Article

Machine Learning Consensus Clustering Approach for Hospitalized Patients with Dysmagnesemia

Charat Thongprayoon [1], Janina Paula T. Sy-Go [1], Voravech Nissaisorakarn [2], Carissa Y. Dumancas [1],
Mira T. Keddis [3], Andrea G. Kattah [1], Pattharawin Pattharanitima [4], Saraschandra Vallabhajosyula [5],
Michael A. Mao [6], Fawad Qureshi [1], Vesna D. Garovic [1], John J. Dillon [1], Stephen B. Erickson [1]
and Wisit Cheungpasitporn [1,*]

[1] Division of Nephrology and Hypertension, Department of Medicine, Mayo Clinic,
    Rochester, MN 55905, USA; charat.thongprayoon@gmail.com (C.T.); sy-go.janina@mayo.edu (J.P.T.S.-G.);
    dumancas.carissa@mayo.edu (C.Y.D.); kattah.andrea@mayo.edu (A.G.K.); Qureshi.Fawad@mayo.edu (F.Q.);
    garovic.Vesna@mayo.edu (V.D.G.); dillon.John@mayo.edu (J.J.D.); erickson.stephen@mayo.edu (S.B.E.)
[2] Division of Nephrology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical
    School, Boston, MA 01702, USA; voravech.niss@gmail.com
[3] Division of Nephrology and Hypertension, Department of Medicine, Mayo Clinic, Phoenix, AZ 85054, USA;
    keddis.Mira@mayo.edu
[4] Department of Internal Medicine, Faculty of Medicine, Thammasat University,
    Pathum Thani 12121, Thailand; pattharawin@hotmail.com
[5] Section of Cardiovascular Medicine, Department of Medicine, Wake Forest University School of Medicine,
    Winston-Salem, NC 27101, USA; svallabh@wakehealth.edu
[6] Division of Nephrology and Hypertension, Department of Medicine, Mayo Clinic,
    Jacksonville, FL 85054, USA; mao.michael@mayo.edu
*   Correspondence: wcheungpasitporn@gmail.com

Abstract: Background: The objectives of this study were to classify patients with serum magnesium
derangement on hospital admission into clusters using unsupervised machine learning approach and
to evaluate the mortality risks among these distinct clusters. Methods: Consensus cluster analysis was
performed based on demographic information, principal diagnoses, comorbidities, and laboratory
data in hypomagnesemia (serum magnesium $\leq$ 1.6 mg/dL) and hypermagnesemia cohorts (serum
magnesium $\geq$ 2.4 mg/dL). Each cluster's key features were determined using the standardized
mean difference. The associations of the clusters with hospital mortality and one-year mortality were
assessed. Results: In hypomagnesemia cohort (*n* = 13,320), consensus cluster analysis identified
three clusters. Cluster 1 patients had the highest comorbidity burden and lowest serum magnesium.
Cluster 2 patients had the youngest age, lowest comorbidity burden, and highest kidney function.
Cluster 3 patients had the oldest age and lowest kidney function. Cluster 1 and cluster 3 were
associated with higher hospital and one-year mortality compared to cluster 2. In hypermagnesemia
cohort (*n* = 4671), the analysis identified two clusters. Compared to cluster 1, the key features of
cluster 2 included older age, higher comorbidity burden, more hospital admissions primarily due
to kidney disease, more acute kidney injury, and lower kidney function. Compared to cluster 1,
cluster 2 was associated with higher hospital mortality and one-year mortality. Conclusion: Our
cluster analysis identified clinically distinct phenotypes with differing mortality risks in hospitalized
patients with dysmagnesemia. Future studies are required to assess the application of this ML
consensus clustering approach to care for hospitalized patients with dysmagnesemia.

Keywords: artificial intelligence; clustering; consensus clustering; dysmagnesemia; electrolytes;
hypomagnesemia; hypermagnesemia; individualized medicine; machine learning; magnesium;
mortality; nephrology; personalized medicine; precision medicine

1. Introduction

Magnesium is one of the most common cations in the body with great physiologic importance, including signal transduction, glycolysis, oxidative phosphorylation, protein synthesis and degradation, and more than 300 intracellular reactions [1–7]. Previous studies have reported hypomagnesemia and hypermagnesemia as common as 11–20% and 9–12%, respectively [8,9]. Dysregulated magnesium level or dysmagnesemia, whether hypomagnesemia [8,10–12] or hypermagnesemia [8,11–14], has been associated with adverse outcomes, including cardiac arrhythmias, respiratory failure, and increased mortality in various patient populations [8,11,13,15–19]. Magnesium homeostasis is regulated by intestinal absorption and renal excretion of magnesium [20], and identifying the phenotype of patients with dysmagnesemia on hospital admission can often be challenging because of lack of detailed clinical information prior to admission and limited data on urinary magnesium.

With the advancement of electronic medical record (EMR) and artificial intelligence, machine learning (ML) approaches have been developed as part of precision medicine to assist in clinical decision-making, including disease detection, medical imaging, and explainable risk prediction [21–29]. In recent years, unsupervised ML algorithms have been utilized to reveal the patterns of diseases such as diabetes and cardiovascular diseases [30–33]. Consensus clustering is an unsupervised ML technique used to identify patterns of data, and provides a visualization tool to inspect cluster numbers, membership, and boundaries [34]. It can be utilized to search for similarities and heterogeneities among data and isolate them into clinically meaningful clusters [22,35]. Recent investigations have demonstrated that ML clustering methods can distinguish meaningful disease subtypes associated with different clinical outcomes [36,37]. Given the heterogeneity of patients with dysmagnesemia on hospital admission [8], the application of ML consensus clustering may help identify distinct phenotypic and clinicopathological clusters of dysmagnesemia that are associated with different clinical outcomes.

In this study, we aimed to identify clinically meaningful clusters of hospitalized patients with dysmagnesemia on hospital admission using an unsupervised ML approach and to assess mortality risks among these distinct clusters.

2. Methods

2.1. Patient Population

Adult patients (age $\geq$ 18 years) admitted to Mayo Clinic in Rochester, Minnesota, USA from January 2009 to 31 December 2013 were screened. Patients with serum magnesium outside the normal reference range (1.7–2.3 mg/dL) on hospital admission were included. Patients who did not have serum magnesium measurement within 24 h of hospital admission or had normal serum magnesium on hospital admission were excluded. Patients were divided into 2 cohorts: (1) hypomagnesemia cohort (serum magnesium $\leq$ 1.6 mg/dL) and (2) hypermagnesemia cohort (serum magnesium $\geq$ 2.4 mg/dL). The Mayo Clinic Institutional Review Board approved this study (IRB number 21-003088 and date of approval; 30 March 2021). All included patients provided research authorization.

2.2. Data Collection

As previously described, we collected pertinent demographic information, principal diagnoses, comorbidities, and laboratory data from our hospital's EMR [8,14]. Only available data within 24 h of hospital admission were incorporated into cluster analysis. If there were multiple laboratory values, the first one within the 24-h time frame was used. We excluded variables with more than 10% missing data. If the variable had missing data less than 10%, the missing data were imputed using Random Forest multiple imputation technique before inputting the data into cluster analysis [38].

2.3. Cluster Analysis

Unsupervised ML consensus clustering analysis was applied to identify clinical clusters of hypomagnesemia and hypermagnesemia cohorts [39]. We utilized a pre-specified subsampling parameter of 80% with 100 iterations. The number of possible clusters (k) was selected to be between 2 to 10 in order to avoid excessive numbers of clusters that would not be clinically useful. The ideal number of clusters was ascertained by evaluating the cumulative distribution function (CDF), consensus matrix (CM) heat map, cluster-consensus plots in the within-cluster consensus scores, and the proportion of ambiguously clustered pairs (PAC) [34,40]. The within-cluster consensus score (range 0–1) is defined as the average consensus value for all pairs of individuals belonging to the same cluster [34]. A value closer to one indicates better cluster stability [34]. PAC (range 0–1) is calculated as the proportion of all sample pairs with consensus values falling within the predetermined boundaries [40]. A value closer to 0 signifies higher cluster stability [40]. The details regarding the consensus cluster algorithms can be found in the online supplementary.

