Artificial Intelligence in Nephrology: Advancements, Opportunities, and Concerns in Hemodialysis

Nataphut Boonvisuth, Khajohn Tiranathanagul

Division of Nephrology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Abstract

Artificial Intelligence (AI) is a rapidly evolving field that is making inroads into various industries, including medicine. The advancements in AI technology have demonstrated their potential to improve diagnostic accuracy, treatment outcomes, and overall patient well-being, making it a valuable tool for healthcare professionals and institutes. This article provides an overview of the history and terminology of AI, including machine learning and deep learning. It also examines the benefits and limitations of AI in medicine, with a specific focus on its application in the field of nephrology. In this area, AI has demonstrated its potential to enhance patient care via clinical decision support systems, particularly in hemodialysis. The article highlights how AI is being used in various aspects of hemodialysis, including anemia management, dialysis adequacy and service planning, arteriovenous access assessment, dry weight prediction, intradialytic adverse event detection, mineral and bone disorder management, mortality and cardiovascular disease prediction, and cognitive function assessment. The goal is to provide readers with a preliminary understanding of AI and its potential to transform the practice of nephrology in the future.

Keywords: machine learning; AI; HD; dialysis; ESKD; ESRD; kidney failure; renal failure

Corresponding author: *Nataphut Boonvisuth*
Email: *datanbbnjunk@yahoo.com*

Received: 5 June 2024; Revised: 11 July 2024; Accepted: 22 July 2024

ปัญญาประดิษฐ์ในโรคไต: ความก้าวหน้า โอกาส และข้อกังวลในการฟอกเลือดด้วยเครื่องไตเทียม

ณรฐพุฒิ บุญวิสุทธิ์, ขจร ตีรณธนากุล

สาขาวิชาโรคไต ภาควิชาอายุรศาสตร์ คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อ

ปัญญาประดิษฐ์เป็นสาขาทางวิทยาศาสตร์ที่มีการพัฒนาอย่างรวดเร็วอย่างยิ่งในช่วงสิบปีที่ผ่านมา และเข้ามามีบทบาทในศาสตร์สาขาต่างๆรวมถึงทางการแพทย์ ความก้าวหน้าในเทคโนโลยีนี้ได้แสดงให้เห็นถึงศักยภาพในการปรับปรุงความแม่นยำในการวินิจฉัย ผลการรักษา ก่อให้เกิดประโยชน์ต่อทั้งผู้ป่วยและแพทย์ผู้ดูแล รวมไปถึงความคุ้มค่าต่อผู้ประกอบการ บทความนี้จะทำให้ผู้อ่านได้ทำความเข้าใจกับคำศัพท์ต่างๆที่เกี่ยวข้องกับปัญญาประดิษฐ์และระบบการเรียนรู้ต่างๆ จากนั้นกล่าวถึงประโยชน์และข้อจำกัดของปัญญาประดิษฐ์ในทางการแพทย์ และจะลงรายละเอียดถึงการประยุกต์ใช้เทคโนโลยีนี้กับสาขาอายุรศาสตร์โรคไต ซึ่งได้แสดงศักยภาพที่โดดเด่นผ่านระบบสนับสนุนการตัดสินใจทางคลินิก โดยเฉพาะอย่างยิ่งการนำไปใช้กับการบำบัดทดแทนไตด้วยการฟอกเลือดด้วยเครื่องไตเทียม ในปัจจุบันเทคโนโลยีนี้ถูกนำมาใช้ในหลายแง่มุมที่เกี่ยวกับการฟอกเลือด ไม่ว่าจะเป็นการจัดการกับภาวะโลหิตจาง การวางแผนและกำหนดนโยบายของศูนย์ฟอกเลือด การประเมินเส้นฟอกเลือด การทำนายน้ำหนักแห้ง การทำนายภาวะแทรกซ้อนขณะฟอกเลือด การจัดการกับความผิดปกติของสมดุลแร่ธาตุและกระดูก การทำนายการเกิดโรคหลอดเลือดหัวใจหรือการเสียชีวิต และสุดท้ายคือการประเมินการรู้ความเข้าใจของผู้ป่วย บทความนี้จึงมีเป้าหมายให้ผู้อ่านมีความเข้าใจเบื้องต้นเกี่ยวกับปัญญาประดิษฐ์ และศักยภาพในการเปลี่ยนแปลงแนวทางปฏิบัติของโรคไตในอนาคต

คำสำคัญ: ไตเรื้อรัง; ฟอกเลือด; ฟอกไต; โรคไตเรื้อรังระยะสุดท้าย

Introduction

Artificial Intelligence (AI) has been gaining significant attention in recent years as a rapidly growing field that has begun to permeate various industries, including medicine. Advancements in AI technology have been shown to have a positive impact on diagnostic accuracy, treatment efficacy, and overall patient outcomes. The integration of AI in the medical field has opened new doors for healthcare professionals and has the potential to revolutionize the way medicine is practiced.

This article provides an overview of AI terminology, highlighting its opportunities and concerns. We also explore the specific use of AI in the complex field of nephrology, with a particular focus on current advancements in hemodialysis.

