Clinical Kidney Journal, 2023, vol. 16, no. 11, 2072–2081

https://doi.org/10.1093/ckj/sfad189
Advance Access Publication Date: 10 August 2023
Original Article

ORIGINAL ARTICLE

# Epidemiology and predictors of hyponatremia in a contemporary cohort of patients with malignancy: a retrospective cohort study

Miho Murashima [1,*], Kaori Ambe [2,*], Yuka Aoki [2], Takahisa Kasugai [1], Tatsuya Tomonari [1], Minamo Ono [1], Masashi Mizuno [1], Masahiro Tohkin [2] and Takayuki Hamano [1,3]

[1]Department of Nephrology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan, [2]Department of Regulatory Science, Nagoya City University Graduate School of Pharmaceutical Sciences, Nagoya, Japan and [3]Department of Nephrology, Osaka University Graduate School of Medicine, Osaka, Japan

*M.Mu. and K.A. contributed equally to the work.
Correspondence to: Miho Murashima; E-mail: mmurashi@med.nagoya-cu.ac.jp

## ABSTRACT

**Background.** Hyponatremia is associated with worse outcomes among patients with malignancy. However, contemporary cohort data on epidemiology and risk factors are lacking.
**Methods.** In this single-centre, retrospective cohort study, patients who received intravenous antineoplastic agents from 2018 to 2020 at Nagoya City University Hospital were enrolled. Associations of demographics, antineoplastic agents, types of malignancy and concomitant medications with hyponatremia, defined as serum sodium concentration ≤130 mmol/l, were analysed by mixed-effects logistic regression and the machine learning–based LightGBM model artificial intelligence technology.
**Results.** Among 2644 patients, 657 (24.8%) developed at least one episode of hyponatremia. Approximately 80% of hyponatremia was due to sodium wasting from the kidneys. Variables associated with hyponatremia both by mixed-effects logistic regression and the LightGBM model were older age, hypoalbuminemia and higher estimated glomerular filtration rate. Among antineoplastic agents, cisplatin {odds ratio [OR] 1.52 [95% confidence interval (CI) 1.18–1.96]}, pembrolizumab [OR 1.42 (95% CI 1.02–1.97)] and bortezomib [OR 3.04 (95% CI 1.96–4.71)] were associated with hyponatremia and these variables also had a positive impact on predicted hyponatremia in the LightGBM model.
**Conclusions.** Hyponatremia was common among patients with malignancy. In addition to older age and poor nutritional status, novel antineoplastic agents, including immune checkpoint inhibitors and bortezomib, should be recognized as risk factors for hyponatremia.

## LAY SUMMARY

In this study we demonstrated that hyponatremia is common among patients with malignancy in a contemporary cohort. Most of hyponatremia was caused by sodium wasting from the kidneys. In addition to known risk factors of hyponatremia, such as older age, poor oral intake, vinca alkaloids, small cell lung cancer and pancreatic cancer,

Received: 24.2.2023; Editorial decision: 11.7.2023

immune checkpoint inhibitors and bortezomib were associated with a higher incidence of hyponatremia both by conventional statistics and by a machine learning prediction model.

**Keywords:** bortezomib, hyponatremia, immune checkpoint inhibitors, machine learning, malignancy

## INTRODUCTION

Hyponatremia affects nearly half of hospitalized patients with cancer [1] and is associated with a poor prognosis [1–4]. Hyponatremia in cancer patients can have a variety of aetiologies, including poor oral intake, syndrome of inappropriate antidiuretic hormone secretion (SIADH) and sodium (Na) wasting from the kidneys due to tubular damage, and often a combination of these [4]. Various factors have been reported to be associated with hyponatremia, including types of malignancy (small cell lung cancer, head and neck cancer and pancreatic cancer) [2–7] and antineoplastic agents (vinca alkaloids [8], cyclophosphamide [9] and cisplatin [10–12]).

Since the publication of these previous studies, many antineoplastic agents, including immune checkpoint inhibitors and molecular targeted agents, have become available in clinical practice. However, studies on epidemiology, clinical characteristics and antineoplastic agents associated with hyponatremia in more contemporary cohorts are lacking. It is probably because of the large number of antineoplastic agents and the complexity of the combination of these agents used to treat different types of malignancy. As a result, recent studies on hyponatremia among patients with malignancy are limited to certain agents [13, 14].

In this study we utilized a mixed effects logistic regression model to identify antineoplastic agents and other factors associated with hyponatremia among a wide range of oncologic patients in real clinical practice. Moreover, we developed a machine learning–based LightGBM model of hyponatremia and explain the model.

## MATERIALS AND METHODS

### Study design

This is a single-centre, retrospective cohort study. The study protocol and a waiver of consent were approved by the Nagoya City University Institutional Review Board (approval 60-21-0003) and the study was conducted in accordance with the Helsinki Declaration.

### Setting and patients

Inclusion criteria were patients treated with intravenous antineoplastic agents from January 2018 to September 2020 at Nagoya City University Hospital. Exclusion criteria were those without measurements of serum Na concentration during the study period.

### Exposures of interests

Exposures of interest were antineoplastic agents listed in Table 1, age, sex, estimated glomerular filtration rate (eGFR), serum albumin, types of malignancy and concomitant medications that could potentially affect electrolytes. eGFR was calculated by an equation developed for the Japanese population. The accuracy within 30% of measured GFR was 75% [15].

### Outcomes

The outcome was hyponatremia, defined as a serum Na concentration ≤130 mmol/l [16, 17]. Serum Na levels were measured by an indirect method. Additional analyses were performed after correcting Na levels for hyperglycaemia {corrected Na (mmol/l) = 1.6*[glucose (mg/dl) – 200]/100} if serum glucose levels were >200 mg/dl [18].

### Statistical analyses

Data were shown as *n* (%), mean [standard deviation (SD)] or median [interquartile range (IQR)]. Descriptive statistics were determined by Fisher's exact test, *t*-test or Mann–Whitney U test as appropriate. To identify variables associated with hyponatremia, a mixed effects logistic regression model was used rather than multivariate logistic regression, as patients had multiple measurements of Na at different timings with different exposures to covariates, while each patient has a different preponderance for developing hyponatremia. Mixed effects logistic regression is suited when there are multiple observations for each unit of observation. This scenario also includes when there is a clustered nature to the data. eGFR and serum albumin values measured at a previous visit were used as covariates (fixed effects) in the model, as kidney function and nutritional status before the development of hyponatremia can determine the likelihood of developing hyponatremia. Age, sex, types of malignancy, antineoplastic agents and other medications were also treated as fixed effects and patients were treated as a random effect. We conducted a sensitivity analysis using a mixed effects Poisson regression model, with the number of Na measurements in each patient (natural log-transformed) as an exposure variable to account for different frequencies of Na measurements among patients.

