

RESEARCH ARTICLE

Validation of the usefulness of artificial neural networks for risk prediction of adverse drug reactions used for individual patients in clinical practice

Shungo Imai^{1*}, Yoh Takekuma², Hitoshi Kashiwagi¹, Takayuki Miyai³, Masaki Kobayashi¹, Ken Iseki¹, Mitsuru Sugawara^{1,2}

1 Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo, Japan, **2** Department of Pharmacy, Hokkaido University Hospital, Sapporo, Japan, **3** Graduate School of Life Science, Hokkaido University, Sapporo, Japan

* s.imai@pharm.hokudai.ac.jp



OPEN ACCESS

Citation: Imai S, Takekuma Y, Kashiwagi H, Miyai T, Kobayashi M, Iseki K, et al. (2020) Validation of the usefulness of artificial neural networks for risk prediction of adverse drug reactions used for individual patients in clinical practice. PLoS ONE 15(7): e0236789. <https://doi.org/10.1371/journal.pone.0236789>

Editor: Le Hoang Son, Vietnam National University, VIET NAM

Received: November 9, 2019

Accepted: July 14, 2020

Published: July 29, 2020

Copyright: © 2020 Imai 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: All relevant data are within the paper.

Funding: S.I received Naomi Hoshino Memorial Grant for Pharmaceutical Initiatives, 2019. URL: <http://www.kp-dousoukai.com/business.html> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Abstract

Artificial neural networks are the main tools for data mining and were inspired by the human brain and nervous system. Studies have demonstrated their usefulness in medicine. However, no studies have used artificial neural networks for the prediction of adverse drug reactions. We aimed to validate the usefulness of artificial neural networks for the prediction of adverse drug reactions and focused on vancomycin-induced nephrotoxicity. For constructing an artificial neural network, a multilayer perceptron algorithm was employed. A 10-fold cross validation method was adopted for evaluating the resultant artificial neural network. In total, 1141 patients who received vancomycin at Hokkaido University Hospital from November 2011 to February 2019 were enrolled. Among these patients, 179 (15.7%) developed vancomycin-induced nephrotoxicity. The top three risk factors of vancomycin-induced nephrotoxicity which are relatively important in the artificial neural networks were average vancomycin trough concentration ≥ 13.0 mg/L and concomitant use of piperacillin-tazobactam and vasopressor drugs. The predictive accuracy of the artificial neural network was 86.3% and that of the multiple logistic regression model (conventional statistical method) was 85.1%. Moreover, area under the receiver operating characteristic curve (AUROC) of the artificial neural network was 0.83. In the 10-fold cross-validation, the accuracy obtained was 86.0% and AUROC was 0.82. The artificial neural network model predicting the vancomycin-induced nephrotoxicity showed good predictive performance. This appears to be the first report of the usefulness of artificial neural networks for an adverse drug reactions risk prediction model.

Introduction

The process of data mining is defined as the use of techniques to identify hidden correlations and patterns from complex datasets. In addition, it has been described as a method for

constructing predictive models based on the discovery of underlying patterns and relationships in large datasets [1].

Artificial neural networks (ANNs) are among the main tools used for data mining. They have a complex computational structure that is inspired by the human brain and nervous system [2]. The structure consists of input and output layers and a hidden layer of units that transform the inputs into something that the output layer can use [3]. ANNs are exceptional tools used for identifying the patterns from complex or numerous datasets to extract and teach the machine to recognise relationships [4–6]. Thus, ANNs are able to incorporate the intricate associations among variables into algorithms. In medical fields, recent studies concerning ANNs have constructed a variety of prediction models: survival prediction of gastric cancer [4], length of stay in an intensive care unit (ICU) [5] and risk of congenital heart disease in pregnant women [6]. Recently, several studies have applied ANNs to investigate adverse drug reactions (ADRs) [7–10]. However, these studies employed ANNs in areas of pharmacovigilance and drug discovery to find a causal relationship between a drug and adverse events [7–10]. Thus, a risk prediction model of ADRs using ANNs that is intended to be used for ‘individual patients in clinical practice’ has not yet been established. Such an ANN would be very useful, so it is important to validate its usefulness when applied to risk prediction models for clinical practice.

In this study, we selected vancomycin (VCM)-induced nephrotoxicity (VIN) for validating the usefulness of ANNs. There are many reports on risk factors for VIN, such as higher concentration (e.g. trough concentrations > 15 or 20 mg/L) [11–13], long-term duration of therapy [14,15], certain hosts (i.e. those with baseline renal impairment and a history of acute kidney injury and those who are critically ill or have septic shock) [16–18] and concomitant medications [i.e. nonsteroidal anti-inflammatory drugs (NSAIDs), furosemide, amphotericin B, aminoglycoside antibiotics and piperacillin–tazobactam (PIPC–TAZ)] [11,19,20]. Thus, risk factors also have been established for the construction of ANNs. In Hokkaido University Hospital, the number of cases of intravenous VCM administration is about 200 patients per year, and this has been estimated to be sufficient for the construction of ANNs [4–6]. Considering the above, VIN was thought to be suitable for verifying the usefulness of an ANN model for the risk prediction of ADR. These risk factors have also been analysed by multiple logistic regression [11–18]. Thus, this conventional statistical method is suitable to validate the ANNs.

Although there are several algorithms for constructing ANNs, we employed a multilayer perceptron (MLP) in this study. MLP is one of the typical supervised learning algorithms in which a small number of parameters can be used to predict outcomes [21,22]. In addition, MLP can be performed by packaging software, such as SPSS (IBM, Tokyo, Japan) and JMP (SAS Institute, Inc., Cary, NC, USA) [4–6,23,24]. Since it does not require complex programming, the methodology established in this research is expected to be easily adaptable to other ADRs by clinicians and pharmacists. Thus, MLP is not new but our approach of applying it to ‘risk prediction of ADR’ is novel.

Therefore, in the present study, our objective was to validate the usefulness of ANNs using MLP algorithm as applied to risk prediction ADRs by constructing a risk prediction model for VIN.