History and Terminology

In 1956, a group of computer scientists introduced the term artificial intelligence (AI), defined later as "a subject about the study and development of computer systems that can copy intelligent human behavior"[1]. To mimic human behavior, this field brings together many disciplines such as computer science, mathematics, data science, and engineering. It focuses on the interaction between humans and computers. An example of an early AI application is the old-fashioned Chatbot, a script-based program that

ผู้ประพันธ์บรรณกิจ: *ณรฐพุฒิ บุญวิสุทธิ์*
อีเมล: *datanbbnjunk@yahoo.com*

รับบทความ: 5 มิถุนายน 2567; ปรับปรุงแก้ไข: 11 กรกฎาคม 2567; รับตีพิมพ์: 22 กรกฎาคม 2567

looks for keywords in human questions and then gives back the answer that has already been coded. However, despite its usefulness, the problem with this approach is its rigidity.

Another subset of AI programs that can think for itself is developed and called Machine learning (ML). ML is a program or model that enables computers to learn from data and even improve themselves without being explicitly programmed. Input and output data are fed through these algorithms, resulting in the development of models. Its debut in the image classification competition in 2012, earned the first title with a significant victory over other conventional mathematics models[2]. There are two types of ML: conventional ML models such as decision trees[3], random forests[4], and support vector machines[5]. These models use simple network architecture, are capable of handling small to medium data sizes, have good performance on structured data (tabular, categorical, numerical), and require low computational resources. The other type is deep learning (DL), which is more complex and requires more computational power.

DL is a type of ML program that relies on building multiple layers of non-linear modules and replacing the processed data with a new layer[6]. This allows more complex functions to be created. The core principle of DL is a multi-layered substructure that is beyond human ability to design, leading to solutions for complex tasks that traditional programs could not do and the discovery of hidden associations that human beings could not see. DL can handle both structured data and unstructured data (image, audio, text) quite well. This can be classified into many types, depending on the structure of the program and the type of data, such as recurrent neural network (RNN)[7] or convolutional neural network (CNN)[8].

Another term that is gaining attention is “Big data”, representing a large volume of structured and unstructured data. To improve the performance of AI systems, large amounts of data are required to train these models. Therefore, big data and AI are closely related and often used together to achieve better results.

Details on each type of model and its infrastructure are not in the scope of this article.

An explanation of each term is shown in **Figure 1.**

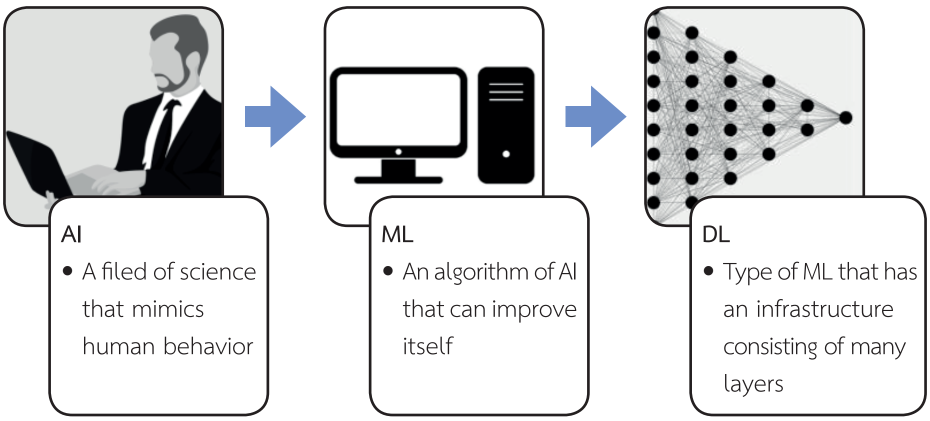

Figure 1 Terminology of each word. Artificial intelligence (AI); Machine learning (ML); Deep learning (DL)

Categories of ML by Learning Method

The concepts of each type of machine learning can be categorized into three ways: supervised learning, unsupervised learning, and reinforcement learning. Each has different approaches and rationales behind it.

Supervised learning is a method of machine learning that uses labeled data, where the output of a model is labeled by a human expert. Its purpose can be classified into prediction (or regression) or classification. For instance, the application for prediction of intradialytic hypotension (IDH) from the trend of vital signs. The target of this model would be to identify the occurrence of IDH events. When the model finished the learning session by recurrent neural network. It could predict the IDH event, by using the next set of vital signs[9]. Another example is in the classification of the severity of lupus pathology, where the model has each class of lupus pathology as a target and can recognize images using a convolutional neural network[10]. The choice of supervised learning algorithm also depends on the type of data used, such as linear regression being commonly used for continuous data, and decision trees/random forests for categorical data. Supervised learning is widely used in various medical fields, as it can handle labeled data. Further details on designing the models can be found in other sources.

On the other hand, unsupervised learning is a method of machine learning that does not involve any labeled data and instead, focuses on finding hidden patterns between input variables. Clustering and association finding are techniques that can be used, such as the K-means method to group cancer genes into different clusters. This technique has been utilized in CKD to identify co-expressed genes with shared biological functions, which has helped in identifying potential therapeutic targets and biomarkers[11]. Dimensionality reduction, another technique, involves reducing the number of input variables by identifying and removing redundant or irrelevant ones while preserving as much of the original information as possible. Principal component analysis (PCA) is a common technique used for dimensionality reduction and has been used in glomerular diseases such as Ig A nephropathy to identify important genes that contribute to the disease[12]. This approach is valuable in the identification of potential therapeutic targets and biomarkers.