2.4. Statistical Analysis

After cluster identification, subsequent analyses were performed to characterize differences among the clusters. Clinical characteristics between the clusters were compared using Student's *t*-test for continuous variables and Chi-squared test for categorical variables. The key features of each cluster were determined using the standardized mean difference in clinical characteristics between each cluster and the overall cohort, and clinical characteristics with absolute standardized mean difference of >0.3 were included. Hospital mortality and one-year mortality were compared among the clusters. Logistic regression was used to assess the association of the cluster with hospital mortality, and odds ratio (OR) with 95% confidence interval (95% CI) was reported. In contrast, Cox proportional hazard regression was used to assess the association of the cluster with one-year mortality, and hazard ratio (HR) with 95% CI was reported. We did not adjust for differences in clinical variables between groups because these variables were utilized through unsupervised machine learning to identify the clusters. We used the ConsensusClusterPlus package (version 1.46.0) for consensus clustering analysis, and the "missForest" package for missing data imputation [41]. We used R, version 4.0.3 (RStudio, Inc., Boston, MA, USA) for all analyses.

3. Results

3.1. Hypomagnesemia Cohort

There were 65,974 hospitalized patients with available admission serum magnesium measurement. A total of 13,320 (20%) patients presented with hypomagnesemia on hospital admission. The mean age was 61 ± 17 years. 47% were male. The mean estimated glomerular filtration rate (eGFR) was 76 ± 31. The mean admission serum magnesium was 1.5 ± 0.2 mg/dL.

The CDF plot displays the consensus distributions for each hypomagnesemia cluster (Figure 1A, Supplementary Figure S1). The delta area plot, in turn demonstrates the relative change in the area under the CDF curve (Figure 1B, Supplementary Figure S2). The largest changes in area occurred between k = 2 and k = 4. Beyond this range, the relative increase in area became significantly smaller. The CM heatmap (Figure 2A, Supplementary Figures S3–S11) reveals that the ML algorithm identified k = 2 and 3 with clear boundaries (Figure 2A), indicating good cluster stability over repeated iterations. K = 2 and 3 also had high stability given their high mean cluster consensus score (Figure 3A). K = 3 exhibited favorably low PACs (Supplementary Figure S12); Thus, the consensus clustering analysis from available hospital admission baseline characteristics identified three clusters that best represented the data pattern of our patients admitted with hypomagnesemia.
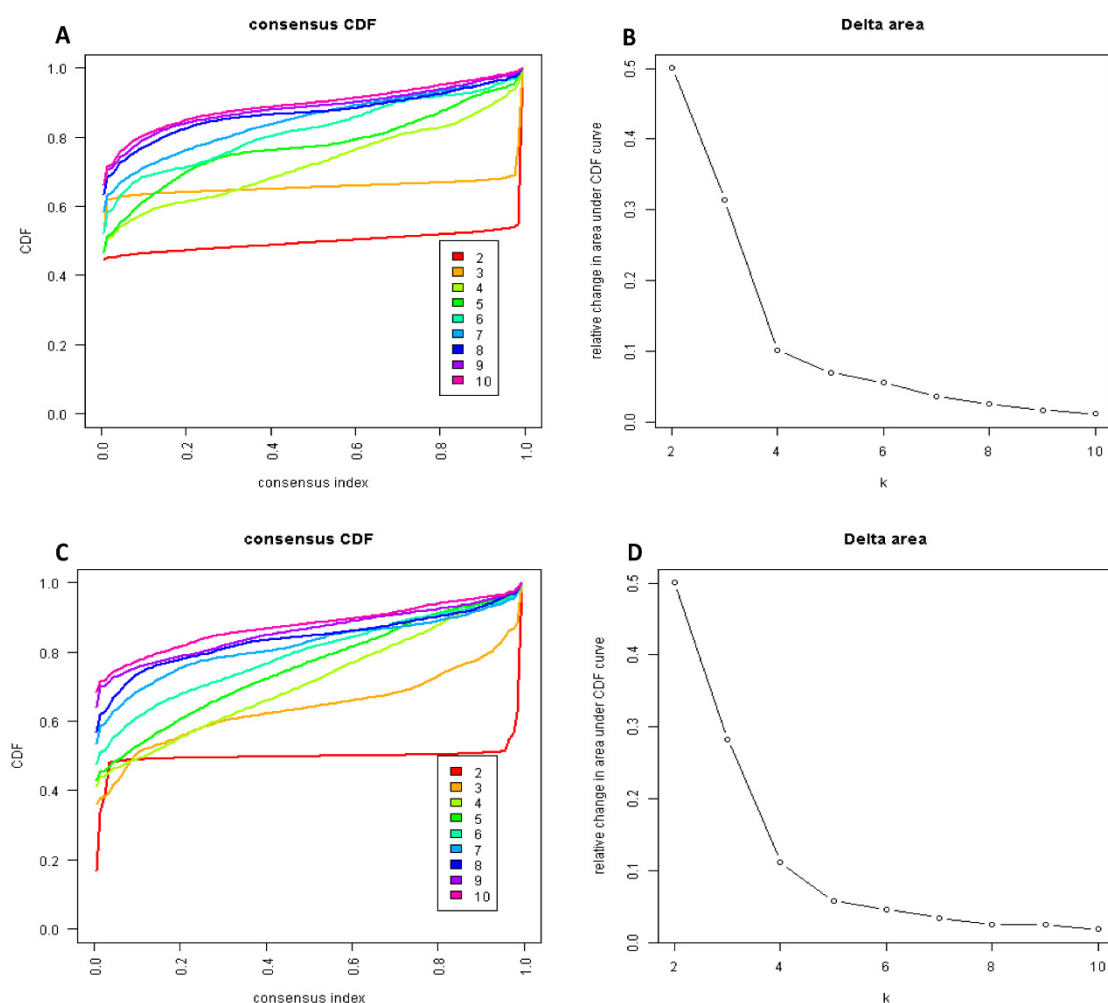
Figure 1. (**A**) Cumulative distribution function (CDF) plot displaying consensus distributions for each cluster (k) for patients with hypomagnesemia; (**B**) Delta area plot reflecting the relative changes in the area under the CDF curve for hypomagnesemia. (**C**) CDF plot displaying consensus distributions for each cluster (k) for patients with hypermagnesemia; (**D**) Delta area plot reflecting the relative changes in the area under the CDF curve for hypermagnesemia.

Cluster 1 had 3446 (26%) patients. Cluster 2 had 4351 (33%) patients. Cluster 3 had 5523 (41%) patients. As shown in Table 1, baseline characteristics significantly differed among the three clusters in the hypomagnesemia cohort.

Based on standardized mean difference shown in Figure 4, cluster 1 was mainly characterized by higher comorbidity burden and lower serum magnesium, albumin, and calcium. On the other hand, cluster 2 was mainly characterized by younger age, lower comorbidity burden, especially less history of hypertension, diabetes mellitus, coronary artery disease, and leukemia/lymphoma, less use of angiotensin converting enzyme inhibitors (ACEI)/angiotensin receptor blockers (ARB), and diuretics, less acute kidney injury (AKI), higher eGFR, and lower serum albumin and calcium. Lastly, cluster 3 was mainly characterized by older age, more history of hypertension, more use of ACEI/ARB, lower eGFR, higher serum potassium, magnesium, albumin, and calcium.

The hospital-mortality was 3.4% in cluster 1, 1.4% in cluster 2, and 2.3% in cluster 3 (<0.001) (Figure 5A). Compared to cluster 2, cluster 1 and cluster 3 had higher odds of hospital mortality with OR of 2.45 (95% CI 1.79–3.35), and 1.63 (95% CI 1.20–2.22) respectively. The one-year mortality was 20.1% in cluster 1, 10.5% in cluster 2, and 16.8% in cluster 3 (Figure 5B).

