0260438> PMID: 34871297
49. Giles B, Meredith P, Robson S, Smith G, Chauhan A, PACIFIC-19 and COG-UK research groups. The SARS-CoV-2 B.1.1.7 variant and increased clinical severity-the jury is out. *Lancet Infect Dis*. 2021;21(9):1213–4. [https://doi.org/10.1016/S1473-3099\(21\)00356-X](https://doi.org/10.1016/S1473-3099(21)00356-X) PMID: 34197767
50. Cetin M, Balci PO, Sivgin H, Cetin S, Ulgen A, Demir HD, et al. Alpha variant (B.1.1.7) of SARS-CoV-2 increases fatality-rate for patients under age of 70 years and hospitalization risk overall. *Acta Microbiol Immuno Hungarica*. 2021;68:153–61.
51. Van Goethem N, Vandromme M, Van Oyen H, Haahrhuis F, Brondeel R, Cateau L, et al. Severity of infection with the SARS-CoV-2 B.1.1.7 lineage among hospitalized COVID-19 patients in Belgium. *PLoS One*. 2022;17(6):e0269138. <https://doi.org/10.1371/journal.pone.0269138> PMID: 35657787
52. Rijnhart JJM, Valente MJ, Smyth HL, MacKinnon DP. Statistical mediation analysis for models with a binary mediator and a binary outcome: the differences between causal and traditional mediation analysis. *Prev Sci*. 2021;24(3):408–18. <https://doi.org/10.1007/s11212-021-01308-6>

53. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *Canadian J Emerg Med*. 2006;8:19–20.
54. Obuchowski NA, Bullen JA. Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine. *Phys Med Biol*. 2018;63(7):07TR01. <https://doi.org/10.1088/1361-6560/aab4b1> PMID: 29512515
55. Li W. Using interpretable random forest to determine observational-data-based cutoff points in medical measurements. Cold Spring Harbor Laboratory; 2025. <https://doi.org/10.1101/2025.07.10.25331301>
56. Kim S. ppcor: an R package for a fast calculation to semi-partial correlation coefficients. *Commun Stat Appl Methods*. 2015;22(6):665–74. <https://doi.org/10.5351/CSAM.2015.22.6.665> PMID: 26688802
57. Li W, Wang M, Irigoyen P, Gregersen PK. Inferring causal relationships among intermediate phenotypes and biomarkers: a case study of rheumatoid arthritis. *Bioinformatics*. 2006;22(12):1503–7. <https://doi.org/10.1093/bioinformatics/btl100>
58. Freudenberg J, Wang M, Yang Y, Li W. Partial correlation analysis indicates causal relationships between GC-content, exon density and recombination rate in the human genome. *BMC Bioinformatics*. 2009;10(S1). <https://doi.org/10.1186/1471-2105-10-s1-s66>
59. Shi B, Choirat C, Coull BA, VanderWeele TJ, Valeri L. CMAverse: a suite of functions for reproducible causal mediation analyses. *Epidemiology*. 2021;32(5):e20–2. <https://doi.org/10.1097/EDE.0000000000001378> PMID: 34028370
60. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S to analyze and compare ROC curves. *BMC Bioinf*. 2011;12:77.
61. Venables WN, Ripley BD. *Modern Applied Statistics*. 4th ed. Springer.
62. Kleiber C, Zeileis A. *Applied econometrics with R*. Springer; 2008.
63. Colquhoun D. The reproducibility of research and the misinterpretation of p-values. *R Soc Open Sci*. 2017;4(12):171085. <https://doi.org/10.1098/rsos.171085>
64. Ioannidis JPA. The proposal to lower P value thresholds to 0.05. *JAMA*. 2018;319(14):1429–30. <https://doi.org/10.1001/jama.2018.1536> PMID: 29566133
65. Li W, Shih A, Freudenberg-Hua Y, Fury W, Yang Y. Beyond standard pipeline and $p < 0.05$ in pathway enrichment analyses. *Comput Biol Chem*. 2021;92:107455. <https://doi.org/10.1016/j.compbiolchem.2021.107455> PMID: 33774420
66. Peters J, Jansing D, Scholkopf B. *Elements of Causal Inference*. MIT Press; 2017.
67. Trost HA, Gaab MR. Plasma osmolality, osmoregulation and prognosis after head injury. *Acta Neurochir (Wien)*. 1992;116(1):33–7. <https://doi.org/10.1007/BF01541250> PMID: 1615766
68. Holtfreter B, Bandt C, Kuhn SO, Grunwald U, Lehmann C, Schutt C, et al. Serum osmolality and outcome in intensive care unit patients. *Acta Obstetrica et Gynecologica Scandinavica*. 2006;50:970–7.