Reinforcement learning is a method of machine learning that focuses on learning through interaction with an environment. This learning algorithm learns from feedback in the form of rewards or punishments. The goal is to learn a policy that maximizes the total reward obtained from the environment. While this approach has been mainly applied in game playing and robotics, there are also some studies in the medical field, such as a study from Valencia[13] that explored the use of a reinforcement learning-based method for personalizing erythropoietin dosage in a dosing algorithm for hemodialysis patients.

In summary, both supervised, unsupervised, and reinforcement learning methods have their unique advantages and disadvantages, and the choice of usage depends on the type of data and the problems that need to be solved. Each of these subjects has far more detail than the scope of this review. **Table 1** provides an overview of the simplified rationales and differences in learning methods. Examples of these methods in both non-medical and nephrology (hemodialysis) fields are shown in **Figure 2**, which can provide a better understanding of these new concepts.

Table 1 Rationales, strengths, and limitations of each type of machine learning model by learning method.

Learning Method	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Rationales	Learning from labeled data	Learning from unlabeled data	Learning without dataset through interaction with environment and feedback
Strengths	Used for classification and prediction Ability to handle missing values	Find hidden patterns and structure Used for clustering and association finding	Learn from the experience and adapt to a complex and dynamic environment Used to optimize a long-term strategy
Limitations	Risk of overfitting Need caution in tuning to generalize to unseen data	Requires large amounts of data Results can be subjective Limited ability to make predictions or infer causality	Requires significant trial and error Can be computationally expensive

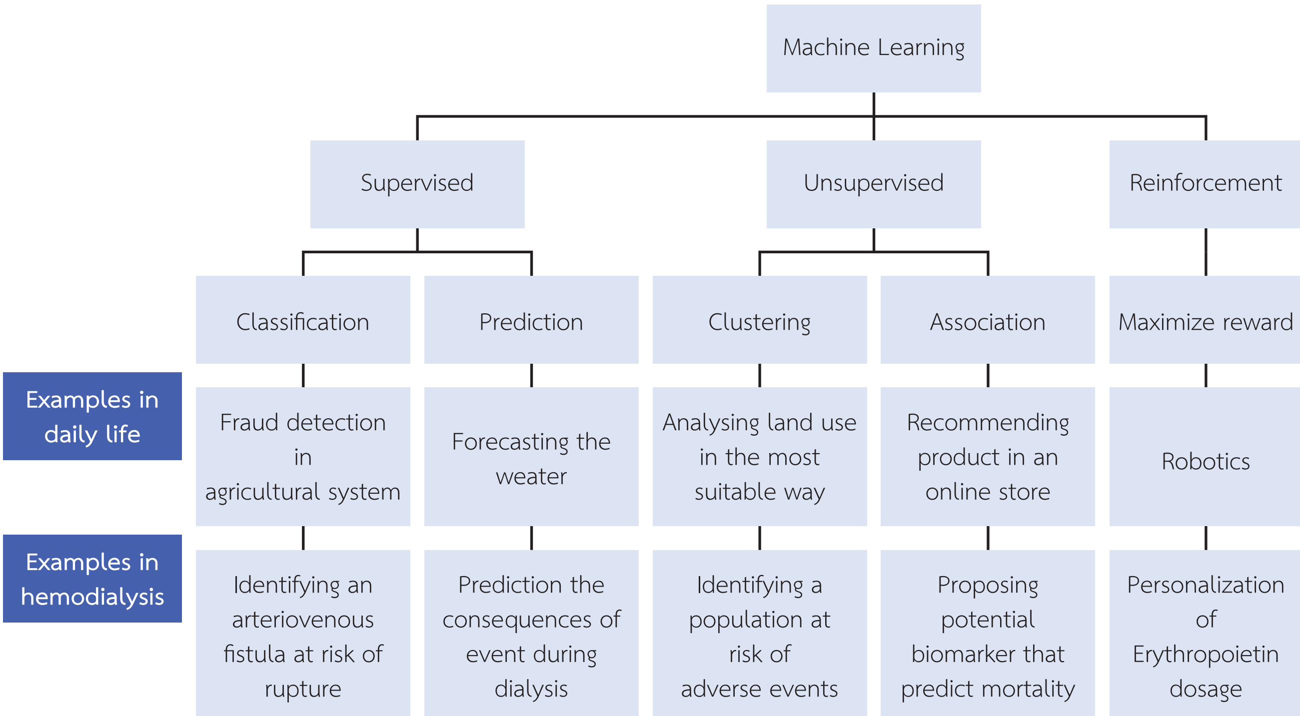

Figure 2 Mind map of learning method of ML and its example

Benefits and drawbacks of Artificial Intelligence

AI has several strengths and weaknesses that are worth considering. Among its advantages, AI can identify hidden correlations and connections, process a diverse range of inputs besides numerical data such as audio, video, images, and text, adapt itself to new incoming data, and play a critical role in precision medicine. AI can also serve as a component of clinical decision support systems (CDSS) which aid physicians in making more accurate decisions for each patient, reducing time, and minimizing human error, while also increasing cost-effectiveness. An example of the clinical implementation of CDSS will be mentioned in the next section.