Statistical analyses were performed using Stata MP version 17 (StataCorp, College Station, TX, USA).

### Machine learning–based prediction model

#### Datasets

We used clinical examination–level data for model construction, removing the data without serum Na level, serum albumin level and eGFR at the previous examination. The positive data were hyponatremia, i.e. serum Na concentration ≤130 mmol/l. The negative data were non-hyponatremia, i.e. serum Na concentration >130 mmol/l. We randomly split these data 4:1, maintaining positive and negative ratios to create two independent datasets, assigning one for the training dataset and one for the test dataset.

#### Classification model for hyponatremia

We constructed the prediction model for classification of hyponatremia using the LightGBM algorithm of machine learning [19]. It is one of the gradient-boosting decision tree algorithms and has high performance among non-neural network systems.

**Table 1: Demographics.**

| Characteristics | No hyponatremia ($n = 1987$) | Hyponatremia[a] ($n = 657$) | $P$-value |
|---|---|---|---|
| Age (years), mean (SD) | 63.9 (14.2) | 66.9 (13.9) | <.001 |
| Male, *n* (%) | 1051 (52.9) | 423 (64.4) | <.001 |
| eGFR (ml/min/1.73 $m^2$)[b], median (IQR) | 71.0 (59.2–83.7) ($n = 1951$) | 69.7 (55.5–84.6) ($n = 643$) | .21 |
| Albumin (g/dl)[b], mean (SD) | 4.0 (0.5) ($n = 1802$) | 3.8 (0.6) ($n = 604$) | <.001 |
| Malignancy, *n* (%) | | | |
| Brain | 17 (0.9) | 5 (0.8) | 1.00 |
| Head and neck | 160 (8.1) | 76 (11.6) | .006 |
| Thyroid | 7 (0.4) | 3 (0.5) | .72 |
| Small cell lung cancer | 34 (1.7) | 17 (2.6) | .16 |
| Non-small cell lung cancer | 105 (5.3) | 52 (7.9) | .01 |
| Lung, unspecified | 203 (1.0) | 91 (13.9) | .01 |
| Mesothelioma | 1 (0.05) | 4 (0.6) | .02 |
| Breast | 362 (18.2) | 31 (4.7) | <.001 |
| Oesophageal | 107 (5.4) | 46 (7.0) | .12 |
| Gastric | 137 (6.9) | 54 (8.2) | .26 |
| Hepatocellular | 93 (4.7) | 21 (3.2) | .10 |
| Pancreatic | 85 (4.3) | 53 (8.1) | <.001 |
| Biliary | 22 (1.1) | 19 (2.9) | .001 |
| Duodenal | 4 (0.2) | 4 (0.6) | .11 |
| Colon | 242 (12.2) | 90 (13.7) | .31 |
| Renal cell | 42 (2.1) | 19 (2.9) | .25 |
| Genitourinary | 168 (8.5) | 53 (8.1) | .76 |
| Prostate | 134 (6.7) | 34 (5.2) | .15 |
| Endometrial | 66 (3.3) | 5 (0.8) | <.001 |
| Cervical | 47 (2.4) | 10 (1.5) | .20 |
| Ovarian | 74 (3.7) | 19 (2.9) | .32 |
| Lymphoma | 213 (10.7) | 89 (13.5) | .05 |
| Leukaemia | 54 (2.7) | 43 (6.5) | <.001 |
| Myeloma | 108 (5.4) | 53 (8.1) | .01 |
| Sarcoma | 52 (2.6) | 18 (2.7) | .87 |
| Neuroendocrine | 11 (0.6) | 6 (0.9) | .40 |
| Non-melanoma skin cancer | 37 (1.9) | 15 (2.3) | .50 |
| Melanoma | 32 (1.6) | 9 (1.4) | .67 |
| Cancer of unknown origin | 26 (1.3) | 11 (1.7) | .49 |
| Antineoplastic agents | | | |
| Vinca alkaloids, *n* (%) | | | |
| Docetaxel | 210 (10.6) | 72 (11.0) | .78 |
| Vincristine | 123 (6.2) | 68 (10.4) | <.001 |
| Irinotecan | 101 (5.1) | 57 (8.7) | .001 |
| Etoposide | 91 (4.6) | 62 (9.4) | <.001 |
| Paclitaxel | 456 (22.9) | 140 (21.3) | .39 |
| Antimetabolites, *n* (%) | | | |
| 5-fluorouracil | 242 (12.2) | 123 (8.7) | <.001 |
| Gemcitabine | 165 (8.3) | 103 (15.7) | <.001 |
| Cytarabine | 69 (3.5) | 55 (8.4) | <.001 |
| Methotrexate | 70 (3.5) | 42 (6.4) | .002 |
| Alkylating agents, *n* (%) | | | |
| Cyclophosphamide | 307 (15.5) | 87 (13.2) | .17 |
| Anthracyclines, *n* (%) | | | |
| Doxorubicin | 266 (13.4) | 79 (12.0) | .37 |
| Platinum, *n* (%) | | | |
| Carboplatin | 266 (13.4) | 121 (18.4) | .002 |
| Cisplatin | 337 (17.0) | 149 (22.7) | .001 |
| Oxaliplatin | 236 (11.9) | 80 (12.2) | .84 |
| Molecular targeted drugs, *n* (%) | | | |
| Bortezomib | 51 (2.6) | 45 (6.8) | <.001 |
| Bevacizumab | 158 (8.0) | 61 (9.3) | .28 |
| Cetuximab | 50 (2.5) | 32 (4.9) | .003 |
| Daratumumab | 47 (2.4) | 20 (3.0) | .34 |
| Trastuzumab | 91 (4.6) | 6 (0.9) | <.001 |

Table 1: (Continued).