Materials and methods

Ethics

This retrospective observational study was conducted in accordance with the guidelines for human studies. The study protocol was approved by the ethics committee of Hokkaido

University Hospital (study protocol NO. 018–0379). Because this study is conducted retrospectively, they approved this study and waived informed consent.

Patients

This single-centre retrospective observational study was conducted at Hokkaido University Hospital. Subjects who had received VCM intravenously from November 2011 to February 2019 were recruited. All data were obtained from the patients' electronic medical records. The inclusion criteria were (1) age ≥ 18 years, (2) measured VCM trough concentration after the third day of administration and (3) dosing period of ≥ 3 days. We excluded patients who had undergone haemodialysis and continuous haemodialysis flow or had nephrotoxicity prior to the measurement of VCM trough concentration. Informed consent was obtained from all patients in the form of opt-out on the web-site in Hokkaido University Hospital.

Criteria of VCM-induced nephrotoxicity

The 2009 vancomycin consensus statement of the Infectious Diseases Society of America [25] has defined nephrotoxicity as a serum creatinine (Scr) increase of ≥ 0.5 mg/dL or $\geq 50\%$ relative to baseline [25]. To evaluate VIN, we extracted the maximum Scr during the administration period.

Data collection

Risk factors for nephrotoxicity were extracted on the basis of previous reports [11–20] and the following potential factors: patient age, sex (male/female), body weight, Scr, creatinine clearance (CCr), duration of therapy, concomitant medications (NSAIDs, furosemide, amphotericin B, aminoglycosides, PIPC–TAZ and vasopressor drugs), residence in the ICU, with or without loading dose and average VCM trough concentration during therapy. Among the concomitant medications, vasopressor drugs were defined as follows: etilefrine, noradrenaline, olprinone, milrinone, dopamine and dobutamine [26]. The loading dose was defined as an initial dose (single or daily) ≥ 1.25 times of the maintenance dosage [26]. Moreover, to evaluate patient characteristics, we collected the days to initial therapeutic drug monitoring (TDM) and initial and maximum VCM trough concentration during therapy. All data were extracted from the beginning of VCM administration, except for the duration of therapy, concomitant medications, residence in ICU, days to initial TDM and VCM trough concentration. Data of concomitant medications and residence in ICU were evaluated during the administration period. To calculate CCr, the Cockcroft–Gault equation was employed [27].

Construction of the ANN and statistical analysis

As described above, MLP was employed for the construction of ANN. The MLP consists of an input layer of nodes containing information, such as risk factors, followed by a hidden layer of nodes that interact with the input variables that are finally transferred to the output layer [21,28]. In the input layer, the number of neurons depends on the number of independent variables, whereas the number of neurons in the output layer correlates with the number of values that need to be predicted [21,28]. The steps of MLP are summarised as follows [21,28]: (1) data is provided to input layer; (2) input layer produces a predicted output layer, which is subtracted from actual output, and error value is estimated; (3) a back propagation adjusts the weights between output and hidden layer nodes, which works backwards through network; (4) when a back propagation is finished, the process starts again; and (5) this process is repeated until error is minimised.

The analysis was performed in three steps according to previous reports [28]. Firstly, univariate logistic regression analysis was performed to identify the potential risk factors of VIN. All continuous variables were converted into categorical variables. The optimal cut-off points were determined from the receiver operating characteristic (ROC) curves using Youden's index [29]. Secondly, the ANN and multivariate logistic regression models were constructed. In this analysis, all of the potential risk factors with P -values ≤ 0.05 in the univariate analysis were used. Finally, the predictive performances of the ANN model and multivariate logistic regression model were compared. To evaluate predictive performances, the accuracy was calculated for each model, and the areas under the ROC curve (AUROC) of the ANN model was evaluated. These indexes were generally considered to be important performance scores in previous studies [28,30–34]. Furthermore, the 10-fold cross validation was performed for internal validation of the ANN model [24,35]. The Hosmer–Lemeshow test was used to evaluate the fitness of the logistic regression model (the cut-off value was $P \geq 0.05$) [36].

Patient characteristics were compared using unpaired, and all tests of significance were two-tailed. For comparing the continuous variables, the Mann–Whitney U -test was used (all continuous variables were non-normally distributed). Categorical variables were compared using Pearson's Chi-squared test or Fisher's exact test. $P \leq 0.05$ was considered to be statistically significant.

All statistical analyses were performed using JMP 14 (SAS Institute Inc., Cary, NC, USA), a statistical software typically used for ANNs [23,24].

Results

Patient characteristics

Out of 1490 initial patients, 1141 were included in the study (Fig 1). Among them, 179 (15.7%) developed VIN. As shown in Table 1, there were significant differences between the patients who developed nephrotoxicity and those who did not in Scr; CCr; duration of therapy; concomitant medications (furosemide, amphotericin B, PIPC–TAZ and vasopressor drugs);

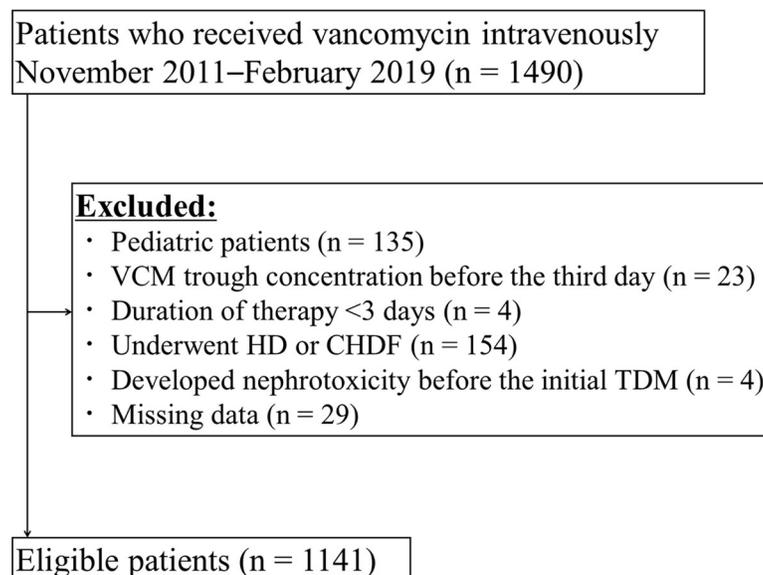


Fig 1. Flowchart of patients included in this study. Vancomycin: VCM, Therapeutic drug monitoring: TDM, Haemodialysis: HD, Continuous haemodialysis flow: CHDF.