However, AI also has certain limitations. One of the major drawbacks of AI is the complexity of its architecture, especially in the case of some machine learning models, such as deep learning. These models are more intricate than traditional models that rely on basic mathematical principles. There is a myriad of algorithms that can be used to solve a single problem, and each algorithm has its particularities that must be considered when selecting the appropriate one. Additionally, modifying the structure of an AI model can be challenging, as it may necessitate retraining the model from scratch. One major limitation of AI is its "black box" nature, where the causal links between visible inputs and outputs are not transparent. While some efforts have been made to clarify this issue, they are not widely generalized[14]. However, many explainable AI techniques have been developed to provide humanly understandable explanations of how an AI model arrives at its decisions. For example, SHapley Additive exPlanation (SHAP) values can help laypeople understand the effect of each parameter on the model. The quality and quantity of input data for the model are also crucial, as AI models require large and clean datasets to perform optimally. Furthermore, AI is prone to unconscious bias. A study from Berkeley found that some financial algorithms charged Latino and African American borrowers higher interest rates unintentionally[15]. Lastly, AI has raised concerns about privacy and ethical issues. As AI continues to be implemented in healthcare, there are concerns about the privacy and ethical implications of collecting and processing large amounts of patient data. The more data that AI systems use, the larger the electronic data will be available not only for developers but also for potential cybercriminals who may try to exploit it for malicious purposes Regulations should be developed to ensure patient privacy and security, including laws and policies that mandate data anonymization, encryption, and access controls. Healthcare organizations should also be transparent with patients about data collection and obtain explicit consent before using it for research or other purposes.

Overall, it is essential to balance the potential benefits of AI with these concerns. It is important to note that AI is not yet a replacement for physician judgment, but rather a tool that enhances and complements physician decisions. A summary of this section can be shown in **Table 2.**

Table 2 Pros and Cons of Artificial Intelligence

Pros	Cons
Identify hidden correlations and connections	Complex architecture
Process diverse inputs beyond numerical data	Algorithm selection challenges
Adapt to new incoming data	Modifying structure can be challenging
Aid in clinical decision-making, reducing time and human error	"Black box" nature makes causal links unclear
Play a critical role in precision medicine	Quality and quantity of input data crucial
Cost-effective	Prone to unconscious bias
	Raises concerns about privacy and ethical issues

How Artificial Intelligence is being used in Nephrology

To help nephrologists know how far this subject is being merged into their career. We searched for the MEDLINE cover period from database inception to Nov 20, 2022. A total of 1687 articles were found using the keywords "Artificial Intelligence" and "Nephrology" or "kidney disease." After an initial screening and exclusion of articles on renal mass and uro-oncology, 342 articles remained that were related to AI in nephrology. The majority of these papers were published after 2010. The papers were then further categorized into 9 domains within nephrology based on our institute subspecialty, including hemodialysis (HD) (n= 65), peritoneal dialysis (PD) (n=5), kidney transplantation (KT) (n=29), acute kidney injury (AKI) (n=71), chronic kidney disease (CKD) (n=72), critical care (n=16), glomerular disease (n=39), nutrition (n=3), and miscellaneous (n=42). The detail of the objective in each domain is shown in **Table 3.**

If we focus on the HD domains. In recent years advancements in AI have pushed the use of models in many ways. We divided these into a total of 8 subsections. As many doctors may not be familiar with this topic, we aim to provide them with a picture of how it relates to parameters encountered in everyday practice. To achieve this, we classify the inputs of the models into six groups: 1) demographic data (age, sex, height, status of diabetes, etc.), 2) laboratory data (sodium, calcium, etc.), 3) time series data (systolic blood pressure, diastolic blood pressure, etc.), 4) prescription data (dosage of medication, route of medication, etc.), 5) dialysis data (dialyzer type, dialysis vintage, vascular access, etc.), and 6) image data (picture of arteriovenous fistula, digital slides of renal pathology, etc.).

Table 3 Literature review in each domain.

Topic	Publications(N)	Objectives
HD	65	Prediction of adequacy of dialysis and service planning, Arteriovenous access assessment, Prediction of mortality, Prediction of CVD event, CDSS for anemia, CDSS for CKD-MBD management, Prediction of dry weight, Prediction of intradialytic adverse events, Prediction of cognitive function
PD	5	Prediction of PET, Prediction of mortality, Prediction of treatment failure, Prediction of CVD event, Finding pathogen-specific inflammatory response
KT	29	Prioritization for allocation of organs, Prediction of time on the waiting list, Prediction of renal survival, Prediction of graft function, Prediction of %IFTA, Prediction of tacrolimus level
AKI	71	Prediction event/severity of AKI in many operations including CABG, TAA repair, heart transplant, PCI, TKA, burn, Prediction event of AKI in CVD patients, Prediction of Contrast-induced nephropathy, Prediction of mortality post-AKI, Finding novel biomarkers
CKD	72	Prediction of RRT initiation, Prediction of DN progression, Prediction of ADPKD progression, Prediction of other CKD progression, Finding potential biomarkers/genes for CKD, CDSS for anemia management, CDSS for CKD-MBD management
Critical care	16	Prediction of renal recovery, Prediction of RRT dependent, Prediction of severe AKI in ICU, Prediction of Vol responsiveness in oliguric patients, Predict the trend of blood urea nitrogen level

Topic	Publications(N)	Objectives
Glomerular disease	39	Diagnosis of SLE, IgAN, MN, and others using data from either a light microscope or electron microscope, Prediction of Nephrotic syndrome in the general population, Prediction of histology of kidney biopsy from clinical parameters, Prediction prognosis of nephrotic syndrome patients
Nutrition	3	Predict the consequence of malnutrition patients, Prediction of potassium amount in the diet of the patient
Miscellaneous	42	Finding potential genes, and biomarkers for diseases e.g. AIN (IL1617), Prediction stone analysis, Prediction of urine 24-hour analysis of stone component, Develop a new Modified protocol for imaging, Image recognition for renal stone, hydronephrosis, and CAKUT, Prediction of hyperkalemia