Figure 2. (**A**) Consensus matrix heat map depicting consensus values on a white to blue color scale of each cluster (k) for patients with hypomagnesemia; (**B**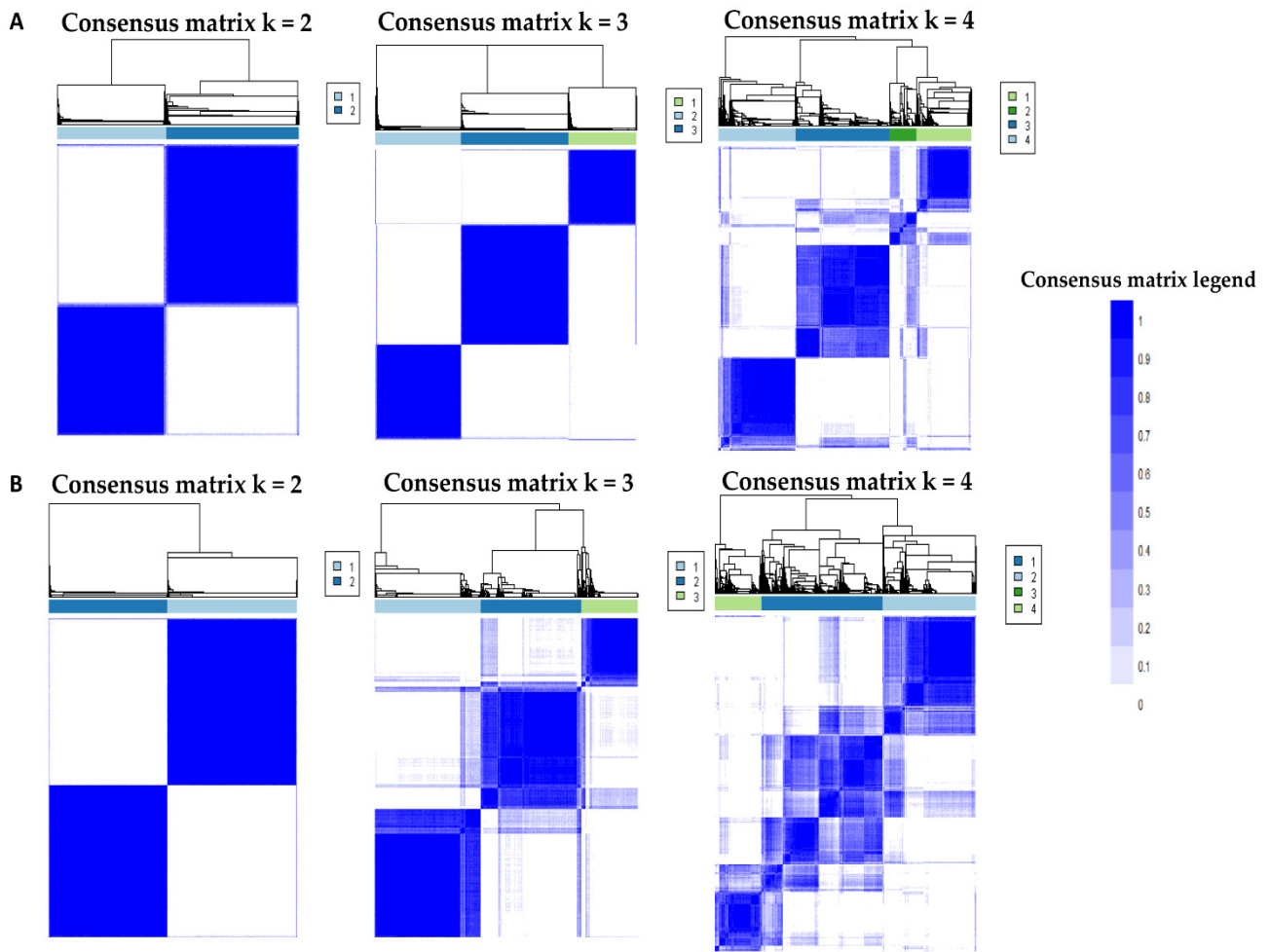) Consensus matrix heat map depicting consensus values on a white to blue color scale of each cluster (k) for patients with hypermagnesemia.
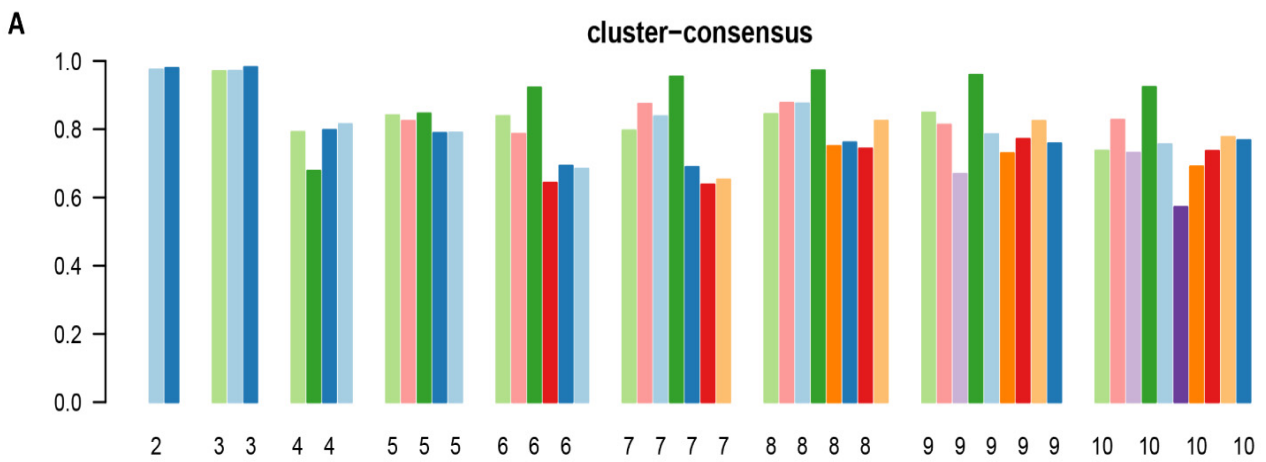

Figure 3. *Cont.*
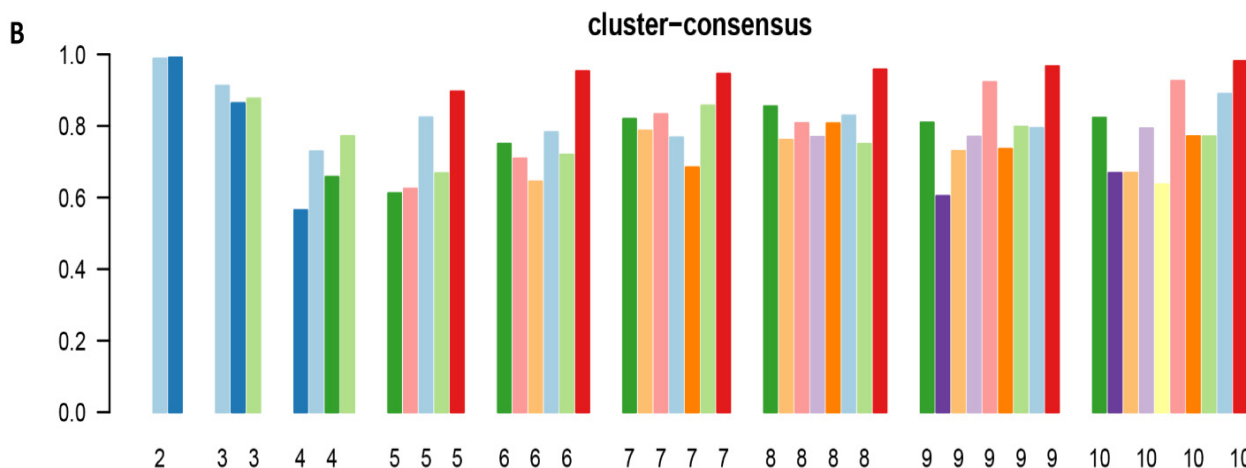
Figure 3. (**A**) The bar plot represents the mean consensus score for different numbers of clusters (K ranges from two to ten) for patients with hypomagnesemia; (**B**) The bar plot represents the mean consensus score for different numbers of clusters (K ranges from two to ten) for patients with hypermagnesemia.

Table 1. Clinical characteristics.