| Characteristics | No hyponatremia ($n = 1987$) | Hyponatremia[a] ($n = 657$) | P-value |
|---|---|---|---|
| Rituximab | 124 (6.2) | 51 (7.8) | .17 |
| Immune checkpoint inhibitors | | | |
| Atezolizumab | 34 (1.7) | 22 (3.3) | .01 |
| Nivolumab | 105 (5.3) | 68 (10.4) | <.001 |
| Pembrolizumab | 85 (4.3) | 54 (8.2) | <.001 |
| Ipilimumab | 23 (1.2) | 10 (1.5) | .47 |
| Use of other medications, *n* (%) | | | |
| ACE inhibitors | 58 (2.9) | 37 (5.6) | .001 |
| ARBs | 302 (15.2) | 133 (20.2) | .002 |
| Loop diuretics | 285 (14.3) | 177 (26.9) | <.001 |
| Thiazide diuretics | 37 (1.9) | 38 (5.8) | <.001 |
| Potassium-sparing diuretics | 94 (4.7) | 73 (11.1) | <.001 |
| Tolvaptan | 4 (0.2) | 8 (1.2) | .003 |
| Acetazolamide | 10 (0.5) | 3 (0.5) | 1.00 |
| Denosumab | 199 (10.0) | 79 (12.0) | .15 |
| Bisphosphonates | 144 (7.2) | 89 (13.5) | <.001 |
| Vitamin D | 106 (5.3) | 48 (7.3) | .06 |

[a]The development of hyponatremia at any time during the observation period.
[b]The first measurement of the values during the observation period.
P-values were by Fisher's exact test, *t*-test or Mann–Whitney U test.

It operates quickly and is strong for large datasets. It is also an excellent algorithm that does not require data standardization and can handle missing values. However, in our dataset, the number of missing values was small and data with missing values were removed for model building. The objective variable was the presence of hyponatremia and the explanatory variables were age, sex, antineoplastic agents, concomitant medications, types of malignancy, serum albumin at a previous visit and eGFR at a previous visit.

First, we constructed a model using a training dataset. Data were pre-processed to address differences in the number of repeated tests per patient and imbalance issues between positive and negative data. Specifically, since the number of measurements per patient varies, the inverse of the number of measurements was used as the weight of the data. In the LightGBM algorithm, the 'weight' parameter and data with higher weights, i.e. patients' data with a small number of inspections overall, were emphasized in the model training phase (https://lightgbm.readthedocs.io/en/latest/Parameters.html). Adjustment for imbalanced data was performed only on the training dataset. A very large number of negative data compared with positive data were reduced by undersampling to equal the number of positives. After negative data clustering by the k-means method, the number of data in each cluster was adjusted to equal the number of positive data by resampling while maintaining the ratio of the number of data in each cluster [20]. We performed 5-fold cross-validation (CV) in the training dataset to identify the optimal hyperparameters through a Bayesian optimization using Optuna [21]. Second, the best-performing parameters were used to predict the test data with a model built on the entire training dataset. We evaluated performance using calculated sensitivity, specificity and receiver operating characteristics curve area under the curve (ROC-AUC) in the training dataset and test datasets. The training dataset evaluation used the average of a 5-fold CV.

Explainable artificial intelligence (AI) is a technique for explaining the predictive results of complex machine learning models. In this study we implemented the Shapley Additive Explanations (SHAP) based on Shapley values, which is a unified approach of any machine learning model, as a method to calculate the contribution of each feature to the predictions [22].

### Software

Predictive model building was conducted using the following library in Python for Windows (version 3.8.5): lightgbm for LightGBM (version 3.2.1), scikit-learn (version 0.24.2), NumPy (version 1.20.3) and Pandas (version 1.2.4) for data processing.

## RESULTS

### Patients' characteristics

Among 2709 patients who received intravenous antineoplastic agents during the study period, 2644 had data for serum Na concentration. Among them, 657 patients (24.8%) developed hyponatremia at least once during the study period. Serum Na was measured 101 622 times (38 measurements per patient on average) and 2854 events of hyponatremia were observed. Demographics for those who developed hyponatremia at least once and those who never developed hyponatremia are shown in Table 1. Those with hyponatremia were significantly older, more likely to be male and more likely to have head and neck, lung, pancreatic, biliary cancer, leukaemia and myeloma. Those with hyponatremia were also more likely to have received vincristine, irinotecan, etoposide, gemcitabine, cytarabine, methotrexate, cisplatin, carboplatin, bortezomib, cetuximab, atezolizumab, nivolumab, pembrolizumab, angiotensin-converting enzyme (ACE) inhibitors, angiotensin receptor blockers (ARBs), loop diuretics, thiazide diuretics, potassium-sparing diuretics, tolvaptan and bisphosphonates.

### Urinary findings and endocrinological evaluations in hyponatremia

Urine Na concentration and urine osmolarity were measured in 215 (32.7%) and 197 (30.0%) patients within 30 days from the first event of hyponatremia. The lowest mean serum Na concentrations among those with and without measurements of urine

**Table 2: Association between demographics, antineoplastic agents and other medications with hyponatremia (mixed effects logistic regression).**