Abbreviations: CDSS, clinical decision support systems; CKD-MBD, chronic kidney disease with the mineral bone disorder; PET, peritoneal equilibration test; CVD, cardiovascular disease; IFTA, interstitial fibrosis, and tubular atrophy; CABG, coronary artery bypass graft; TAA, Thoracic aortic aneurysm; PCI, percutaneous coronary intervention; TKA, total knee arthroplasty; RRT, renal replacement therapy; DN, diabetes nephropathy; ADPKD, Autosomal dominant polycystic kidney disease; SLE, systemic lupus erythematosus; IgAN, Ig A nephropathy; MN, membranous nephropathy; AIN, acute interstitial nephritis; CAKUT, congenital anomalies of the kidney and urinary tract.

Managment of anemia

The utilization of AI through CDSS in the prediction of erythropoiesis-stimulating agent (ESA) responsiveness is a particularly promising example of how AI can aid clinicians. It is by far the field with the most published articles by our search. The prediction of ESA responsiveness is a challenging task due to the complex interplay between various objective data such as the dosage of ESA, the route of administration, iron status, and the level of parathyroid hormone, all of which ultimately impact hemoglobin levels. While conventional statistical models were developed in the early 1990s, they were not successful in clinical applications[16]. In 2003, a study group from the University of Valencia used ML to uncover these hidden associations[17]. Since then, many ML models have been developed. An observational study in 2016 found that clinical adjustments with CDSS, specifically the "Anemia Control Model (ACM) by Fresenius Medical Care," led to a higher percentage of hemoglobin on target (76.6% versus 70.6%), less deviation in hemoglobin fluctuation (0.95 g/dl to 0.83 g/dl), and a lower mean erythropoietin dosage (0.63 to 0.46 mcg/kg/month), as shown in **Figure 3**[18]. The implementation of these ML algorithms in the form of CDSS serves as a prime illustration of the potential of AI in clinical decision-making. The utilization of these models has been shown to result in more precise treatment targeting, reducing the occurrence of adverse events for the patients, streamlining the prescription process for the physicians, and increasing cost-effectiveness for healthcare institutions. It is expected that in the future, the use of these AI-powered CDSS will become increasingly prevalent among nephrologists in their management of patients. The ability of these models to process large and diverse sets of data, adapt to new information, and identify complex correlations make them a valuable asset in the field of healthcare.

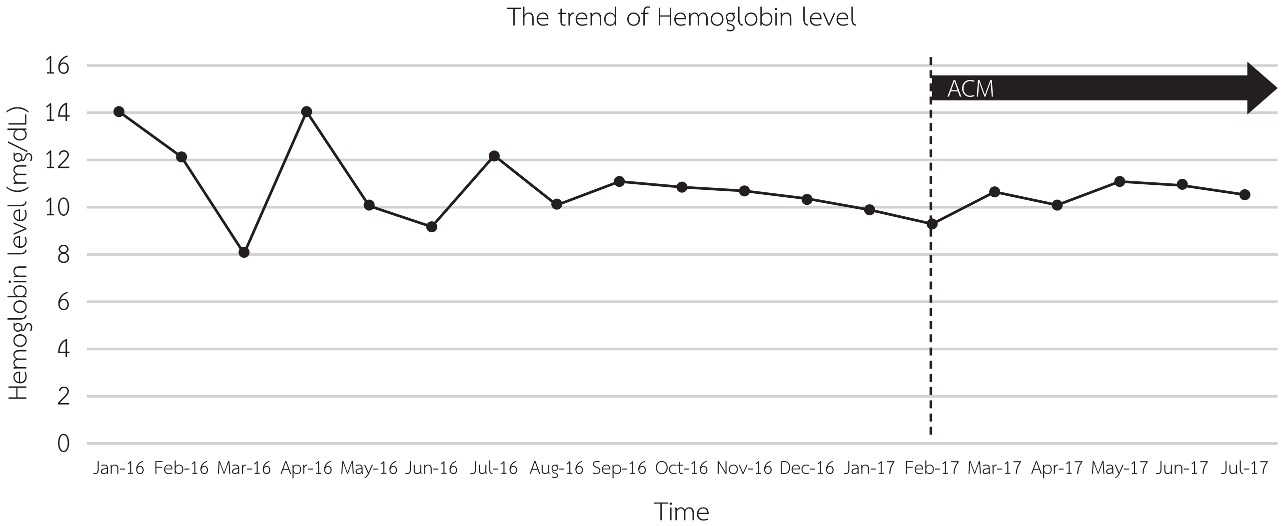

Figure 3 Temporal evolution of hemoglobin of the same patient.
The vertical dotted line represents the time of ACM introduction. Adapted from the study of Carlo Barbieri[18].

Dialysis adequacy and service planning

The utilization of AI in the management of dialysis units has proven to be a valuable asset in the face of the large quantity of data collected during dialysis sessions. By utilizing AI as an auditing system, it is possible to summarize these sessions from a clinical quality perspective, such as assessing the likelihood of clotting, session interruption, or the adequacy of the dialysis. This application of AI has been studied, with a total of 12 articles dedicated to this specific use. The majority of these studies utilize a diverse range of input data, including demographic data, laboratory data, time series data, and dialysis data, to construct a risk profile for each patient. This risk profile can then be used to prioritize surveillance management in clinics with limited resources, ultimately improving patient outcomes and streamlining clinical decision-making.