<https://doi.org/10.1371/journal.pone.0236789.g001>

Table 1. Comparison of the characteristics of patients with and without nephrotoxicity.

Characteristic	Total (n = 1141)	With nephrotoxicity (n = 179)	Without nephrotoxicity (n = 962)	P-value
Age (years), median (range)	65 (18–96)	65 (18–96)	66 (18–95)	$P = 0.21^c$
Age ≥ 67 years, n (%)	504 (44.2)	72 (40.2)	432 (44.9)	$P = 0.25^a$
Sex (male), n (%)	728 (63.8)	111 (62.0)	617 (64.1)	$P = 0.59^a$
Body weight (kg), median (range)	57.0 (28.3–127.0)	57.9 (28.3–98.1)	56.7 (29.1–127.0)	$P = 0.37^c$
Body weight ≥ 57.2 kg, n (%)	566 (49.6)	99 (55.3)	467 (48.5)	$P = 0.10^a$
Serum creatinine (mg/dL), median (range)	0.67 (0.16–5.15)	0.62 (0.24–4.57)	0.68 (0.16–5.15)	$P < 0.01^c$ *
Serum creatinine ≥ 0.68 mg/dL, n (%)	564 (49.4)	71 (39.7)	493 (51.3)	$P < 0.01^a$ *
CCr (mL/min), median (range)	85.9 (7.3–569.6)	96.3 (7.3–315.2)	84.0 (10.0–569.6)	$P < 0.01^c$ *
CCr < 88.8 mL/min, n (%)	607 (53.2)	75 (41.9)	532 (55.3)	$P < 0.01^a$ *
Duration of therapy (days), median (range)	9 (3–88)	12 (3–88)	8 (3–83)	$P < 0.01^c$ *
Duration of therapy ≥ 10 days, n (%)	533 (46.7)	114 (63.7)	419 (43.6)	$P < 0.01^a$ *
Concomitant medications, n (%)				
NSAIDs	541 (47.4)	92 (51.4)	449 (46.7)	$P = 0.24^a$
Furosemide	392 (34.4)	108 (60.3)	284 (29.5)	$P < 0.01^a$ *
Piperacillin–Tazobactam	188 (16.5)	57 (31.8)	131 (13.6)	$P < 0.01^a$ *
Amphotericin B	21 (1.84)	11 (6.15)	10 (1.04)	$P < 0.01^b$ *
Aminoglycoside antibiotics	26 (2.28)	7 (3.91)	19 (1.98)	$P = 0.17^b$
Vasopressor drugs	149 (13.1)	48 (26.8)	101 (10.5)	$P < 0.01^a$ *
Residence in intensive care unit, n (%)	145 (12.7)	33 (18.4)	112 (11.6)	$P = 0.01^a$ *
Duration of initial TDM (days), median (range)	3 (3–10)	3 (3–7)	3 (3–10)	$P = 0.63^c$
Initial VCM trough concentration (mg/L), median (range)	10.6 (2.1–39.4)	12.8 (3.8–39.4)	10.4 (2.1–36.0)	$P < 0.01^c$ *
Maximum VCM trough concentration (mg/L), median (range)	13.5 (2.1–72.2)	21.5 (5.7–72.2)	12.6 (2.1–36.0)	$P < 0.01^c$ *
Average VCM trough concentration (mg/L), median (range)	11.6 (2.1–42.1)	15.1 (4.7–42.1)	11.2 (2.1–29.5)	$P < 0.01^c$ *
Average VCM trough concentration ≥ 13 mg/L, n (%)	449 (39.4)	114 (63.7)	335 (34.8)	$P < 0.01^a$ *
With loading dose, n (%)	187 (16.4)	23 (12.8)	164 (17.0)	$P = 0.16^a$

Creatinine clearance: CCr, Vancomycin: VCM, Nonsteroidal anti-inflammatory drugs: NSAIDs, Therapeutic drug monitoring: TDM

a)Chi-squared test

b)Fisher's exact test

c)Mann–Whitney *U*-test.

**P*-values ≤ 0.05 were considered statistically significant.

<https://doi.org/10.1371/journal.pone.0236789.t001>

residence in the ICU; and the initial, maximum and average VCM trough concentrations during therapy.

Univariate analysis

In the univariate analysis (Table 2), Scr ≥ 0.68 mg/dL, CCr < 88.8 mL/min, duration of therapy ≥ 10 days, concomitant medications furosemide, amphotericin B, PIPC–TAZ and vasopressor drugs, residence in the ICU and average VCM trough concentration ≥ 13.0 mg/L were significant factors ($P \leq 0.05$). However, Scr is usually strongly associated with CCr. In this study, Scr was also excluded. Thus, these factors, excluding Scr, were used to construct the ANN and multiple logistic regression models.

Construction of the ANN model

The ANN model predicting the VIN is shown in Fig 2. Based on the univariate analysis, the eight independent variables were applied, and the dependent variable was the presence or

Table 2. Univariate analysis of risk factors for nephrotoxicity.