| Variables | OR (95% CI) |
|---|---|
| Age | **1.03 (1.02–1.05)** |
| Male | 1.09 (0.82–1.45) |
| eGFR (per ml/min/1.73 $m^2$)[a] | **1.002 (1.000–1.004)** |
| Albumin (per g/dl)[a] | **0.38 (0.35–0.41)** |
| Malignancy | |
| Brain | 1.47 (0.31–7.03) |
| Head and neck | 1.28 (0.87–1.86) |
| Thyroid | 3.92 (0.68–22.69) |
| Small cell lung cancer | **2.42 (1.12–5.22)** |
| Non-small cell lung cancer | 1.34 (0.81–2.25) |
| Lung, unspecified | 1.12 (0.76–1.66) |
| Mesothelioma | 4.37 (0.63–16.55) |
| Breast | **0.54 (0.31–0.95)** |
| Oesophageal | 1.02 (0.62–1.68) |
| Gastric | 0.91 (0.59–1.40) |
| Hepatocellular | 0.56 (0.30–1.07) |
| Pancreatic | **1.77 (1.10–2.82)** |
| Biliary | **3.40 (1.68–6.89)** |
| Duodenal | 0.70 (0.14–3.58) |
| Colon | 1.13 (0.78–1.62) |
| Renal cell | 1.02 (0.47–2.19) |
| Genitourinary | 0.92 (0.58–1.46) |
| Prostate | **0.54 (0.32–0.92)** |
| Endometrial | 0.37 (0.12–1.16) |
| Cervical | 1.99 (0.78–5.06) |
| Ovarian | 1.90 (0.90–3.99) |
| Lymphoma | 1.16 (0.80–1.70) |
| Leukaemia | 2.56 (0.89–2.74) |
| Myeloma | 0.60 (0.35–1.04) |
| Sarcoma | 0.47 (0.21–1.07) |
| Neuroendocrine | 2.18 (0.60–7.93) |
| Non-melanoma skin cancer | 0.58 (0.25–1.34) |
| Melanoma | 0.60 (0.21–1.68) |
| Cancer of unknown origin | 0.61 (0.24–1.61) |
| Antineoplastic agents | |
| Vinca alkaloids | |
| Docetaxel | **1.09 (1.27–2.26)** |
| Vincristine | 1.61 (0.99–2.62) |
| Irinotecan | **1.65 (1.09–2.50)** |
| Etoposide | 0.70 (0.48–1.01) |
| Paclitaxel | 0.94 (0.73–1.21) |
| Antimetabolites | |
| 5-fluorouracil | 0.80 (0.61–1.19) |
| Gemcitabine | 1.09 (0.84–1.42) |
| Cytarabine | 0.91 (0.60–1.38) |
| Methotrexate | 0.93 (0.61–1.43) |
| Alkylating agents | |
| Cyclophosphamide | **1.61 (1.10–2.37)** |
| Anthracyclines | |
| Doxorubicin | 0.70 (0.49–1.00) |
| Platinum | |
| Carboplatin | 0.83 (0.64–1.09) |
| Cisplatin | **1.52 (1.18–1.96)** |
| Oxaliplatin | **0.65 (0.45–0.95)** |
| Molecular targeted drugs | |
| Bortezomib | **3.04 (1.96–4.71)** |
| Bevacizumab | 0.70 (0.47–1.04) |
| Cetuximab | **1.66 (1.04–2.65)** |
| Daratumumab | 1.03 (0.65–1.64) |
| Trastuzumab | 1.74 (0.69–4.42) |
| Rituximab | **0.64 (0.42–0.97)** |
| Immune checkpoint inhibitors | |
| Atezolizumab | 1.38 (0.78–2.45) |
| Nivolumab | **1.37 (1.00–1.88)** |
| Pembrolizumab | **1.42 (1.02–1.97)** |
| Ipilimumab | **2.30 (1.00–5.29)** |
| Use of other medications | |
| ACE inhibitors | **1.66 (1.10–2.50)** |
| ARBs | 1.13 (0.87–1.46) |
| Loop diuretics | **1.86 (1.49–2.32)** |
| Thiazide diuretics | 0.73 (0.46–1.15) |
| Potassium-sparing diuretics | **1.47 (1.04–2.08)** |
| Tolvaptan | 0.52 (0.21–1.26) |
| Acetazolamide | 1.21 (0.24–6.14) |
| Denosumab | 1.28 (0.97–1.70) |
| Bisphosphonates | 1.28 (0.94–1.74) |
| Vitamin D | 0.79 (0.52–1.17) |

[a]The values at the previous visits were used.
Values in bold are statistically significant.

Na concentration were 126 mmol/l (SD 4) and 127 mmol/l (SD 4), respectively ($P < .001$). Although statistically significant, the differences in serum Na concentration between those with and without urine Na or urine osmolarity measurements were small. Urine Na concentration and urine osmolarity closest to the first event of hyponatremia were identified. Among those with data for urine Na or osmolarity, 80.5% had a urine Na concentration $\geq$30 mEq/l and 99.0% had a urine osmolarity of >100 mOsm/kg $H_2O$, suggesting that most cases of hyponatremia were due to Na wasting from the kidneys [16]. The number of patients with measurements of cortisol and free thyroxine within 30 days from the development of hyponatremia was 27 and 58, of which 2 and 3, respectively, had values below the reference ranges.

## Factors associated with hyponatremia by mixed effects models

By mixed effects logistic regression analyses, older age, eGFR, lower serum albumin levels, small cell lung cancer, pancreatic cancer, biliary cancer, docetaxel, irinotecan, cyclophosphamide, cisplatin, bortezomib, cetuximab, nivolumab, pembrolizumab, ipilimumab, ACE inhibitors, loop diuretics and potassium-sparing diuretics were associated with a higher incidence of hyponatremia. Association with hyponatremia was the strongest for biliary cancer among various malignancies and the strongest for bortezomib among antineoplastic agents. On the other hand, breast cancer, prostate cancer, oxaliplatin and rituximab were associated with a lower incidence of hyponatremia (Table 2). Correcting Na levels for hyperglycaemia and further adjustment for the amount and kinds of intravenous fluids within 2 days prior to Na measurements did not substantially change the results, while normal saline or lactated Ringer's was associated with a lower incidence of hyponatremia and half normal saline was associated with a higher incidence of hyponatremia. Sensitivity analyses by mixed effects Poisson regression analyses yielded similar results (Supplementary Table S1).

## Machine learning–based prediction model

We developed a machine learning–based classification model for hyponatremia based on clinical examination–level data.