Assessment of arteriovenous access

The assessment of good arteriovenous fistula (AVF) and arteriovenous graft (AVG) is crucial for successful hemodialysis and is typically based on clinical examination by skilled medical healthcare professionals. However, this process can be time-consuming and may require additional visits for the patient. A study conducted by the Renal Research Institute has demonstrated that the utilization of ML algorithms can aid in the classification of AVF/AVG images into three stages of aneurysm/pseudoaneurysm, which have different levels of risk for AV access rupture, a serious complication that can lead to morbidity and mortality[19]. This ML model has been shown to accurately predict these stages with a 93% accuracy. The input is image data from a smartphone. This allows patients to easily send a picture of their AVF for evaluation during interdialytic days, without the need for an additional visit, thus promoting proactive management and reducing the burden on patients. Another study from the same group has focused on AVF patency, using demographic data (age, sex), laboratory data (albumin, ferritin, PTH), underlying disease (status of diabetes), and dialysis data (number of vascular access use, number of clots, dialysis vintage, Kt/v, etc.) to predict AVF failure within 90 days with a good AUC of 0.80[20].

The utilization of photoplethysmography (PPG) sensor systems is widespread in various applications, such as determining oxygen saturation and heart rate. A study conducted in Taiwan has explored the use of PPG sensors as a portable device for the assessment of AVF[21]. The study, which is a proof-of-concept, developed an ML model that transforms the PPG wave into a velocity comparable to that of doppler ultrasound. This model can classify AVF with abnormalities with a high accuracy of 89%. This technology could be a promising alternative to the traditional clinical examination, as it allows for a more convenient, non-invasive, and accurate assessment of AVF.

Dry weight prediction

The determination of dry weight (DW) is a crucial aspect of hemodialysis, as it serves as the foundation for determining the appropriate ultrafiltration volume for each patient. The maintenance of an accurate DW is vital as deviations, whether too high or too low, can lead to a variety of negative consequences. Despite the importance of accurate DW determination, traditional clinical assessment methods, such as those utilizing history taking and physical examination, have been known to be both inaccurate and insensitive. In recent years, several studies have attempted to replicate these methods through the utilization of ML techniques, with the ultimate aim of determining DW that is comparable to clinical-assessment DW[22, 23]. Inputs for these ML models typically include demographic data, laboratory data, underlying disease, and dialysis-related data. However, none of these models incorporate parameters from time-series data as an input. The performance of these ML-based models has been variable, and the application of these models in clinical settings remains a topic of ongoing debate.

Intradialytic adverse event detection

As mentioned above, the determination of an appropriate DW for patients undergoing hemodialysis remains a contentious issue, with various methods employed to assess optimal DW. Without knowing the true DW, some models aim for the prediction of consequences instead. The prediction of intradialytic adverse events, such as intradialytic hypotension (IDH) and cramping, has also been explored by 2 fashions. Predictive models utilizing static input parameters before the dialysis session have been implemented on a pilot basis. These models utilize a variety of input data, including demographic data (age, sex, height, etc.), laboratory results (sodium, calcium, etc.), underlying medical conditions (status of diabetes, hypertension, etc.), time series data as a timestamp (systolic blood pressure and diastolic blood pressure before dialysis session), and dialysis-specific data (dialyzer type, dialysis vintage, etc.). However, models utilizing non-static parameters (change of timestamps data during hemodialysis sessions such as systolic blood pressure, diastolic blood pressure, heart rate, and respiratory rate), have proven to be more effective in proactively mitigating adverse events, with many models demonstrating excellent performance in terms of area under the curve (AUC) for adverse events (0.83-0.90)[24, 25]. By utilizing such assistive tools, healthcare professionals can reduce their workload and more efficiently monitor adverse events.

Management of mineral and bone disorder

Regarding mineral and bone disorder (MBD) management, these studies focus on complex interactions between calcium, phosphate, vitamin D, PTH, etc. Although it uses the same concept as anemia management, the number of papers in this field is limited to six papers. One study from Taiwan, which utilizes underlying disease and laboratory data to predict in-range PTH, demonstrates promising results with an AUROC of 0.83 ($p=0.003$)[26]. Some studies have also shown that ML can predict vitamin D deficiency with higher accuracy than conventional logistic regression[27]. However, it should be noted that these studies are currently lacking in external validation, and their generalizability is thus limited.

Prediction of cardiovascular disease and mortality

In the realm of hemodialysis, the initiation of treatment is a crucial juncture at which patients are often forced to make decisions based on their intuition and the potential impact of treatment on their quality of life and survival. A period of a few months following the initiation of hemodialysis is a particularly high risk for mortality. Through the utilization of ML models, it has been demonstrated that the prediction of mortality in this early stage can be accomplished with greater accuracy than traditional statistical models (AUC 0.82 vs AUC 0.69-0.72)[28]. This can greatly aid patients in making informed decisions regarding their treatment. Furthermore, ML models have also been developed to predict 90-day and 5-year mortality with greater accuracy than conventional statistical models[29], which can facilitate patient-nephrologist shared decision-making for the continuation or cessation of dialysis with increased confidence.