Patient Characteristics	Hypomagnesemia Cohort (*n* = 13,320)				Hypermagnesemia Cohort (*n* = 4671)		
	Cluster 1 (*n* = 3446)	Cluster 2 (*n* = 4351)	Cluster 3 (*n* = 5523)	*p*-Value	Cluster 1 (*n* = 2445)	Cluster 2 (*n* = 2226)	*p*-Value
Age (years)	66.0 ± 13.5	48.2 ± 15.9	67.9 ± 13.4	<0.001	59.5 ± 17.4	71.9 ± 14.6	<0.001
Male sex	1458 (42)	1848 (42)	2987 (54)	<0.001	1444 (59)	1399 (63)	0.008
Race -White -Black -Others	3208 (93) 40 (1) 198 (6)	3810 (88) 109 (3) 432 (10)	5196 (94) 64 (1) 263 (5)	<0.001	2198 (90) 53 (2) 194 (8)	2016 (91) 43 (2) 167 (8)	0.72
Principal diagnosis -Cardiovascular	408 (12)	443 (10)	1084 (20)		863 (35)	718 (32)	
-Endocrine/metabolic	133 (4)	136 (3)	218 (4)		83 (3)	126 (6)	
-Gastrointestinal	433 (13)	829 (19)	466 (8)	<0.001	247 (10)	231 (10)	<0.001
-Genitourinary	184 (5)	94 (2)	332 (6)		16 (0.7)	335 (15)	
-Hematology/oncology	1163 (34)	1082 (25)	1632 (30)		283 (12)	157 (7)	
-Infectious disease	334 (10)	226 (5)	159 (3)		67 (3)	141 (6)	
-Respiratory	131 (4)	85 (2)	254 (5)		152 (6)	139 (6)	
-Injury/poisoning	363 (11)	828 (19)	649 (12)		293 (12)	187 (8)	
-Other	297 (9)	628 (14)	729 (13)		441 (18)	192 (9)	
Charlson Comorbidity Score	3.4 ± 3.0	1.3 ± 1.9	2.7 ± 2.6	<0.001	1.4 ± 1.9	3.2 ± 2.8	<0.001
Comorbidities -Hypertension	2400 (70)	944 (22)	4101 (74)	<0.001	1088 (44)	1621 (73)	<0.001
-Diabetes mellitus	1169 (34)	455 (10)	1835 (33)	<0.001	282 (12)	862 (39)	<0.001
-Coronary artery disease	788 (23)	190 (4)	1493 (27)	<0.001	451 (18)	933 (42)	<0.001
-Congestive heart failure	215 (6)	67 (2)	398 (7)	<0.001	203 (8)	614 (28)	<0.001
-Peripheral vascular disease	170 (5)	50 (1)	256 (5)	<0.001	49 (2)	204 (9)	<0.001

Table 1. *Cont.*

| Patient Characteristics	Hypomagnesemia Cohort (*n* = 13,320)				Hypermagnesemia Cohort (*n* = 4671)		
	Cluster 1 (*n* = 3446)	Cluster 2 (*n* = 4351)	Cluster 3 (*n* = 5523)	*p*-Value	Cluster 1 (*n* = 2445)	Cluster 2 (*n* = 2226)	*p*-Value
-Stroke	240 (7)	96 (2)	547 (10)	<0.001	164 (7)	312 (14)	<0.001
-End-stage kidney disease	278 (8)	38 (1)	458 (8)	<0.001	19 (0.8)	270 (12)	<0.001
-Dementia	36 (1)	14 (0.3)	80 (1)	<0.001	26 (1)	79 (4)	<0.001
-COPD	410 (12)	136 (3)	676 (12)	<0.001	232 (9)	393 (18)	<0.001
-Cirrhosis	178 (5)	132 (3)	187 (3)	<0.001	58 (2)	132 (6)	<0.001
-Cancer	1699 (49)	1087 (25)	2071 (37)	<0.001	414 (17)	597 (27)	<0.001
-Leukemia/lymphoma	142 (4)	123 (3)	293 (5)	<0.001	117 (5)	90 (4)	0.22
Alcohol use	234 (7)	393 (9)	213 (4)	<0.001	162 (7)	107 (5)	0.008
Laboratory test							
-eGFR (mL/min/1.73 m^2)	67 ± 29	101 ± 23	63 ± 26	<0.001	81 ± 25	28 ± 16	<0.001
-Potassium (mEq/L)	3.9 ± 0.6	3.9 ± 0.5	4.3 ± 0.6	<0.001	4.2 ± 0.6	4.8 ± 0.9	<0.001
-Magnesium (mg/dL)	1.3 ± 0.2	1.4 ± 0.1	1.5 ± 0.1	<0.001	2.6 ± 0.4	2.7 ± 0.4	<0.001
-Albumin (g/dL)	3.2 ± 0.5	3.1 ± 0.5	3.7 ± 0.4	<0.001	3.8 ± 0.4	3.4 ± 0.5	<0.001
-Total calcium (mg/dL)	8.3 ± 0.8	8.2 ± 0.6	9.2 ± 0.7	<0.001	9.3 ± 0.6	9.0 ± 0.8	<0.001
-Phosphorus (mg/dL)	3.7 ± 1.0	3.5 ± 0.8	3.8 ± 0.9	<0.001	3.7 ± 0.7	5.0 ± 1.8	<0.001
Medication							
-ACEI/ARB	1629 (47)	557 (13)	3047 (55)	<0.001	789 (32)	1172 (53)	<0.001
-Diuretics	1524 (44)	699 (16)	2602 (47)	<0.001	1007 (41)	1150 (70)	<0.001
Acute kidney injury	830 (24)	179 (4)	1431 (26)	<0.001	156 (6)	1947 (87)	<0.001
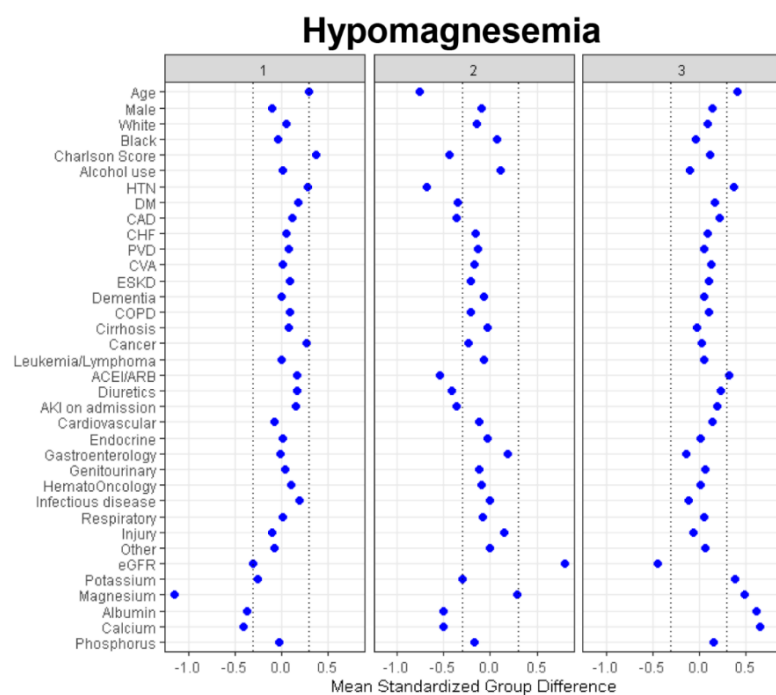
Figure 4. *Cont.*
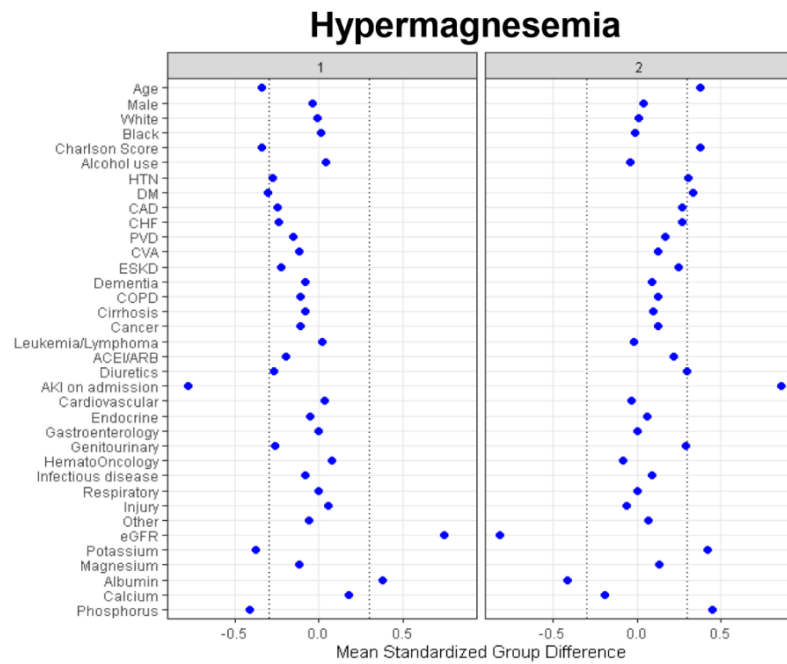
Figure 4. The standardized differences across three clusters for each of baseline parameters for patients with hypomagnesemia and hypermagnesemia. The *x*-axis represents the standardized differences value, and the y axis represents baseline variables. The dashed vertical lines signify the standardized differences cutoffs of <−0.3 or >0.3. Abbreviations: AG, anion gap; AKI, acute kidney injury; BMI, body mass index; CHF, congestive heart failure; Cl, chloride; COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; DM, diabetes mellitus; ESKD, end stage kidney disease; GFR, glomerular filtration rate; GI, gastrointestinal; Hb, hemoglobin; HCO3, bicarbonate; K, potassium; ID, infectious disease; MI, myocardial infarction; Na, sodium; PVD, peripheral vascular disease; RS, respiratory system; SID, strong ion difference.
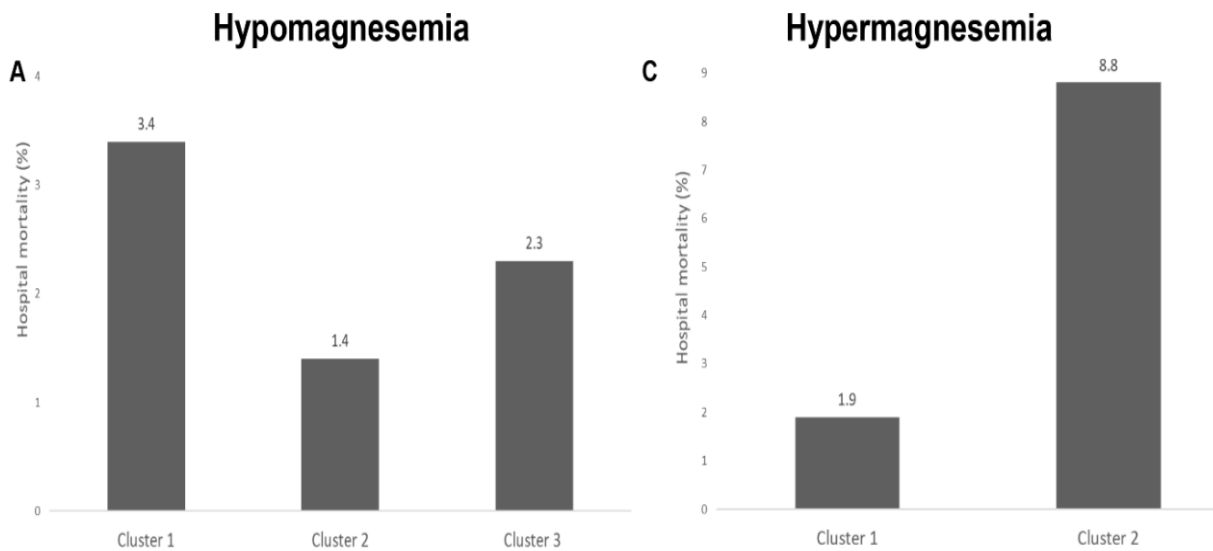