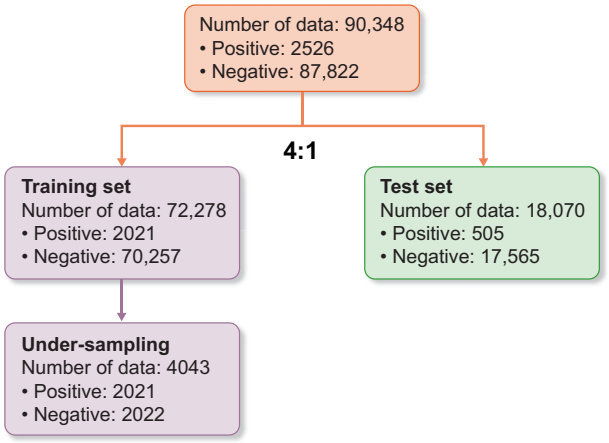


**Figure 1:** The flow of data for a machine learning–based LightGBM model.

This model can predict whether or not hyponatremia is likely to occur at the next examination. The total number of data was 90 348. We detected 2526 hyponatremia-positive data and 87 822 hyponatremia-negative data. For the training dataset, a total of 72 278 data existed in the original dataset, which was comprised of 2021 positive and 70 257 negative data. Subsequently, with undersampling, 4043 data remained in the dataset (2021 positive and 2022 negative data). The test dataset included 18 070 data, which comprised 505 positive and 17 565 negative data (Fig. 1). For the LightGBM model, the training dataset, which resolved the data imbalance, had a sensitivity of 0.776, specificity of 0.706, positive predictive value (PPV) of 0.725, negative predictive value (NPV) of 0.761 and ROC-AUC of 0.818. In contrast, the test dataset had a sensitivity of 0.834, specificity of 0.688, PPV of 0.071, NPV of 0.993 and ROC-AUC of 0.827. The test dataset remained imbalanced, with an extremely low number of positive data compared with negative data, resulting in a low PPV. Similar ROC-AUCs in the training and test datasets suggest that the predictive model was able to avoid overfitting (too much learning in the training dataset) [23]. In the test dataset, the sensitivity and NPV were high, so our model provides reliable predictions with few misses. We performed a Bayesian optimization using an Optuna search in various ranges and evaluated the log loss function of five cross-validations as the evaluation criterion. The seven adjusted parameter values adopted were the following: 'learning_rate', 0.01; 'num_leaves', 31; 'min_data_in_leaf', 16; 'lambda_1', 3.8773179281026535e-07; 'lambda_2', 2.062350675098732e-07; 'feature_fraction', 0.48000000000000004; 'bagging_fraction', 0.9253610810045878; and 'bagging_freq', 0.8. Other parameters were set to the default values.

Figure 2 shows the main variables of the LightGBM model. The ranking of the SHAP values in the final model with the test dataset is shown. The strongest predictors of hyponatremia were albumin levels at a previous visit, followed by age, eGFR at a previous visit, sex, loop diuretics and cisplatin. Concomitant use of certain medications (carboplatin, gemcitabine, pembrolizumab and vincristine) and types of malignancy (lung cancer, oesophageal cancer and pancreatic cancer) were also among the main predictors of hyponatremia. Figure 2 shows the SHAP value plot and SHAP values, which can show whether the variables contributed positively or negatively to the prediction. Variables are sorted by their mean absolute SHAP values in descending order, with the most important variable at the top. Each dot corresponds to one data in our study. The colour of

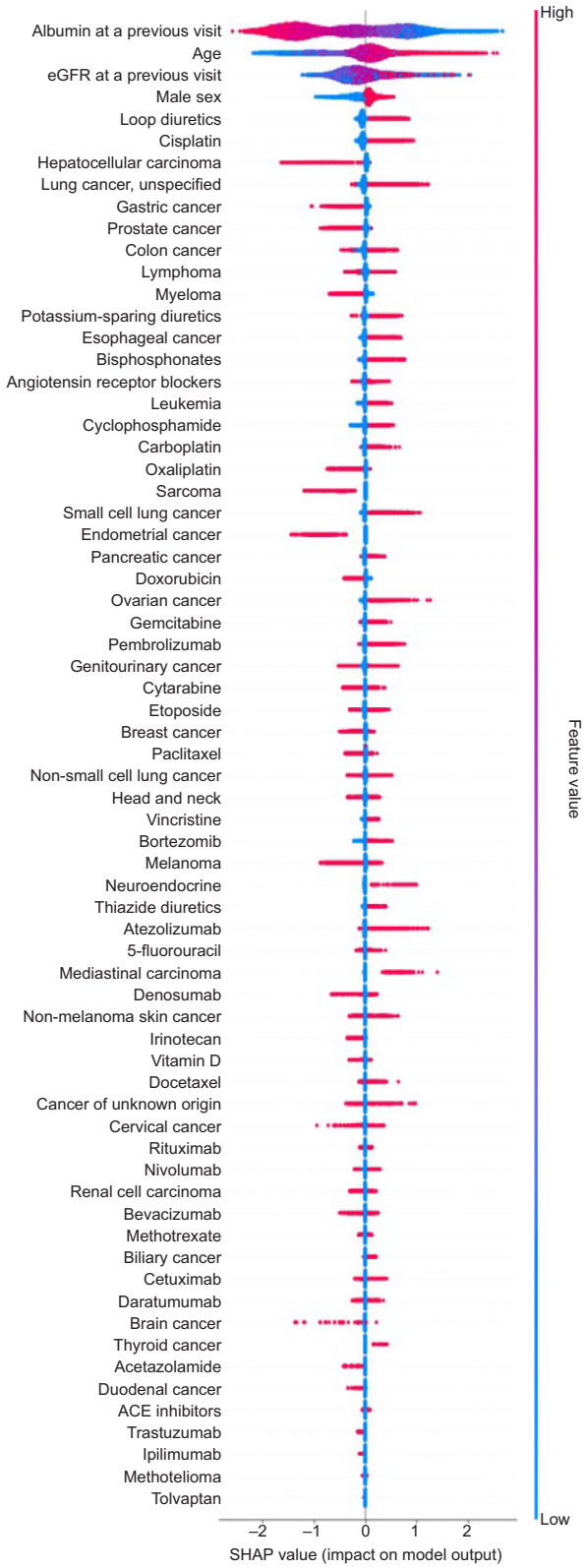


**Figure 2:** SHAP summary plot of the LightGBM model for the test dataset. Variables are sorted by their mean absolute SHAP values in descending order with the most important variable at the top. The colour of the dot indicates the magnitude of each variable. Higher values are red, lower values are blue. If the SHAP value on the horizontal axis is positive, it increases the predictive value, i.e. it predicts that hyponatremia is likely to occur. This plot shows that on average 'albumin at a previous visit' is the most important feature and patients with low albumin values (blue) are more likely to develop hyponatremia.