In addition to predicting mortality, some studies in this area have also focused on the prediction of other hard outcomes such as cardiovascular events[30, 31]. Input data of this model consists of a wide variety of variables including demographic data (age, sex, ethnicity, region), laboratory data (hemoglobin, hematocrit, leukocyte, creatinine, eosinophil, iron status, albumin, electrolyte, uric, LDH, LFT, ESR, CRP, etc.), time-series data (vital signs, UFR, etc.), underlying disease (diabetes, hypertension, cystic kidney, etc.), and dialysis data (dialysis vintage, vascular access, Kt/v, etc.). Large-scale implementation of these models has been undertaken in Taiwan[32].

Furthermore, unsupervised learning has been utilized in some models[29, 31] to identify potential biomarkers that may serve as key pathways for predicting mortality or cardiovascular events, such as IL-12p70, AST, and miRNA. However, further research in the realm of basic sciences is needed to fully understand the associations between these biomarkers and clinical outcomes before they can be implemented in real clinical settings.

Cognitive function assessment

Cognitive impairment is a prevalent issue among individuals with End-Stage Renal Disease (ESRD), which is associated with poor outcomes. While the domain of memory is impaired earliest in Alzheimer's disease, cognitive impairment in ESRD usually presents with impairment of orientation, attention, and executive ability[33]. One-third of ESRD patients in their 50s and 60s have executive function impairment, with a prevalence of around 60% in patients over 60 years of age. Various factors may contribute to this phenomenon, including demographic data such as age and educational level, underlying diseases such as atherosclerosis and diabetes, and dialysis parameters such as retention of uremic toxins, hemodynamic instability, and anemia. However, its mechanism remains poorly understood[34]. Diagnosis of cognitive impairment is traditionally made subjectively by a neurologist using clinical scores such as the Montreal Cognitive Assessment Scale (MoCA). With advancements in neuroimaging technology, a study group from China has developed several ML models that predict MoCA scores using only image data from functional magnetic resonance imaging (fMRI)[35]. These models have shown good accuracy in predicting MoCA scores (RMSE 0.88, RMSE 2.40) and may help researchers better understand the relationship between cognitive function and neuroimaging in ESRD patients. However, their practical application is limited due to the cost and inaccessibility of fMRI screening, and they have yet to be validated with external datasets.

Details of important publications are provided in **Table 4.**

Conclusion

AI is transforming the way nephrologists diagnose and treat patients. The field of nephrology has been merged with AI in various ways. The promising concept of CDSS has been proven in beneficence not only to patients but also to physicians, and institutes. This is slowly changing the landscape of the management in the dialysis unit and will probably have a significant impact on the way we practice in the coming years.

Table 4 Summary of important publications

Subsections	Author, year of publication	Method of research#	Patient (N)	Input variables	AI technique$	Objectives	Results	Remark
Anemia management	Martin-Guerrero JD, et al., 2003	Prospective study	110	Demographic data (age, weight) laboratory data (hemoglobin, hematocrit, ferritin) prescription data (dosage of iron, dosage of erythropoietin its isoform, and its administration)	Profile-dependent support vector regressor (pdSVR) Multilayer perceptron neural network (MLP)	Optimize ESA dosage, improve patient quality of life, and avoid adverse effects of EPO	Profile-dependent support vector regressor improved patient quality of life and reduced treatment costs	Data from a single center
Anemia management	Barbieri C, et al., 2016	Retrospective study	752	Demographic data (age, sex, height, weight) laboratory data (hemoglobin, mean corpuscular volume ferritin, transferrin saturation index, albumin, white blood cells count, phosphate, C-reactive protein) prescription data (dosage of iron, dosage of erythropoietin its isoform, and its administration)	"Anemia Control Model, ACM" (ANN)	Percentage of hemoglobin values on target, median darbepoetin dose, and individual hemoglobin fluctuation	ACM decreased median darbepoetin consumption, increased on-target hemoglobin, and decreased hemoglobin fluctuation	The multicenter international observational study, not randomized, also includes prevalent HD patients, which may limit the generalizability
Arteriovenous access assessment	Krackov W, et al., 2021	Prospective study	30	Image data (extracted from panning video)	CNN	Classification of aneurysm grading	The model accurately (>90%) classified the grading of aneurysm	In a small study, the Quality of captured images may interfere with performance. Not yet evaluated the clinical impact of its classification

Subsections	Author, year of publication	Method of research#	Patient (N)	Input variables	AI technique$	Objectives	Results	Remark
Arteriovenous access assessment	Peralta R, et al., 2021	Retrospective study	13,369	Demographic data (age, sex, DM status, dialysis vintage,) laboratory data (hemoglobin, mean corpuscular volume ferritin, parathyroid hormone, albumin, C-reactive protein) dialysis data (dialysis treatment time, AVF vintage, ultrafiltration volume, blood flow rate, Kt/V, days since last use AVF, number of AVF use in 6mo)	“Arteriovenous fistula failure model; AVM-FM” (XG-Boost)	Prediction of composite AVF failure endpoint	The model can predict AVF failure with a high AUC-ROC score of 0.80 (95% CI 0.79-0.81)	Requires periodic data as input
DW prediction	Olivier Niel, et al., 2018	Prospective study	14	Time series data (blood pressure post dialysis) Dialysis data (bioimpedance analysis, relative blood volume measured by blood volume monitoring)	Multilayer perceptron neural network (MLP)	Prediction of Clinical DW	Predicted DW had a Mean difference +0.497 kg (LOA -1.33 to +1.29 kg)	Single-center, clinical DW is subjective and highly variable.
DW prediction	Xiaoyi Guo, et al., 2021	Prospective study	476	Demographic data (age, sex, dialysis vintage, body mass index) Time series data (diastolic blood pressure, systolic blood pressure, years of dialysis (YD), heart rate (HR), and body mass index	Multiple Laplacian-regularized radial basis function networks (MLapRBFN) Multi-kernel ridge regression (MKRR) Multiple kernel support vector regression (MKSVR) Linear regression (LR)	Prediction of Clinical DW	The prediction of clinical DW using ANN resulted in a mean difference of -0.04% with an agreement test, and the limits of agreement ranged from -4.4 to +4.33%.	Single-center, clinical DW is subjective and highly variable.