Characteristic	OR	95% CI	P-value
Age \geq 67 years	0.83	0.60–1.14	$P = 0.25$
Sex (male)	0.91	0.66–1.27	$P = 0.57$
Body weight \geq 57.2 kg	1.31	0.95–1.81	$P = 0.10$
Serum creatinine \geq 0.68 mg/dL	0.63	0.45–0.87	$P < 0.01^\dagger$
CCr $<$ 88.8 mL/min	0.58	0.42–0.81	$P < 0.01^\dagger$
Duration of therapy \geq 10 days	2.27	1.63–3.16	$P < 0.01^\dagger$
Concomitant medications			
NSAIDs	1.24	0.90–1.70	$P = 0.19$
Furosemide	3.63	2.61–5.05	$P < 0.01^\dagger$
Amphotericin B	6.23	2.61–14.91	$P < 0.01^\dagger$
Aminoglycoside antibiotics	1.95	0.69–5.47	$P = 0.21$
Piperacillin–Tazobactam	2.96	2.06–4.27	$P < 0.01^\dagger$
Vasopressor drugs	3.12	2.12–4.61	$P < 0.01^\dagger$
Residence in intensive care unit	1.72	1.12–2.63	$P = 0.01^\dagger$
Average VCM trough concentration \geq 13 mg/L	3.28	2.35–4.58	$P < 0.01^\dagger$
With loading dose	0.72	0.45–1.15	$P = 0.17$

Creatinine clearance: CCr, Vancomycin: VCM, Odds ratio: OR, 95% Confidence interval: 95% CI

$^\dagger P$ -values ≤ 0.05 were included in the artificial neural network and multiple logistic regression analysis.

<https://doi.org/10.1371/journal.pone.0236789.t002>

absence of nephrotoxicity. The ANN model consists of an input layer, a hidden layer and an output layer. The input and output layers contained eight and one neuron, respectively. The relative importance of the independent variables in the ANN model is presented in Fig 3. The top three factors for VIN were average VCM trough concentration ≥ 13.0 mg/L, concomitant use of PIPC–TAZ and vasopressor drugs.

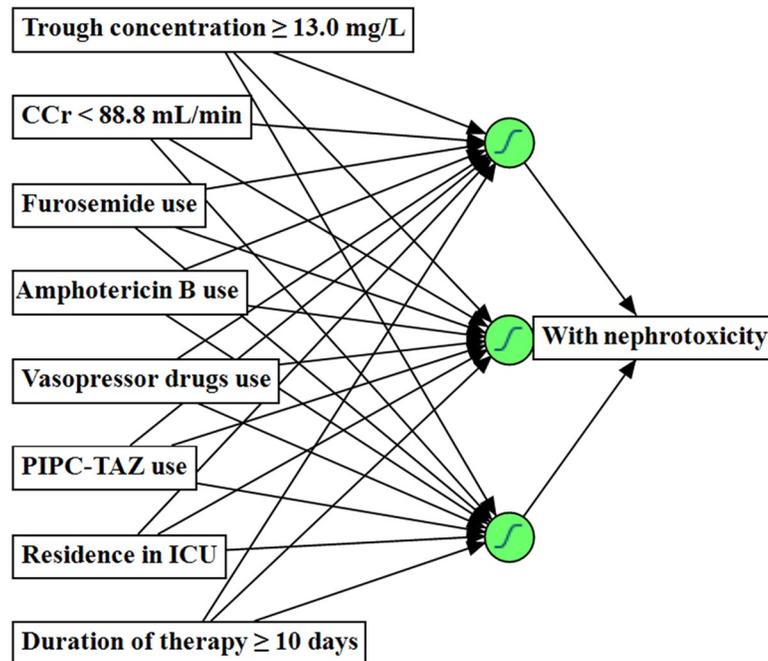


Fig 2. ANN model predicting VCM-induced nephrotoxicity. Creatinine clearance: CCr, Average vancomycin trough concentration: Trough concentration, Intensive care unit: ICU, Piperacillin–Tazobactam: PIPC–TAZ.

<https://doi.org/10.1371/journal.pone.0236789.g002>

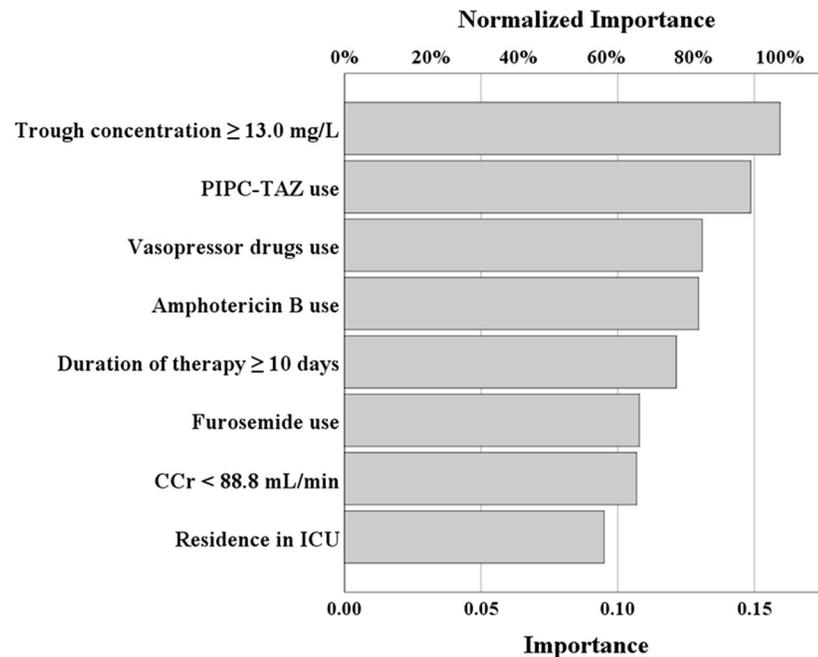


Fig 3. Relative importance of the independent variables in the ANN model. Average vancomycin trough concentration: Trough concentration, Creatinine clearance: CCr, Intensive care unit: ICU, Piperacillin–Tazobactam: PIPC–TAZ.

<https://doi.org/10.1371/journal.pone.0236789.g003>

Multiple logistic regression analysis

As shown in Table 3, in the multiple logistic regression analysis using a stepwise approach, CCr < 88.8 mL/min, duration of therapy ≥ 10 days, concomitant medications (furosemide, amphotericin B, PIPC–TAZ and vasopressor drugs) and average VCM trough concentration ≥ 13.0 mg/L were extracted as the independent risk factors of VIN.