**Table 3: Comparison of predictors for hyponatremia by mixed effects logistic regression and AI.**

| Variables significantly associated with hyponatremia by mixed effects logistic regression and with high SHAP scores | Variables not significantly associated with hyponatremia by mixed effects logistic regression but with high SHAP score and with the point of estimates at the same direction of impact on model output by AI | Variables with the point of estimates at the different directions of impact on model output by AI |
|---|---|---|
| Albumin levels | Male sex | Carboplatin |
| Age | Hepatocellular carcinoma | Lymphoma |
| eGFR | Lung cancer, unspecified | |
| Loop diuretics | Gastric cancer | |
| Cisplatin | Colon cancer | |
| SCLC | Gemcitabine | |
| Prostate cancer | Myeloma | |
| Pembrolizumab | ARB | |
| Pancreas | Bisphosphonates | |
| Potassium-sparing diuretics | Oesophagus | |
| Oxaliplatin | NSCLC | |
| Bortezomib | Vincristine | |
| | Sarcoma | |
| | Leukaemia | |
| | Endometrial cancer | |

SCLC: small cell lung cancer; NSCLC: non-small cell lung cancer.

the dot indicates the magnitude of the value of the variable. Higher values are red, lower values are blue. If the SHAP value on the horizontal axis is positive, it increases the predictive value, i.e. it predicts that hyponatremia is likely to occur. Supplementary Table S2 shows the mean SHAP value of each variable.

### Comparison of results by a mixed effects model and by AI

Table 3 shows the comparison of results by a mixed effects model and an AI technology-based LightGBM model. Variables associated with a higher incidence of hyponatremia both by mixed effects model and LightGBM model were older age, lower albumin, higher eGFR, loop diuretics, cisplatin, small cell lung cancer, pembrolizumab, pancreatic cancer, potassium-sparing diuretics and bortezomib. In contrast, carboplatin was associated with a lower incidence of hyponatremia by the mixed-effects logistic regression model but a higher incidence of hyponatremia by the LightGBM model.

## DISCUSSION

This study showed that hyponatremia is common among a contemporary cohort of patients undergoing treatment for malignancy affecting ≈25% of them. Most hyponatremia cases were likely due to Na wasting from the kidneys. In addition to known risk factors for hyponatremia, we have shown that novel antineoplastic agents, including bortezomib, cetuximab and immune checkpoint inhibitors, were significantly associated with hyponatremia.

The incidence of hyponatremia was similar to that in previous studies, although the aetiology of hyponatremia seems to be different versus old studies. In our study, 25% of patients with malignancy (both inpatients and outpatients) developed hyponatremia during treatment, which is similar to previous studies showing that 11–34% of hospitalized cancer patients had hyponatremia (Na <130 mEq/l) [1, 3]. In terms of the aetiology of hyponatremia, one previous study showed that the causes of hyponatremia among cancer patients were as follows: SIADH 30%, volume depletion 28%, diuretic use 14% and hypervolemia 8% [24]. In our study, ≈20% of patients with hyponatremia had urine Na <30 mmol/l and urine osmolarity >100 mOsm/kg $H_2O$, which was compatible with either volume depletion or impaired kidney perfusion due to heart failure, nephrotic syndrome or cirrhosis. A total of 80% of patients with hyponatremia in our study had urine Na >30 mmol/l and urine osmolarity >100 mOsm/kg $H_2O$, suggesting that the aetiologies of their hyponatremia were SIADH, diuretic use or Na wasting nephropathy from tubular injury [12, 16]. The proportion of patients with hyponatremia from Na wasting from the kidneys seems to be higher than in the previous study. However, it should be noted that only one-third of patients with hyponatremia had urine Na and urine osmolarity measured and the number of patients with measurements of cortisol and thyroid function was small. Our data suggest that the importance of differentiating the causes and managing hyponatremia is underrecognized.

In this study, we used mixed effects logistic regression and the mixed effects Poisson model to increase the statistical power and make it possible to include a large number of antineoplastic agents as potential predictors of hyponatremia. By using mixed effects Poisson models, we also considered the effect of different frequencies of Na measurements. Another difficulty in identifying predictors of hyponatremia among patients with malignancy is collinearity, i.e. certain antineoplastic agents are used for certain malignancies or certain antineoplastic agents are used in combination with a particular antineoplastic agent. In such cases, it is difficult to determine whether a certain antineoplastic agent or a certain malignancy is associated with hyponatremia. LightGBM includes the least absolute shrinkage and selection operator, which incorporates regularization terms and processes to select feature values, in its algorithm. Thus it suffers less from the problems of collinearity [25]. Although the mixed effects models and the LightGBM model cannot provide a causal relationship, similar results obtained by two different methods made our results robust. Although collinearity is not a big issue when it comes to prediction, which is the main purpose of machine learning models, even in machine learning

models, if there is a pair of feature values with a strong correlation, the impact of the feature value and the action of each feature value is proportionally divided between the pairs, and both are underestimated. We therefore checked the correlation coefficients of the explanatory variables used in the model, none of which were >0.8. Variables of low importance, i.e. SHAP values of 0 (testicular cancer, zoledronate, intestinal cancer), were also removed from the final model. In a machine learning model, the impact of each variable on the prediction model could be evaluated by SHAP values. SHAP values used in this study evaluate the importance of the output resulting from the inclusion of variable 'A' for all combinations of features other than 'A'. It ensures high local accuracy, stability against missing data and consistency in feature impact.

Some of the factors identified to be associated with hyponatremia in our study were known to be risk factors of hyponatremia. Older age [26–28], diuretics [26–28], ACE inhibitors [28, 29], ARBs [29], small cell lung cancer [2–5], pancreatic cancer [2–4, 7], biliary tract cancer [2], vinca alkaloids [4, 8], cyclophosphamide [4, 9] and cisplatin [4, 12] were reported to be risk factors for hyponatremia. Hypoalbuminemia likely reflects poor oral intake, which is a risk factor for hyponatremia [4, 26]. Although advanced kidney diseases are associated with impaired ability of urine dilution [30], and thus a risk factor for hyponatremia, in our study, higher eGFR was significantly associated with hyponatremia. In this study the median eGFR was ≈70 ml/min/1.73 $m^2$, suggesting that most of the patients had preserved kidney function. Higher eGFR in our study might reflect lower muscle mass and creatinine due to poor oral intake, which is one of the risk factors for hyponatremia and leads to an overestimation of eGFR. Also, patients with low muscle mass are likely to have low bone mass, which is the largest store of Na in the body. It is known that antidiuretic hormone activates osteoclasts and causes bone resorption, which provides Na from bone to circulation [31, 32]. Those with low bone mass might not be able to release enough Na from bone even in the setting of hyponatremia with elevated antidiuretic hormone.