69. Nicholson T, Bennett K, Silke B. Serum osmolality as an outcome predictor in hospital emergency medical admissions. *Eur J Intern Med*. 2012;23(2):e39–43. <https://doi.org/10.1016/j.ejim.2011.06.014> PMID: 22284254
70. Rohla M, Freynhofer MK, Tentzeris I, Farhan S, Wojta J, Huber K, et al. Plasma osmolality predicts clinical outcome in patients with acute coronary syndrome undergoing percutaneous coronary intervention. *Eur Heart J Acute Cardiovasc Care*. 2014;3(1):84–92. <https://doi.org/10.1177/2048872613516018> PMID: 24337920
71. Cheng A, Hu L, Wang Y, Huang L, Zhao L, Zhang C, et al. Diagnostic performance of initial blood urea nitrogen combined with D-dimer levels for predicting in-hospital mortality in COVID-19 patients. *Int J Antimicrob Agents*. 2020;56(3):106110. <https://doi.org/10.1016/j.ijantimicag.2020.106110> PMID: 32712332
72. Rodrigues HCN, Silva ML, Mantovani M dos S, da Silva JM, Domingues MFP, Tanni SE, et al. Higher urea-to-albumin ratio is associated with mortality risk in critically ill COVID-19 patients. *Clin Nutrition ESPEN*. 2023;56:P9–12.
73. Zhang J, Kong W, Xia P, Xu Y, Li L, Li Q, et al. Impaired fasting glucose and diabetes are related to higher risks of complications and mortality among patients with coronavirus disease 2019. *Front Endocrinol*. 2020;11:525.
74. Cai Y, Shi S, Yang F, Yi B, Chen X, Li J, et al. Fasting blood glucose level is a predictor of mortality in patients with COVID-19 independent of diabetes history. *Diabetes Res Clin Pract*. 2020;169:108437. <https://doi.org/10.1016/j.diabres.2020.108437> PMID: 32920103
75. Tzoulis P, Waung JA, Bagkeris E, Hussein Z, Biddanda A, Cousins J, et al. Dysnatremia is a predictor for morbidity and mortality in hospitalized patients with COVID-19. *J Clin Endocrinol Metab*. 2021;106(6):1637–48. <https://doi.org/10.1210/clinem/dgab107> PMID: 33624101

76. de La Flor JC, Gomez-Berrocal A, Marschall A, Valga F, Linares T, Albarracin C, et al. The impact of the correction of hyponatremia during hospital admission on the prognosis of SARS-CoV-2 infection. *Med Clin (Engl Ed)*. 2022;159(1):12–8. <https://doi.org/10.1016/j.medcle.2021.07.021> PMID: 35784827
77. Fang Y, Shen X, Dou A, Xie H, Xie K. Association between osmolality trajectories and mortality in patients with sepsis: a group-based trajectory model in large ICU open access databases. *Front Med (Lausanne)*. 2025;12:1538322. <https://doi.org/10.3389/fmed.2025.1538322> PMID: 40357298
78. Wu W, Wang C, Zhang Y, Xie Y, Li X. Analysis of the correlation between the group-based trajectory modeling of serum osmolality and prognosis in patients with sepsis-associated encephalopathy at 72 h after admission. *BMC Inf Dis*. 2025;25:106.
79. McLean KA, O'Neill PA, Davies I, Morris J. Influence of age on plasma osmolality: a community study. *Age Ageing*. 1992;21(1):56–60. <https://doi.org/10.1093/ageing/21.1.56>
80. Lin Z, Long F, Yang Y, Chen X, Xu L, Yang M. Serum ferritin as an independent risk factor for severity in COVID-19 patients. *J Infect*. 2020;81(4):647–79. <https://doi.org/10.1016/j.jinf.2020.06.053> PMID: 32592705
81. Cheng L, Li H, Li L, Liu C, Yan S, Chen H, et al. Ferritin in the coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis. *J Clin Lab Anal*. 2020;34(10):e23618. <https://doi.org/10.1002/jcla.23618> PMID: 33078400