Figure 5. *Cont*.
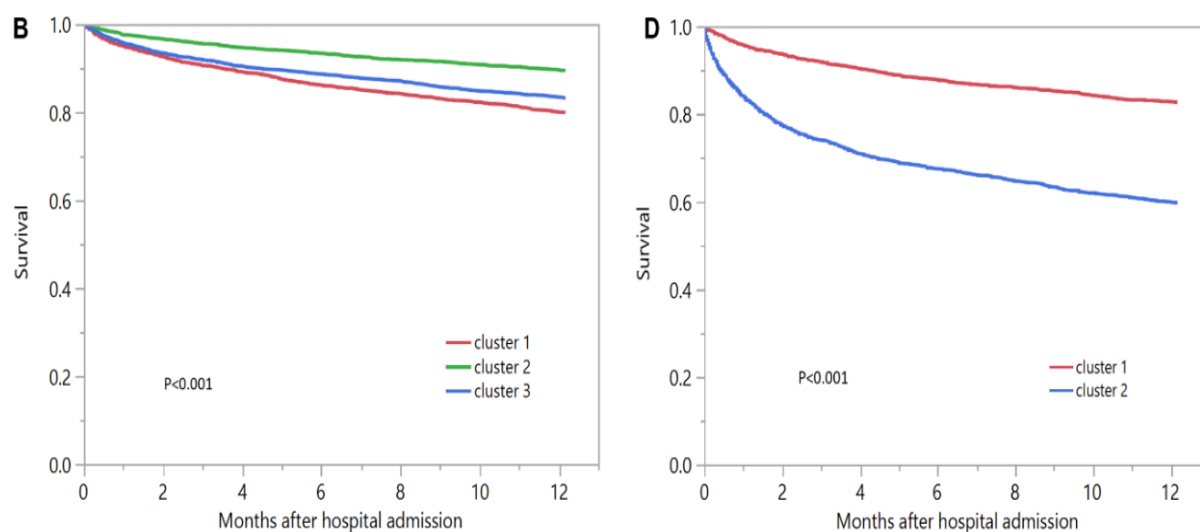
Figure 5. (**A**) Hospital mortality and (**B**) One-year mortality among different clusters of admission hypomagnesemia; (**C**) Hospital mortality and (**D**) One-year mortality among different clusters of admission hypermagnesemia.

Compared to cluster 2, cluster 1 and cluster 3 also had higher risk of one-year mortality with HR of 2.04 (95% CI 1.79–2.32) and 1.68 (95% CI 1.49–1.90) respectively (Table 2).

Table 2. Mortality outcomes according to clusters.

	Hospital Mortality	OR (95% CI)	1-Year Mortality	HR (95% CI)
(a) Hypomagnesemia cohort				
Cluster 1	3.4%	2.45 (1.79–3.35)	20.1%	2.04 (1.79–2.32)
Cluster 2	1.4%	1 (ref)	10.5%	1 (ref)
Cluster 3	2.3%	1.63 (1.20–2.22)	16.8%	1.68 (1.49–1.90)
(b) Hypermagnesemia cohort				
Cluster 1	1.9%	1 (ref)	17.3%	1 (ref)
Cluster 2	8.8%	5.04 (3.63–6.98)	40.3%	2.85 (2.51–3.24)

3.2. Hypermagnesemia Cohort

A total of 4671 (7%) patients presented with hypermagnesemia on hospital admission. The mean age was 65 ± 17 years. 61% were male. The mean eGFR was 56 ± 34. The mean admission serum magnesium was 2.6 ± 0.4 mg/dL.

The CDF plot displays the consensus distributions for each hypermagnesemia cluster (Figure 1C, Supplementary Figure S13). The delta area plot, in turn demonstrates the relative change in the area under the CDF curve (Figure 1D, Supplementary Figure S14). The largest changes in area occurred between k = 2 and k = 4. Beyond this range, the relative increase in area became significantly smaller. The CM heatmap (Figure 2B, Supplementary Figures S15–S23) reveals that the ML algorithm identified k = 2 with clear boundaries (Figure 2B), indicating good cluster stability over repeated iterations. K = 2 also had high stability given its high mean cluster consensus score (Figure 3B). Favorably low PACs were demonstrated for k = 2 (Supplementary Figure S24). Thus, the consensus clustering analysis from available hospital admission baseline characteristics identified two clusters that optimally represented the data pattern of our patients admitted with hypermagnesemia.

Cluster 1 had 2445 (52%) patients while cluster 2 had 2226 (48%) patients. As shown in Table 1, clinical characteristics were significantly different between the two identified clusters in the hypermagnesemia cohort. Based on standardized mean difference shown in Figure 4, the key features of patients in cluster 2, compared to those of patients in cluster 1,

included older age, higher comorbidity burden, especially more history of hypertension and diabetes mellitus, more AKI, more hospital admissions primarily due to kidney disease, and lower eGFR and serum albumin but higher potassium and phosphorus.

Cluster 1 had hospital mortality of 1.9%, while cluster 2 had hospital mortality of 8.8% ($p < 0.001$) (Figure 5C). Compared to cluster 1, cluster 2 had higher odds of hospital mortality with OR of 5.04 (95% CI 3.63–6.98). Cluster 1 had one-year mortality of 17.3%, whereas cluster 2 had one-year mortality of 40.3% ($p < 0.001$) (Figure 5D). Compared to cluster 1, cluster 2 also had higher risk of one-year mortality with HR of 2.85 (95% CI 2.51–3.24) (Table 2).

4. Discussion

ML consensus clustering algorithms offer the ability to efficiently analyze and identify clusters of patients with different characteristics in a large amount of data [22,35,42,43]. In this study, the unsupervised ML consensus clustering approach was utilized to distinguish patients with dysmagnesemia into distinct clusters. Among patients with hypomagnesemia on hospital admission, age, comorbidity burden, kidney function (with baseline eGFR and AKI on admission as surrogate markers), and degree of hypomagnesemia were important features to differentiate the phenotypes. Similarly, among patients with hypermagnesemia on hospital admission, age, comorbidity burden, kidney function, and principal genitourinary diagnosis were important features to differentiate the phenotypes.