In addition to previously known risk factors for hyponatremia, we identified that immune checkpoint inhibitors, cetuximab and bortezomib were significantly associated with hyponatremia. Hyponatremia during the treatment with these agents has been reported [13, 14, 33–36] and hyponatremia is described as one of the side effects for atezolizumab, bortezomib, pembrolizumab and nivolumab, but not for cetuximab, in the summaries of product characteristics. However, the strength of associations between these agents and hyponatremia have not been recognized.

In our study, breast cancer, prostate cancer, oxaliplatin and rituximab were associated with a significantly lower incidence of hyponatremia. The reasons for associations were unclear. Also, the point of estimates by the mixed effects model and the direction of impact on the LightGBM model by AI was different for carboplatin. This could be due to co-linearity in the mixed effects model.

The strength of our study is the inclusion of a wide range of oncologic patients, the use of the mixed-effects model to enable the inclusion of a large number of covariates, including types of malignancy and antineoplastic agents, and the use of AI to account for co-linearity between certain malignancies and certain antineoplastic agents. Furthermore, the machine learning–based LightGBM model has the advantage of predicting hyponatremia. This AI model is expected to support the early detection of hyponatremia in clinical practice.

The limitations of the study need to be acknowledged. The study is a single-centre retrospective study and external validity is uncertain. Even with the use of AI, it is impossible to prove a causal relationship between antineoplastic agents and hyponatremia. The serum Na levels were measured indirectly and the number of patients with serum osmolarity measurements within 1 week of development of hyponatremia was only 28. Pseudohyponatremia could have been included in hyponatremia in this study. Although we included medications that can potentially affect serum Na levels, it was impossible to include all the medications that can potentially cause SIADH.

In conclusion, hyponatremia was common among patients undergoing treatment for malignancy. Most cases of hyponatremia were likely due to Na wasting from the kidneys. In addition to older age, poor nutritional status and medications known to cause hyponatremia, we demonstrated that immune checkpoint inhibitors, cetuximab and bortezomib were independently associated with hyponatremia. We should recognize that these agents are associated with hyponatremia and vigilant monitoring of serum Na is warranted.

## SUPPLEMENTARY DATA

Supplementary data are available at *ckj* online.

## ACKNOWLEDGEMENTS

Part of the study results were presented as an abstract at the Annual Meeting of the Japanese Society of Nephrology in 2021, American Society of Nephrology, Kidney Week in 2021 and Regulatory Science Forum for Young Investigators in 2022.

## FUNDING

This work was supported by a Grant-in-Aid for Research from Nagoya City University (grant 2122104) and Grant-in-Aid for Scientific Research (KAKENHI) (C) (grant 21K08258).

## AUTHORS' CONTRIBUTIONS

M.Mu. was responsible for conception of the study design, data synthesis, data analysis and drafting of the manuscript. K.A. was responsible for data analysis and drafting of the manuscript. Y.A. was responsible for data analysis. T.K., T.T., M.O. and M.Mi. were responsible for review and approval of the manuscript. M.T. and T.H. were responsible for supervision, review and approval of the manuscript.

## DATA AVAILABILITY STATEMENT

The data are available upon reasonable request to the authors.

## CONFLICT OF INTEREST STATEMENT

T.H. has received honoraria from Kissei Pharmaceutical, Otsuka Pharmaceutical, Chugai Pharmaceutical and Kyowa Kirin; research grants from Otsuka Pharmaceutical and Chugai Pharmaceutical and lecture fees from Chugai Pharmaceutical, Sumitomo Pharma, Kyowa Kirin, Kissei Pharmaceutical, Ono Pharmaceutical, Otsuka Pharmaceutical, AstraZeneca KK and Astellas Pharma.