Validation of the ANN and multiple logistic regression models

The predictive accuracy of the ANN model was 86.3% and that of the multiple logistic regression model (conventional statistical method) was 85.1%. In addition, AUROC of the ANN model was 0.83 (Fig 4). In the 10-fold cross-validation, accuracy and AUROC were 86.0% and

Table 3. Multivariate analysis of risk factors for nephrotoxicity.

Characteristic	OR	95% CI	P-value
CCr < 88.8 mL/min	0.41	0.29–0.60	$P < 0.01^*$
Duration of therapy ≥ 10 days	2.32	1.60–3.36	$P < 0.01^*$
Concomitant medications			
Furosemide	2.54	1.75–3.68	$P < 0.01^*$
Amphotericin B	3.43	1.29–9.11	$P = 0.01^*$
Piperacillin–Tazobactam	3.36	2.22–5.06	$P < 0.01^*$
Vasopressor drugs	2.78	1.75–4.41	$P < 0.01^*$
Average VCM trough concentration ≥ 13 mg/L	3.60	2.49–5.20	$P < 0.01^*$

Creatinine clearance: CCr, Vancomycin: VCM, Odds ratio: OR, 95% Confidence interval: 95% CI

*P-values ≤ 0.05 were considered statistically significant.

<https://doi.org/10.1371/journal.pone.0236789.t003>

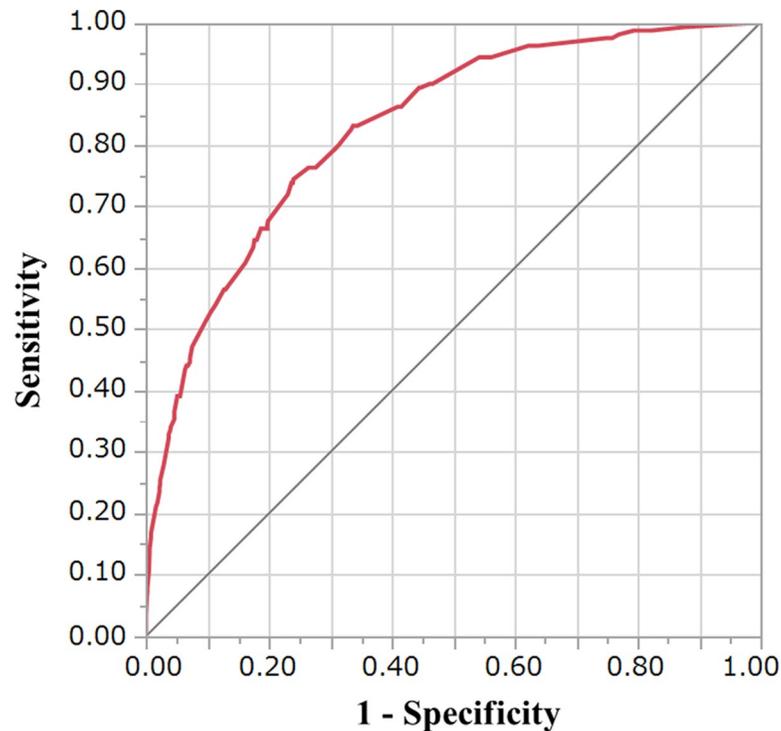


Fig 4. Receiver operating characteristic curve of the ANN model. The area under the receiver operating characteristic curve was 0.83.

<https://doi.org/10.1371/journal.pone.0236789.g004>

0.82, respectively. In the multiple logistic regression model, the Hosmer–Lemeshow test gave a *P*-value of 0.66.

Discussion

To the best of our knowledge, this is the first study to validate the usefulness of ANNs applied to a risk prediction model of ADRs for individual patients in clinical practice by constructing a risk prediction model of VIN. In the ANN model, the predictive accuracy was 86.3% and the AUROC was 0.83. These indexes were also used in some previous reports that regarded them as important performance scores [28,30–34]. The AUROC of the ANN model (0.83) indicated moderate accuracy based on the criteria reported by Akobeng [29]. Furthermore, when compared with the results of previous reports, our results are favourable. For example, Pergialiotis et al. built an ANN model to predict endometrial cancer in postmenopausal women and achieved an accuracy of 85.4% [2]. Paydar et al. developed a prediction model of pregnancy outcomes among pregnant women with systemic lupus erythematosus and achieved an accuracy of 90.9% [35]. Hassanipour et al. conducted a systematic review of ten studies that used ANNs to predict health-related outcomes in traumatic patients [30]. They compared AUROC and accuracy between these ten studies, and the AUROC ranged from 0.73 to 0.97, with accuracies from 80.9% to 98.4%. Considering these values, our predictive performances were reasonably accurate. In addition, the accuracy and AUROC in the 10-fold cross-validation was 86.0% and 0.82, respectively, which were favourable [24,35].

In this study, the accuracy of the ANN model (86.3%) was slightly higher than that of the multiple logistic regression model (85.1%). Comparison of the predictive performances of ANNs and logistic regression models has been reported by several previous studies. In the

above-mentioned systematic review [30], ANNs had a high level of accuracy and was statistically significant (odds ratio: 1.09). Further, similar results have been obtained in other previous reports [2,31,37,38]. Thus, clinical application of ANNs may enable more accurate prediction of ADRs than logistic regression model. In addition, this approach can be applied to other ADRs and developed further. Meanwhile, logistic regression model is appropriate if the primary endpoint is extracting dependent factors affecting ADRs because ANNs cannot analyse individual factors (e.g., calculating odds ratio) [38].