Applying the unsupervised consensus clustering approach to the patient characteristics at the time of hospital admission, we identified three clinically distinct clusters of patients with concomitant hypomagnesemia. The three clusters demonstrated different characteristics and were associated with different hospital and one-year mortality risks. Even though the majority of the patients across all three clusters presented with similar conditions (mainly hematology/oncology-, cardiovascular-, and gastrointestinal-related conditions), the three clusters demonstrated different clinical outcomes.

Cluster 2, the reference cluster of hypomagnesemia, consisted of patients with younger age with the lowest comorbidity burden compared to other clusters. They also had higher eGFR, lower use of ACEI/ARB and diuretics, and lower incidence of AKI. Interestingly, they had the highest prevalence of alcohol use. They also had the lowest serum albumin, calcium, and phosphate compared to other clusters. Calcium and magnesium are electrolytes that normally bind to albumin. Alteration of circulating albumin levels alters the measured levels of these electrolytes. Measured or total calcium and magnesium alike are lower in the setting of concomitant low serum albumin [44–46]. In addition, it is possible that alcoholism and malnutrition played important roles in the development of hypomagnesemia in this patient population because of poor oral intake. Furthermore, these patients also had the highest principal diagnosis of gastrointestinal conditions on admission, which could have potentially resulted in gastrointestinal magnesium loss or redistribution of magnesium triggered by acute pancreatitis [47,48]. Given the fact that patients in cluster 2 were younger and had the lowest comorbidity burden, they had the lowest in-hospital and one-year mortality risks among the three clusters.

Compared to the patients in cluster 2 of hypomagnesemia, those in clusters 1 and 3 were older and had higher comorbidity burden, reduced kidney function (lower baseline eGFR and higher incidence of AKI), and higher use of ACEI/ARB and diuretics. Patients in cluster 1 had the highest comorbidity burden and the highest prevalence of cancer among the three clusters. They also presented more frequently with a principal diagnosis of infectious disease on hospital admission. Furthermore, they had the lowest serum magnesium. A number of cancer-specific therapies can cause hypomagnesemia via renal magnesium wasting, including platinum-based chemotherapy, anti-epidermal growth factor receptor (EGFR) monoclonal antibodies, inhibitors of human epidermal growth factor receptor 2 (HER2), and calcineurin inhibitors [49]. In addition, cancer patients frequently use medications that cause or exacerbate hypomagnesemia, such as proton pump inhibitors (PPIs), diuretics, and chemotherapy [49]. These patients in cluster 1 had

the highest in-hospital and one-year mortality risks among the three clusters, which could potentially be due to poor outcomes among cancer patients with hypomagnesemia [8,50]. Previous studies also reported that hypomagnesemia is associated with increased mortality risk among patients with infection [10,51,52], which is one of the main characteristics of the patients in this cluster.

Patients in cluster 3 of hypomagnesemia were the oldest among the three clusters. They had the highest prevalence of hypertension, coronary artery disease, congestive heart failure, and principal diagnosis of cardiovascular disease on hospital admission. They also had the highest use of diuretics, which could lead to urinary magnesium wasting. Previous studies have shown that hypomagnesemia in patients with cardiovascular diseases carries a high mortality rate [53,54]. While these patients in cluster 3 had increased in-hospital mortality and one-year mortality compared to those in cluster 2, they had lower mortality when compared to those in cluster 1 (phenotype of cancer patients with hypomagnesemia) despite having the oldest age group among the three clusters.

Within the cohort of hypermagnesemia on hospital admission, we identified two clinically distinct clusters by ML consensus clustering approach. Compared to cluster 1, the key features of cluster 2 included older age, higher comorbidity burden, more hospital admissions primarily due to kidney disease, more AKI, and lower eGFR. Because of the reduction in kidney function, these patients in cluster 2 likely had reduced ability to renally excrete magnesium, resulting in hypermagnesemia. Reported clinical cases include the administration of antacids or magnesium supplements in older patients or those with reduced kidney function [55]. Conversely, patients in cluster 1 were younger but had higher alcohol use and were more likely to be admitted with principal diagnoses of injury/poisoning and hematology/oncology. As such, it is possible that hypermagnesemia among patients in cluster 1 could have resulted from excessive tissue injury or breakdown (such as tumor lysis syndrome and burn injury) [56,57]. Compared to those in cluster 1, patients in cluster 2 had higher hospital mortality and one-year mortality.

The strengths of this study include a large sample size and unbiased data manipulation, and easy reproducibility by unsupervised ML consensus clustering. There are also, however, limitations that should be noted. First, the data were abstracted from a single-center, and our patient population was predominantly Caucasian, which might limit the extrapolation of our findings to other populations. Second, consensus clustering was performed on hospital admission and did not include data before or during hospitalization, which could affect hospitalization-related outcomes. Third, we did not have data on magnesium supplements or other medications that might have affected serum magnesium, such as antibiotics, proton pump inhibitors, and chemotherapy. Fourth, some relevant laboratory results were not available, including 24-h urinary magnesium, fractional urinary excretion of magnesium, and genetic testing (for conditions that might have caused hypomagnesemia in adults, such as Gitelman syndrome). These investigations are not commonly performed on hospital admission and were thus not included in our ML clustering algorithm. Therefore, future studies are required to assess whether these variables could have improved the discriminatory ability of these clusters we identified. Nevertheless, we included readily available data at the time of hospital admission and successfully identified distinct clusters of dysmagnesemia associated with different clinical outcomes.

5. Conclusions

In summary, we present an unsupervised ML consensus clustering analysis of hospitalized patients with dysmagnesemia. We discovered three distinct phenotypes of admission hypomagnesemia and two distinct phenotypes of admission hypermagnesemia with different hospital and one-year mortality risks. With the advancement of EMR and artificial intelligence, future studies are required to evaluate and validate the application of this ML consensus clustering approach to care for hospitalized patients with dysmagnesemia.

Author Contributions: Conceptualization, C.T., J.P.T.S.-G., V.N., C.Y.D., M.T.K., A.G.K., P.P., S.V., M.A.M., F.Q., V.D.G., J.J.D., S.B.E. and W.C.; Data curation, C.T. and W.C.; Formal analysis, C.T. and W.C.; Investigation, C.T., P.P. and W.C.; Methodology, C.T., P.P. and W.C.; Project administration, J.P.T.S.-G. and V.N.; Resources, W.C.; Software, W.C.; Supervision, M.T.K., A.G.K., S.V., M.A.M., F.Q., V.D.G., J.J.D., S.B.E. and W.C.; Validation, C.T., V.N., C.Y.D., P.P. and W.C.; Visualization, C.T. and W.C.; Writing—original draft, C.T. and W.C.; Writing—review & editing, C.T., J.P.T.S.-G., V.N., C.Y.D., M.T.K., A.G.K., P.P., S.V., M.A.M., F.Q., V.D.G., J.J.D., S.B.E. and W.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The Mayo Clinic Institutional Review Board approved this study (IRB number 21-003088 and date of approval; 3/30/2021). All included patients provided research authorization.

Informed Consent Statement: Patient consent was waived due to the minimal risk nature of observational chart review study.

Data Availability Statement: Data is available upon reasonable request to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ACEI	angiotensin converting enzyme inhibitors
AG	anion gap
AKI	acute kidney injury
ARB	angiotensin receptor blockers
BMI	body mass index
CDF	cumulative distribution function
CHF	congestive heart failure
CI	confidence interval
Cl	chloride
CM	consensus matrix;
COPD	chronic obstructive pulmonary disease
CVA	cerebrovascular accident
DM	diabetes mellitus
eGFR	estimated glomerular filtration rate
EMR	electronic medical record
ESKD	end stage kidney disease
GI	gastrointestinal
Hb	hemoglobin
HCO3	bicarbonate
HR	hazard ratio
ID	infectious disease
K	potassium
MI	myocardial infarction
ML	machine learning
Na	sodium
OR	odds ratio
PAC	proportion of ambiguously clustered pairs
PVD	peripheral vascular disease
RS	respiratory system
SID	strong ion difference