## REFERENCES


1. Doshi SM, Shah P, Lei X *et al*. Hyponatremia in hospitalized cancer patients and its impact on clinical outcomes. *Am J Kidney Dis* 2012;**59**:222–8. https://doi.org/10.1053/j.ajkd.2011.08.029
2. Fucá G, Mariani L, Lo Vullo S *et al*. Weighing the prognostic role of hyponatremia in hospitalized patients with metastatic solid tumors: the HYPNOSIS study. *Sci Rep* 2019;**9**:12993. https://doi.org/10.1038/s41598-019-49601-3
3. Zeinah GF, Al-Kindi SG, Hassan AA *et al*. Hyponatremia in cancer: association with type of cancer and mortality. *Eur J Cancer Care (Engl)* 2015;**24**:224–31. https://doi.org/10.1111/ecc.12187
4. Workeneh BT, Jhaveri KD, Rondon-Berrios H. Hyponatremia in the cancer patient. *Kidney Int* 2020;**98**:870–82. https://doi.org/10.1016/j.kint.2020.05.015
5. Lassen U, Osterlind K, Hansen M *et al*. Long-term survival in small-cell lung cancer: posttreatment characteristics in patients surviving 5 to 18+ years—an analysis of 1,714 consecutive patients. *J Clin Oncol* 1995;**13**:1215–20. https://doi.org/10.1200/JCO.1995.13.5.1215
6. Talmi YP, Wolf GT, Hoffman HT *et al*. Elevated arginine vasopressin levels in squamous cell cancer of the head and neck. *Laryngoscope* 1996;**106**:317–21. https://doi.org/10.1097/00005537-199603000-00014
7. O'Neil BH, Scott AJ, Ma WW *et al*. A phase II/III randomized study to compare the efficacy and safety of rigosertib plus gemcitabine versus gemcitabine alone in patients with previously untreated metastatic pancreatic cancer. *Ann Oncol* 2015;**26**:1923–9. https://doi.org/10.1093/annonc/mdv264
8. Robertson GL, Bhoopalam N, Zelkowitz LJ. Vincristine neurotoxicity and abnormal secretion of antidiuretic hormone. *Arch Intern Med* 1973;**132**:717–20. https://doi.org/10.1001/archinte.1973.03650110061013
9. Gilbar PJ, Richmond J, Wood J *et al*. Syndrome of inappropriate antidiuretic hormone secretion induced by a single dose of oral cyclophosphamide. *Ann Pharmacother* 2012;**46**:e23. https://doi.org/10.1345/aph.1R296
10. Tan AC, Marx GM. Cisplatin-induced syndrome of inappropriate antidiuretic hormone secretion (SIADH) with life-threatening hyponatremia. *BMJ Case Rep* 2018;**2018**:bc2017222948. https://doi.org/10.1136/bcr-2017-222948
11. Littelwood TJ, Smith AP. Syndrome of inappropriate antidiuretic hormone secretion due to treatment of lung cancer with cisplatin. *Thorax* 1984;**39**:636–7. https://doi.org/10.1136/thx.39.8.636
12. Oronssky B, Caroen S, Oronsky A *et al*. Electrolyte disorders with platinum-based chemotherapy: mechanisms, manifestations, and management. *Cancer Chemother Pharmacol* 2017;**80**:895–907. https://doi.org/10.1007/s00280-017-3392-8
13. Seethapathy H, Rusibamayila N, Chute DF *et al*. Hyponatremia and other electrolyte abnormalities in patients receiving immune checkpoint inhibitors. *Nephrol Dial Transplant* 2021;**36**:2241–7. https://doi.org/10.1093/ndt/gfaa272
14. Wanchoo R, Sakhiy V, Jhaveri KD. Immune checkpoint inhibitor-associated electrolyte disorders: query of the Food and Drug Administration Adverse Event Reporting System. *Kidney Int* 2021;**100**:945–7. https://doi.org/10.1016/j.kint.2021.06.001
15. Matsuo S, Imai E, Horio M *et al*. Collaborators developing the Japanese equation for estimated GFR. *Am J Kidney Dis* 2009;**53**:982–92. https://doi.org/10.1053/j.ajkd.2008.12.034
16. Spasovski G, Vanholder R, Allolio B *et al*. Clinical practice guideline on diagnosis and treatment of hyponatremia. *Nephrol Dial Transplant* 2014;**29**:i1–39. https://doi.org/10.1093/ndt/gfu040
17. National Cancer Institute. *Common Terminology Criteria for Adverse Events (CTCAE) version* 5.0. Washington, DC: US Department of Health and Human Services; 2017. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf
18. Katz MA. Hyperglycemia-induced hyponatremia-calculation of expected serum sodium depression. *N Engl J Med* 1973;**289**:843–4. https://doi.org/10.1056/NEJM197310182891607
19. Ke G, Meng Q, Finley T *et al*. LightGBM: a highly efficient gradient boosting decision tree. *Adv Neur Inform Process Syst* 2018;**30**:3148–56.
20. Likas A, Vlassis N, Verbeek J. The global k-means clustering algorithm. *Pattern Recognit* 2003;**36**:451–61. https://doi.org/10.1016/S0031-3203(02)00060-2
21. Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. *Adv Neur Inform Process Syst* 2012;**25**:2960–8.
22. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neur Inform Process Syst* 2017;**30**:4768–77.
23. Domingos P. A few useful things to know about machine learning. *Commun ACM* 2012;**55**:78–87. https://doi.org/10.1145/2347736.2347755
24. Berghmams T, Paesmans M, Body JJ. A prospective study on hyponatremia in medical cancer patients: epidemiology, aetiology, and differential diagnosis. *Support Care Cancer* 2000;**8**:192–7. https://doi.org/10.1007/s005200050284
25. Amin MN, Salami BA, Zahid M *et al*. Investigating the bond strength of FRP laminates with concrete using LIGHT GBM and SHAPASH analysis. *Polymers* 2022;**14**:4717. https://doi.org/10.3390/polym14214717
26. Adrogué HJ, Tucker BM, Madias NE. Diagnosis and management of hyponatremia: a review. *JAMA* 2022;**328**:280–91. https://doi.org/10.1001/jama.2022.11176
27. Shysh AC, Ismail Z, Sidhu D *et al*. Factors associated with hyponatremia in patients newly prescribed citalopram: a retrospective observational study. *Drugs Real World Outcomes* 2021;**8**:555–63. https://doi.org/10.1007/s40801-021-00257-4
28. Ravioli S, Bahmad S, Funk GC *et al*. Risk of electrolyte disorders, syncope, and falls in patients taking thiazide diuretics: results of a cross-sectional study. *Am J Med* 2021;**134**:1148–54. https://doi.org/10.1016/j.amjmed.2021.04.007
29. Falhammar H, Skov J, Calissendorff J *et al*. Associations between antihypertensive medications and severe hyponatremia: a Swedish population-based case-control study. *J Clin Endocrinol Metab* 2020;**105**:e3696–705. https://doi.org/10.1210/clinem/dgaa194
30. Tannen RL, Regal EM, Dunn MJ *et al*. Vasopressin-resistant hyposthenuria in advanced chronic renal disease. *N Engl J Med* 1969;**280**:1135–41. https://doi.org/10.1056/NEJM196905222802101
31. Barsony J, Sugimura Y, Verbalis JG. Osteoclast response to low extracellular sodium and the mechanism of hyponatremia-induced bone loss. *J Biol Chem* 2011;**286**:10864–75. https://doi.org/10.1074/jbc.M110.155002
32. Tamma R, Cuscito LS, Corcelli PL *et al*. Regulation of bone remodeling by vasopressin explains the bone loss in hyponatremia. *Proc Natl Acad Sci USA* 2013;**110**:18644–9. https://doi.org/10.1073/pnas.1318257110
33. Berardi R, Santoni M, Rinaldi S *et al*. Risk of hyponatremia in cancer patients treated with targeted therapies: a systematic review and meta-analysis of clinical trials. *PLoS*

*One* 2016;**11**:e0152079. https://doi.org/10.1371/journal.pone.0152079

34. O'Connor-Byrne N, Glavey S, Tudor R *et al*. Bortezomib-induced hyponatremia: tolvaptan therapy permits continuation of lenalidomide, bortezomib and dexamethasone therapy in relapsed myeloma. *Exp Hematol Oncol* 2019;**8**:4. https://doi.org/10.1186/s40164-019-0128-y
35. Shimazu Y, Fujimura K. Hyponatremia in a multiple myeloma patient treated with bortezomib. *Hematol Oncol* 2017;**35**:929–30. https://doi.org/10.1002/hon.2381
36. Brodmann S, Klaas EG, Cathomas R *et al*. Severe hyponatremia in a patient with mantle cell lymphoma treated with bortezomib: a case report and review of the literature. *Onkologie* 2007;**30**:651–4.

*Received: 24.2.2023; Editorial decision: 11.7.2023*