As shown in Fig 3, an average VCM trough concentration ≥ 13.0 mg/L was extracted as the most important factor of VIN in the ANN, which was consistent with the multiple logistic regression analysis (Table 3). A high VCM trough concentration is known to be a common risk factor of VIN, and cut-off values are usually >15 or 20 mg/L [11–13,16]. On the other hand, our result of ≥ 13.0 mg/L was lower than these (cut-off points were determined from the ROC using the Youden's index [29]), which was assumed to be caused by differences in the target trough concentrations. In previous reports, the target trough concentrations were also set to 15 – 20 mg/L [16,19,25]. In our hospital, target trough levels were set to 10 – 20 mg/L based on the TDM practice guidelines in Japan [39]. Thus, these target trough levels were lower than those of 15 – 20 mg/L in previous reports [16,19,25], which may be the reason of the lower cut-off value of VIN. PIPC–TAZ use was extracted as the second most important risk factor in ANN model. Recently, concomitant use of PIPC–TAZ has received attention for its association with VIN [20,40,41]. Although this mechanism remains unclear, VIN is obviously increased by PIPC–TAZ use, and our results supported those of the previous reports. Generally, baseline renal impairment, like that in patients with chronic kidney disease, is associated with VIN [16]. However, our result was inconsistent with this ($\text{CCr} < 88.8$ mL/min, odds ratio = 0.41, 95% confidence interval, 0.29–0.60, Table 3). This is thought to have been caused by the 'actual Scr use' in the CCr calculations. Smythe et al. evaluated the accuracy of CCr estimates generated for elderly patients and recommended rounding the Scr to 1.0 mg/dL for low Scr values [42]. In addition, rounding the Scr to 0.6 mg/dL was recommended by Winter [43]. Thus, if an adjusted Scr was employed, this result would not have been obtained. However, an adjustment method for Scr has not become well established, so we used the actual values in the present study. Therefore, investigation of the accuracy of CCr calculations should be investigated in future research.

Accordingly, we also used ANNs to successfully build a risk prediction model of VIN. However, compared with logistic regression analysis, ANNs have several disadvantages. Firstly, ANNs have a 'black box' nature; that is, ANNs cannot explain any insights into the structure of the function being approximated [44]. This is in contrast with a logistic regression model that can provide such information. Secondly, ANNs have a risk of overtraining and a possibility of overfitting the model, which may provide an overconfident prediction [45]. Finally, for clinical applications, ANNs require special statistical analysis software. Thus, it would currently be difficult to use our models widely. However, Pergialiotis V et al. explained that these problems can be solved using a larger number of patients (except for requiring the special statistical analysis software) because although a small dataset may not be applicable to large cohorts, the reverse is always possible [2]. Thus, establishment of larger databases, such as one in a multi-centre study, is necessary for the construction of safer ANN models.

Our study had several limitations. First, this study was conducted at a single centre. Second, factors that have been reported previously, such as septic shock, history of acute kidney injury and acute physiology and chronic health evaluation II scores, could not be evaluated [16–18]. In addition, risk factors of concomitant medications and residence in ICU were extracted during the administration period, and trough concentrations were evaluated using average values. Thus, our models included factors that could not be evaluated at the time of use. However, this

study aimed to validate ANNs for the prediction of ADRs, so we thought that our study design was the best.

In this study, the ANN model predicting VIN exhibited good predictive performance. Thus, our results indicate the usefulness of ANNs as risk prediction models of ADRs for individual patients in clinical practice. These models would enable clinician and pharmacists to predict ADRs and to easily make decisions such as drug selections. Furthermore, some advanced ANN algorithms, such as recurrent neural network [7,8], can also be employed for this purpose in future. Thus, by performing multi-centre study and using advanced ANN algorithms, reliable risk prediction models need to be built.

Author Contributions

Conceptualization: Shungo Imai.

Data curation: Shungo Imai, Takayuki Miyai.

Formal analysis: Shungo Imai.

Funding acquisition: Shungo Imai, Masaki Kobayashi, Ken Iseki.

Investigation: Shungo Imai, Yoh Takekuma, Takayuki Miyai.

Methodology: Shungo Imai, Yoh Takekuma, Hitoshi Kashiwagi, Takayuki Miyai.

Project administration: Shungo Imai.

Resources: Shungo Imai, Ken Iseki.

Software: Shungo Imai, Masaki Kobayashi, Ken Iseki.

Supervision: Mitsuru Sugawara.

Visualization: Shungo Imai.

Writing – original draft: Shungo Imai.

Writing – review & editing: Yoh Takekuma, Hitoshi Kashiwagi, Takayuki Miyai, Masaki Kobayashi, Ken Iseki, Mitsuru Sugawara.

References

1. Hiramatsu N, Kurosaki M, Sakamoto N, Iwasaki M, Sakamoto M, Suzuki Y, et al. Pretreatment prediction of anemia progression by pegylated interferon alpha-2b plus ribavirin combination therapy in chronic hepatitis C infection: decision-tree analysis. *J Gastroenterol*. 2011; 46: 1111–1119. <https://doi.org/10.1007/s00535-011-0412-z> PMID: 21681410
2. Pergialiotis V, Pouliakis A, Parthenis C, Damaskou V, Chrelias C, Papantoniou N, et al. The utility of artificial neural networks and classification and regression trees for the prediction of endometrial cancer in postmenopausal women. *Public Health* 2018; 164: 1–6. <https://doi.org/10.1016/j.puhe.2018.07.012> PMID: 30149185
3. Huang S, Xu Y, Yue L, Wei S, Liu L, Gan X, et al. Evaluating the risk of hypertension using an artificial neural network method in rural residents over the age of 35 years in a Chinese area. *Hypertens Res*. 2010; 33: 722–726. <https://doi.org/10.1038/hr.2010.73> PMID: 20505678
4. Yazdani Charati J, Janbabaie G, Alipour N, Mohammadi S, Ghorbani Gholiabad S, Fendereski A. Survival prediction of gastric cancer patients by Artificial Neural Network model. *Gastroenterol Hepatol Bench*. 2018; 11: 110–117. PMID: 29910851
5. LaFaro RJ, Pothula S, Kubal KP, Inchiosa ME, Pothula VM, Yuan SC, et al. Neural network prediction of ICU length of stay following cardiac surgery based on pre-incision variables. *PLoS One*. 2015; 10: e0145395. <https://doi.org/10.1371/journal.pone.0145395> PMID: 26710254
6. Li H, Luo M, Zheng J, Luo J, Zeng R, Feng N, et al. An artificial neural network prediction model of congenital heart disease based on risk factors: A hospital-based case-control study. *Medicine (Baltimore)*. 2017; 96: e6090.