References

1. Glasdam, S.M.; Glasdam, S.; Peters, G.H. The Importance of Magnesium in the Human Body: A Systematic Literature Review. *Adv. Clin. Chem.* **2016**, *73*, 169–193. [CrossRef] [PubMed]
2. Cheungpasitporn, W.; Thongprayoon, C.; Kittanamongkolchai, W.; Srivali, N.; Edmonds, P.J.; Ungprasert, P.; O'Corragain, O.A.; Korpaisarn, S.; Erickson, S.B. Proton pump inhibitors linked to hypomagnesemia: A systematic review and meta-analysis of observational studies. *Ren. Fail.* **2015**, *37*, 1237–1241. [CrossRef]
3. Cheungpasitporn, W.; Thongprayoon, C.; Mao, M.A.; Srivali, N.; Ungprasert, P.; Varothai, N.; Sanguankeo, A.; Kittanamongkolchai, W.; Erickson, S.B. Hypomagnesaemia linked to depression: A systematic review and meta-analysis. *Intern. Med. J.* **2015**, *45*, 436–440. [CrossRef] [PubMed]
4. Qu, X.; Jin, F.; Hao, Y.; Li, H.; Tang, T.; Wang, H.; Yan, W.; Dai, K. Magnesium and the risk of cardiovascular events: A meta-analysis of prospective cohort studies. *PLoS ONE* **2013**, *8*, e57720. [CrossRef]
5. de Baaij, J.H.; Hoenderop, J.G.; Bindels, R.J. Magnesium in man: Implications for health and disease. *Physiol. Rev.* **2015**, *95*, 1–46. [CrossRef] [PubMed]
6. Cheungpasitporn, W.; Thongprayoon, C.; Harindhanavudhi, T.; Edmonds, P.J.; Erickson, S.B. Hypomagnesemia linked to new-onset diabetes mellitus after kidney transplantation: A systematic review and meta-analysis. *Endocr. Res.* **2016**, *41*, 142–147. [CrossRef]
7. Schiroli, D.; Marraccini, C.; Zanetti, E.; Ragazzi, M.; Gianoncelli, A.; Quartieri, E.; Gasparini, E.; Iotti, S.; Baricchi, R.; Merolle, L. Imbalance of Mg Homeostasis as a Potential Biomarker in Colon Cancer. *Diagnostics* **2021**, *11*, 727. [CrossRef]
8. Cheungpasitporn, W.; Thongprayoon, C.; Qian, Q. Dysmagnesemia in Hospitalized Patients: Prevalence and Prognostic Importance. *Mayo Clin. Proc.* **2015**, *90*, 1001–1010. [CrossRef]
9. Wong, E.T.; Rude, R.K.; Singer, F.R.; Shaw, S.T., Jr. A high prevalence of hypomagnesemia and hypermagnesemia in hospitalized patients. *Am. J. Clin. Pathol.* **1983**, *79*, 348–352. [CrossRef]
10. Thongprayoon, C.; Cheungpasitporn, W.; Erickson, S.B. Admission hypomagnesemia linked to septic shock in patients with systemic inflammatory response syndrome. *Ren. Fail.* **2015**, *37*, 1518–1521. [CrossRef]
11. Malinowska, J.; Małecka, M.; Ciepiela, O. Variations in Magnesium Concentration Are Associated with Increased Mortality: Study in an Unselected Population of Hospitalized Patients. *Nutrients* **2020**, *12*, 1836. [CrossRef] [PubMed]
12. Cheungpasitporn, W.; Thongprayoon, C.; Chewcharat, A.; Petnak, T.; Mao, M.A.; Davis, P.W.; Bathini, T.; Vallabhajosyula, S.; Qureshi, F.; Erickson, S.B. Hospital-Acquired Dysmagnesemia and In-Hospital Mortality. *Med. Sci.* **2020**, *8*, 37. [CrossRef] [PubMed]
13. Haider, D.G.; Lindner, G.; Ahmad, S.S.; Sauter, T.; Wolzt, M.; Leichtle, A.B.; Fiedler, G.M.; Exadaktylos, A.K.; Fuhrmann, V. Hypermagnesemia is a strong independent risk factor for mortality in critically ill patients: Results from a cross-sectional study. *Eur. J. Intern. Med.* **2015**, *26*, 504–507. [CrossRef] [PubMed]
14. Cheungpasitporn, W.; Thongprayoon, C.; Bathini, T.; Hansrivijit, P.; Vaitla, P.; Medaura, J.; Vallabhajosyula, S.; Chewcharat, A.; Mao, M.A.; Erickson, S.B. Impact of admission serum magnesium levels on long-term mortality in hospitalized patients. *Hosp. Pract. (1995)* **2020**, *48*, 80–85. [CrossRef]
15. Thongprayoon, C.; Hansrivijit, P.; Petnak, T.; Mao, M.A.; Bathini, T.; Duriseti, P.; Vallabhajosyula, S.; Qureshi, F.; Erickson, S.B.; Cheungpasitporn, W. Impact of serum magnesium levels at hospital discharge and one-year mortality. *Postgrad. Med.* **2021**. [CrossRef] [PubMed]
16. Cheungpasitporn, W.; Thongprayoon, C.; Erickson, S.B. Admission hypomagnesemia and hypermagnesemia increase the risk of acute kidney injury. *Ren. Fail.* **2015**, *37*, 1175–1179. [CrossRef]
17. Thongprayoon, C.; Cheungpasitporn, W.; Srivali, N.; Erickson, S.B. Admission serum magnesium levels and the risk of acute respiratory failure. *Int. J. Clin. Pract.* **2015**, *69*, 1303–1308. [CrossRef]
18. Naksuk, N.; Hu, T.; Krittanawong, C.; Thongprayoon, C.; Sharma, S.; Park, J.Y.; Rosenbaum, A.N.; Gaba, P.; Killu, A.M.; Sugrue, A.M.; et al. Association of Serum Magnesium on Mortality in Patients Admitted to the Intensive Cardiac Care Unit. *Am. J. Med.* **2017**, *130*, 229.e213–229.e225. [CrossRef] [PubMed]
19. Cheungpasitporn, W.; Thongprayoon, C.; Ungprasert, P.; Wijarnpreecha, K.; Mao, M.A.; Aeddula, N.R.; Kaewput, W.; Bathini, T.; Kroner, P.T. Hepatitis A hospitalizations among kidney transplant recipients in the United States: Nationwide inpatient sample 2005-2014. *Eur. J. Gastroenterol. Hepatol.* **2020**, *32*, 650–655. [CrossRef]
20. Jahnen-Dechent, W.; Ketteler, M. Magnesium basics. *Clin. Kidney J.* **2012**, *5*, i3–i14. [CrossRef]
21. Thongprayoon, C.; Kaewput, W.; Kovvuru, K.; Hansrivijit, P.; Kanduri, S.R.; Bathini, T.; Chewcharat, A.; Leeaphorn, N.; Gonzalez-Suarez, M.L.; Cheungpasitporn, W. Promises of Big Data and Artificial Intelligence in Nephrology and Transplantation. *J. Clin. Med.* **2020**, *9*, 1107. [CrossRef]
22. MacEachern, S.J.; Forkert, N.D. Machine learning for precision medicine. *Genome* **2021**, *64*, 416–425. [CrossRef]
23. Viton, F.; Elbattah, M.; Guérin, J.-L.; Dequen, G. Multi-channel ConvNet Approach to Predict the Risk of in-Hospital Mortality for ICU Patients. In Proceedings of the DeLTA, 1 November 2021; pp. 98–102.
24. Jiang, Z.; Bo, L.; Xu, Z.; Song, Y.; Wang, J.; Wen, P.; Wan, X.; Yang, T.; Deng, X.; Bian, J. An explainable machine learning algorithm for risk factor analysis of in-hospital mortality in sepsis survivors with ICU readmission. *Comput. Methods Progr. Biomed.* **2021**, *204*, 106040. [CrossRef] [PubMed]

25. Malyugin, B.; Sakhnov, S.; Izmailova, S.; Boiko, E.; Pozdeyeva, N.; Axenova, L.; Axenov, K.; Titov, A.; Terentyeva, A.; Zakaraiia, T.; et al. Keratoconus Diagnostic and Treatment Algorithms Based on Machine-Learning Methods. *Diagnostics* **2021**, *11*, 1933. [CrossRef] [PubMed]

26. Wang, T.; Chen, Z.; Shang, Q.; Ma, C.; Chen, X.; Xiao, E. A Promising and Challenging Approach: Radiologists' Perspective on Deep Learning and Artificial Intelligence for Fighting COVID-19. *Diagnostics* **2021**, *11*, 1924. [CrossRef] [PubMed]

27. Fernández-Gutiérrez, F.; Kennedy, J.I.; Cooksey, R.; Atkinson, M.; Choy, E.; Brophy, S.; Huo, L.; Zhou, S.M. Mining Primary Care Electronic Health Records for Automatic Disease Phenotyping: A Transparent Machine Learning Framework. *Diagnostics* **2021**, *11*, 1908. [CrossRef] [PubMed]