7. Wunnava S, Qin X, Kakar T, Sen C, Rundensteiner EA, Kong X. Adverse drug event detection from electronic health records using hierarchical recurrent neural networks with dual-level embedding. *Drug Saf.* 2019; 42: 113–122. <https://doi.org/10.1007/s40264-018-0765-9> PMID: 30649736
8. Cocos A, Fiks AG, Masino AJ. Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in twitter posts. *J Am Med Inform Assoc.* 2017; 24: 813–821. <https://doi.org/10.1093/jamia/ocw180> PMID: 28339747
9. Dey S, Luo H, Fokoue A, Hu J, Zhang P. Predicting adverse drug reactions through interpretable deep learning framework. *BMC Bioinformatics.* 2018; 19: 476. <https://doi.org/10.1186/s12859-018-2544-0> PMID: 30591036
10. Zhang L, Zhang H, Ai H, Hu H, Li S, Zhao J, et al. Applications of machine learning methods in drug toxicity prediction. *Curr Top Med Chem.* 2018; 18: 987–997. <https://doi.org/10.2174/1568026618666180727152557> PMID: 30051792
11. Huang M, Wu H, Zhou J, Xu M, Zhou S. Efficacy of vancomycin on gram-positive bacterial infection in elderly critical patients and risk factors associated with nephrotoxicity. *Arch Iran Med.* 2018; 21: 349–355. PMID: 30113856
12. Cano EL, Haque NZ, Welch VL, Cely CM, Peyrani P, Scerpella EG, et al. Incidence of nephrotoxicity and association with vancomycin use in intensive care unit patients with pneumonia: retrospective analysis of the IMPACT-HAP Database. *Clin Ther.* 2012; 34: 149–157. <https://doi.org/10.1016/j.clinthera.2011.12.013> PMID: 22284995
13. Bosso JA, Nappi J, Rudisill C, Wellein M, Bookstaver PB, Swindler J, et al. Relationship between vancomycin trough concentrations and nephrotoxicity: a prospective multicenter trial. *Antimicrob Agents Chemother.* 2011; 55: 5475–5479. <https://doi.org/10.1128/AAC.00168-11> PMID: 21947388
14. Contreiras C, Legal M, Lau TT, Thalakada R, Shalansky S, Ensom MH. Identification of risk factors for nephrotoxicity in patients receiving extended-duration, high-trough vancomycin therapy. *Can J Hosp Pharm.* 2014; 67: 126–132. <https://doi.org/10.4212/cjhp.v67i2.1340> PMID: 24799722
15. Hall RG, Hazlewood KA, Brouse SD, Giuliano CA, Haase KK, Frei CR, et al. Empiric guideline-recommended weight-based vancomycin dosing and nephrotoxicity rates in patients with methicillin-resistant *Staphylococcus aureus* bacteremia: A retrospective cohort study. *BMC Pharmacol Toxicol.* 2013; 14: 12. <https://doi.org/10.1186/2050-6511-14-12> PMID: 23402420
16. Lodise TP, Lomaestro B, Graves J, Drusano GL. Larger vancomycin doses (at least four grams per day) are associated with an increased incidence of nephrotoxicity. *Antimicrob. Agents Chemother.* 2008; 52: 1330–1336. <https://doi.org/10.1128/AAC.01602-07> PMID: 18227177
17. Minejima E, Choi J, Beringer P, Lou M, Tse E, Wong-Beringer A. Applying new diagnostic criteria for acute kidney injury to facilitate early identification of nephrotoxicity in vancomycin-treated patients. *Antimicrob Agents Chemother.* 2011; 55: 3278–3283. <https://doi.org/10.1128/AAC.00173-11> PMID: 21576448
18. Cappelletty D, Jablonski A, Jung R. Risk factors for acute kidney injury in adult patients receiving vancomycin. *Clin Drug Investig.* 2014; 34: 189–193. <https://doi.org/10.1007/s40261-013-0163-0> PMID: 24385282
19. van Hal SJ, Paterson DL, Lodise TP. Systematic review and meta-analysis of vancomycin-induced nephrotoxicity associated with dosing schedules that maintain troughs between 15 and 20 milligrams per liter. *Antimicrob Agents Chemother.* 2013; 57: 734–744. <https://doi.org/10.1128/AAC.01568-12> PMID: 23165462
20. Hammond DA, Smith MN, Li C, Hayes SM, Lusardi K, Bookstaver P B. Systematic review and meta-analysis of acute kidney injury associated with concomitant vancomycin and piperacillin/tazobactam. *Clin Infect Dis.* 2017; 64: 666–674. <https://doi.org/10.1093/cid/ciw811> PMID: 27940946
21. Loftus TJ, Brakenridge SC, Croft CA, Smith RS, Efron PA, Moore FA, et al. Neural network prediction of severe lower intestinal bleeding and the need for surgical intervention. *J Surg Res.* 2017; 212: 42–47. <https://doi.org/10.1016/j.jss.2016.12.032> PMID: 28550920
22. Pasini A. Artificial neural networks for small dataset analysis. *J Thorac Dis.* 2015; 7: 953–960. <https://doi.org/10.3978/j.issn.2072-1439.2015.04.61> PMID: 26101654
23. Pralle RS, Weigel KW, White HM. Predicting blood β -hydroxybutyrate using milk Fourier transform infrared spectrum, milk composition, and producer-reported variables with multiple linear regression, partial least squares regression, and artificial neural network. *J Dairy Sci.* 2018; 101: 4378–4387. <https://doi.org/10.3168/jds.2017-14076> PMID: 29477523
24. Ing EB, Miller NR, Nguyen A, Su W, Bursztyn LLC, Poole M, et al. Neural network and logistic regression diagnostic prediction models for giant cell arteritis: development and validation. *Clin Ophthalmol.* 2019; 13: 421–430. <https://doi.org/10.2147/OPHTH.S193460> PMID: 30863010
25. Rybak MJ, Lomaestro BM, Rotschafer JC, Moellering RC, Craig WA, Billeter M, et al. Vancomycin therapeutic guidelines: A summary of consensus recommendations from the infectious diseases Society of

- America, the American Society of Health-System Pharmacists, and the Society of Infectious Diseases Pharmacists. *Clin Infect Dis*. 2009; 49: 325–327. <https://doi.org/10.1086/600877> PMID: 19569969
26. Imai S, Yamada T, Ishiguro N, Miyamoto T, Kagami K, Tomiyama N, et al. Validating the effectiveness of switching the vancomycin TDM analysis software based on the predictive accuracy. *Yakugaku Zasshi*. 2017; 137: 1185–1192. <https://doi.org/10.1248/yakushi.17-00080> PMID: 28867705
 27. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron*. 1976; 16: 31–41. <https://doi.org/10.1159/000180580> PMID: 1244564
 28. Raghupathi V, Raghupathi W. Preventive healthcare: A neural network analysis of behavioral habits and chronic diseases. *Healthcare (Basel)*. 2017; 5: E8.
 29. Akobeng A K. Understanding diagnostic tests 3: Receiver operating characteristic curves. *Acta Paediatr*. 2017; 96: 644–647.
 30. Hassanipour S, Ghaem H, Arab-Zozani M, Seif M, Fararouei M, Abdzadeh E, et al. Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: A systematic review and meta-analysis. *Injury*. 2019; 50: 244–250. <https://doi.org/10.1016/j.injury.2019.01.007> PMID: 30660332
 31. Lee HC, Yoon SB, Yang SM, Kim WH, Ryu HG, Jung CW, et al. Prediction of acute kidney injury after liver transplantation: Machine learning approaches vs. logistic regression model. *J Clin Med*. 2018; 7: E428. <https://doi.org/10.3390/jcm7110428> PMID: 30413107
 32. Hsieh MH, Hsieh MJ, Chen CM, Hsieh CC, Chao CM, Lai CC. Comparison of machine learning models for the prediction of mortality of patients with unplanned extubation in intensive care units. *Sci Rep*. 2018; 8: 17116. <https://doi.org/10.1038/s41598-018-35582-2> PMID: 30459331
 33. DeGregory KW, Kuiper P, DeSilvio T, Pleuss JD, Miller R, Roginski JW, et al. A review of machine learning in obesity. *Obes Rev*. 2018; 19: 668–685. <https://doi.org/10.1111/obr.12667> PMID: 29426065
 34. Duan H, Sun Z, Dong W, Huang Z. Utilizing dynamic treatment information for MACE prediction of acute coronary syndrome. *BMC Med Inform Decis Mak* 2019; 19: 5. <https://doi.org/10.1186/s12911-018-0730-7> PMID: 30626381
 35. Paydar K, Niakan Kalhori SR, Akbarian M, Sheikhtaheri A. A clinical decision support system for prediction of pregnancy outcome in pregnant women with systemic lupus erythematosus. *Int J Med Inform*. 2017; 97: 239–246. <https://doi.org/10.1016/j.ijmedinf.2016.10.018> PMID: 27919382
 36. Hosmer DW, Lemeshow S. *Applied logistic regression*, Second Edition (ed. Noel A. et al.). 32–46 (Wiley & Sons, Inc, 2005).
 37. Jang DH, Kim J, Jo YH, Lee JH, Hwang JE, Park SM, et al. Developing neural network models for early detection of cardiac arrest in emergency department. *Am J Emerg Med*. 2020; 38: 43–49. <https://doi.org/10.1016/j.ajem.2019.04.006> PMID: 30982559
 38. Lin CC, Ou YK, Chen SH, Liu YC, & Lin J. Comparison of artificial neural network and logistic regression models for predicting mortality in elderly patients with hip fracture. *Injury*. 2010; 41: 869–873. <https://doi.org/10.1016/j.injury.2010.04.023> PMID: 20494353
 39. Matsumoto K, Takesue Y, Ohmagari N, Mochizuki T, Mikamo H, Seki M, et al. Practice guidelines for therapeutic drug monitoring of vancomycin: a consensus review of the Japanese Society of Chemotherapy and the Japanese Society of Therapeutic Drug Monitoring. *J Infect Chemother*. 2013; 19: 365–380. <https://doi.org/10.1007/s10156-013-0599-4> PMID: 23673472
 40. Chen XY, Xu RX, Zhou X, Liu Y., Hu CY, Xie XF. Acute kidney injury associated with concomitant vancomycin and piperacillin/tazobactam administration: a systematic review and meta-analysis. *Int Urol Nephrol*. 2018; 50: 2019–2026. <https://doi.org/10.1007/s11255-018-1870-5> PMID: 29752626
 41. Burgess LD, Drew RH. Comparison of the incidence of vancomycin-induced nephrotoxicity in hospitalized patients with and without concomitant piperacillin-tazobactam. *Pharmacotherapy*. 2014; 34: 670–676. <https://doi.org/10.1002/phar.1442> PMID: 24855041
 42. Smythe M, Hoffman J, Kizy K, Dmuchowski C. Estimating creatinine clearance in elderly patients with low serum creatinine concentrations. *Am J Hosp Pharm*. 1994; 51: 198–204. PMID: 8160670
 43. Winter ME. *Basic clinical pharmacokinetics*. 3rd ed (ed. Koda-kimble M A., Young L Y). 474–499 (Applied Therapeutics, Inc, 1994).
 44. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol*. 1996; 49: 1225–1231. [https://doi.org/10.1016/s0895-4356\(96\)00002-9](https://doi.org/10.1016/s0895-4356(96)00002-9) PMID: 8892489
 45. Aston ML, Wener MH, Thomas RG, Hunder GG, Bloch DA. Overtraining in neural networks that interpret clinical data. *Clin Chem*. 1993; 39: 1998–2004. PMID: 8375090