28. Park, D.; Jeong, E.; Kim, H.; Pyun, H.W.; Kim, H.; Choi, Y.J.; Kim, Y.; Jin, S.; Hong, D.; Lee, D.W.; et al. Machine Learning-Based Three-Month Outcome Prediction in Acute Ischemic Stroke: A Single Cerebrovascular-Specialty Hospital Study in South Korea. *Diagnostics* **2021**, *11*, 1909. [CrossRef]

29. Murdaca, G.; Caprioli, S.; Tonacci, A.; Billeci, L.; Greco, M.; Negrini, S.; Cittadini, G.; Zentilin, P.; Ventura Spagnolo, E.; Gangemi, S. A Machine Learning Application to Predict Early Lung Involvement in Scleroderma: A Feasibility Evaluation. *Diagnostics* **2021**, *11*, 1880. [CrossRef]

30. Nedyalkova, M.; Madurga, S.; Simeonov, V. Combinatorial K-Means Clustering as a Machine Learning Tool Applied to Diabetes Mellitus Type 2. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1919. [CrossRef]

31. Nedyalkova, M.; Barazorda-Ccahuana, H.L.; Sârbu, C.; Madurga, S.; Simeonov, V. Fuzzy partitioning of clinical data for DMT2 patients. *J. Environ. Sci. Health A Tox. Hazard. Subst. Environ. Eng.* **2020**, *55*, 1450–1458. [CrossRef]

32. Lachmann, M.; Rippen, E.; Schuster, T.; Xhepa, E.; von Scheidt, M.; Pellegrini, C.; Trenkwalder, T.; Rheude, T.; Stundl, A.; Thalmann, R.; et al. Subphenotyping of Patients With Aortic Stenosis by Unsupervised Agglomerative Clustering of Echocardiographic and Hemodynamic Data. *JACC Cardiovasc. Interv.* **2021**, *14*, 2127–2140. [CrossRef] [PubMed]

33. Yoon, Y.E.; Baskaran, L.; Lee, B.C.; Pandey, M.K.; Goebel, B.; Lee, S.E.; Sung, J.M.; Andreini, D.; Al-Mallah, M.H.; Budoff, M.J.; et al. Differential progression of coronary atherosclerosis according to plaque composition: A cluster analysis of PARADIGM registry data. *Sci. Rep.* **2021**, *11*, 17121. [CrossRef] [PubMed]

34. Wilkerson, M.D.; Hayes, D.N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **2010**, *26*, 1572–1573. [CrossRef] [PubMed]

35. Thongprayoon, C.; Mao, M.A.; Keddis, M.T.; Kattah, A.G.; Chong, G.Y.; Pattharanitima, P.; Nissaisorakarn, V.; Garg, A.K.; Erickson, S.B.; Dillon, J.J.; et al. Hypernatremia subgroups among hospitalized patients by machine learning consensus clustering with different patient survival. *J. Nephrol.* **2021**. [CrossRef] [PubMed]

36. Thongprayoon, C.; Dumancas, C.Y.; Nissaisorakarn, V.; Keddis, M.T.; Kattah, A.G.; Pattharanitima, P.; Petnak, T.; Vallabhajosyula, S.; Garovic, V.D.; Mao, M.A.; et al. Machine Learning Consensus Clustering Approach for Hospitalized Patients with Phosphate Derangements. *J. Clin. Med.* **2021**, *10*, 4441. [CrossRef]

37. Xue, Y.; Shen, J.; Hong, W.; Zhou, W.; Xiang, Z.; Zhu, Y.; Huang, C.; Luo, S. Risk stratification of ST-segment elevation myocardial infarction (STEMI) patients using machine learning based on lipid profiles. *Lipids Health Dis.* **2021**, *20*, 48. [CrossRef]

38. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

39. Monti, S.; Tamayo, P.; Mesirov, J.; Golub, T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* **2003**, *52*, 91–118. [CrossRef]

40. Șenbabaoğlu, Y.; Michailidis, G.; Li, J.Z. Critical limitations of consensus clustering in class discovery. *Sci. Rep.* **2014**, *4*, 6207. [CrossRef]

41. Stekhoven, D.J.; Bühlmann, P. MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012**, *28*, 112–118. [CrossRef]

42. Alyousef, A.A.; Nihtyanova, S.; Denton, C.; Bosoni, P.; Bellazzi, R.; Tucker, A. Nearest Consensus Clustering Classification to Identify Subclasses and Predict Disease. *J. Healthc. Inform. Res.* **2018**, *2*, 402–422. [CrossRef] [PubMed]

43. Wang, Y.; Zhao, Y.; Therneau, T.M.; Atkinson, E.J.; Tafti, A.P.; Zhang, N.; Amin, S.; Limper, A.H.; Khosla, S.; Liu, H. Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. *J. Biomed. Inform.* **2020**, *102*, 103364. [CrossRef]

44. Huijgen, H.J.; Soesan, M.; Sanders, R.; Mairuhu, W.M.; Kesecioglu, J.; Sanders, G.T. Magnesium levels in critically ill patients. What should we measure? *Am. J. Clin. Pathol.* **2000**, *114*, 688–695. [CrossRef] [PubMed]

45. Zaloga, G.P. Interpretation of the serum magnesium level. *Chest* **1989**, *95*, 257–258. [CrossRef] [PubMed]

46. Zaloga, G.P.; Wilkens, R.; Tourville, J.; Wood, D.; Klyme, D.M. A simple method for determining physiologically active calcium and magnesium concentrations in critically ill patients. *Crit. Care Med.* **1987**, *15*, 813–816. [CrossRef] [PubMed]

47. Al Alawi, A.M.; Majoni, S.W.; Falhammar, H. Magnesium and Human Health: Perspectives and Research Directions. *Int. J. Endocrinol.* **2018**, *2018*, 9041694. [CrossRef] [PubMed]

48. Pham, P.C.; Pham, P.A.; Pham, S.V.; Pham, P.T.; Pham, P.M.; Pham, P.T. Hypomagnesemia: A clinical perspective. *Int. J. Nephrol. Renovasc. Dis.* **2014**, *7*, 219–230. [CrossRef]

49. Thongprayoon, C.; Cheungpasitporn, W.; Hansrivijit, P.; Thirunavukkarasu, S.; Chewcharat, A.; Medaura, J.; Mao, M.A.; Kashani, K.B. Association of serum magnesium level change with in-hospital mortality. *BMJ Evid. Based Med.* **2020**, *25*, 206–212. [CrossRef]

50. Gile, J.J.; Lopez, C.L.; Ruan, G.J.; Hathcock, M.A.; Abeykoon, J.P.; Heimgartner, J.R.; Baumann, N.A.; McMahon, M.M.; Micallef, I.N.; Johnston, P.B.; et al. Hypomagnesemia at the time of autologous stem cell transplantation for patients with diffuse large B-cell lymphoma is associated with an increased risk of failure. *Blood Cancer J.* **2021**, *11*, 65. [CrossRef]
51. Limaye, C.S.; Londhey, V.A.; Nadkart, M.Y.; Borges, N.E. Hypomagnesemia in critically ill medical patients. *J. Assoc. Physicians India* **2011**, *59*, 19–22.
52. Jiang, P.; Lv, Q.; Lai, T.; Xu, F. Does Hypomagnesemia Impact on the Outcome of Patients Admitted to the Intensive Care Unit? A Systematic Review and Meta-Analysis. *Shock* **2017**, *47*, 288–295. [CrossRef]
53. Safavi, M.; Honarmand, A. Admission hypomagnesemia–impact on mortality or morbidity in critically ill patients. *Middle East. J. Anaesthesiol.* **2007**, *19*, 645–660. [PubMed]
54. Del Gobbo, L.C.; Imamura, F.; Wu, J.H.; de Oliveira Otto, M.C.; Chiuve, S.E.; Mozaffarian, D. Circulating and dietary magnesium and risk of cardiovascular disease: A systematic review and meta-analysis of prospective studies. *Am. J. Clin. Nutr.* **2013**, *98*, 160–173. [CrossRef]
55. Cao, Z.; Bideau, R.; Valdes, R., Jr.; Elin, R.J. Acute hypermagnesemia and respiratory arrest following infusion of $MgSO_4$ for tocolysis. *Clin. Chim. Acta* **1999**, *285*, 191–193. [CrossRef]
56. Bokhari, S.R.; Siriki, R.; Teran, F.J.; Batuman, V. Fatal Hypermagnesemia Due to Laxative Use. *Am. J. Med. Sci.* **2018**, *355*, 390–395. [CrossRef] [PubMed]
57. Van Laecke, S. Hypomagnesemia and hypermagnesemia. *Acta Clin. Belg.* **2019**, *74*, 41–47. [CrossRef]