Subsections	Author, year of publication	Method of research#	Patient (N)	Input variables	AI technique$	Objectives	Results	Remark
Intradialytic adverse event detection	Liu YS, et al., 2021	Prospective study	108	Demographic data (age, sex, dialysis vintage, predialysis body weight) Time series data (blood pressure, heart rate, venous pressure, transmembranous pressure Prescription data (blood flow rate, ultrafiltration volume)	Support vector machines (SVM) Bayes point machine (BPM) Boosted decision tree (BDT)	Prediction of Intradialytic hypotension, Cramp, Intradialytic hypertension	Support vector machines can predict outcomes with an AUC of 0.83 - 0.93	Single center excluded some ultrafiltration-related features to reduce computing load, which may lead to loss of predictive accuracy
Intradialytic adverse event detection	Thakur SS, et al., 2018	Prospective study	109	Time series data (respiration rate, heart rate, body movement rate during dialysis session)	Adaptive Boosting (AdaBoost) k-Nearest Neighbors (k-NN) Support Vector Machine (SVM) Random forest (RF)	Prediction of a composite of various dialysis adverse events	The adaptive boosting model can predict dialysis-related adverse events with an ROC AUC of 90.16%	Not include time-series data from blood pressure monitoring
Mineral and bone disorder management	Wang YF, et al., 2006	Prospective study	161	Demographic data (age, sex, status of diabetes, hypertension) laboratory data (hemoglobin, albumin, calcium, phosphorus, alkaline phosphatase)	ANN	Prediction of in-range PTH	The model can predict in-range PTH with a high ROC AUC of 0.83 (0.72-0.94)	Had external validation in another institute
Mortality and Cardiovascular Disease Prediction	Rankin S, et al., 2022	Prospective study	1,150,195	Demographic data (age, sex, height, BMI, comorbid, time of hospitalization, nurse home status, nephrologist care status, ambulate status) Prescription data (erythropoietin use) Dialysis data (vascular access maturation) laboratory data (glomerular filtration rate, hemoglobin, albumin, creatinine)	eXtreme Gradient Boosting (XGBoost)	Prediction of death within 90 days of dialysis initiation	The model can discriminate mortality risk with excellent calibration and performs well across key subgroups	Predicting mortality can be a challenging task, especially when the mortality rate in the dataset is relatively low (8% of the USRDS cohort)

Subsections	Author, year of publication	Method of research#	Patient (N)	Input variables	AI technique$	Objectives	Results	Remark
Mortality and Cardiovascular Disease Prediction	Mezzatesta S, et al., 2019	Prospective study	522	Demographic data (age, sex, dialysis vintage, educational status, marital status, smoking, diabetes, malignancy, hypertension, cardiovascular comorbidities) Prescription data (use of antihypertensive drugs) laboratory data (hemoglobin, calcium, phosphate, albumin, liver enzymes, C-Reactive Protein, serum glucose, HbA1C)	Support Vector Machines (SVM) K-Nearest Neighbors (k-NN) Classification and Regression Trees (CART)	Prediction of death or cardiovascular disease initiation	SVM predicts outcomes with a high accuracy of 92%	External validation in American and Italian cohort
Cognitive function assessment	Zhang Y, et al., 2022	Prospective study	50	Image data (functional magnetic resonance imaging)	Gaussian process weighted least squares support vector machine (GPWLSV) Least squares support vector regression machine (LSSVM) Gaussian process support vector machine (GPSV) Gaussian process least squares support vector machine (GPLSV)	Prediction of the Montreal Cognitive Assessment Scale	GPWLSV can predict clinical score with root mean square error 2.40, mean absolute error 2.06	Small sample size, single-center study

The research methods used in the studies are categorized as either prospective or retrospective. Prospective studies involve developing and training the AI model using a dataset or simulation, followed by its evaluation before and after implementation in a real-world setting. Retrospective studies, on the other hand, evaluate an existing AI model based on its performance in a real-world setting. $ The AI technique column displays the various models used in each study, with the first name indicating the one with the best performance. The list of ML models included pdSVR, XGBoost, MKRR, MKSVR, LR, SVM, BPM, BDT, AdaBoost, k-NN, RF, CART, GPWLSV, LSSVM, GPSV, and GPLSV. The list of DL models included MLP, MLapRBFN, CNN, and ANN.

References

1. Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. J Family Med Prim Care. 2019;8(7): 2328-31.
2. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. Int J Comput Vis. 2015;115(3):211-52.
3. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci. 1997;55(1):119-39.
4. Breiman L. Random forests. Machine learning. 2001;45:5-32.
5. Vapnik VN. A note on one class of perceptrons. Automat Rem Control. 1964;25:821-37.
6. Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. Neural computation. 2006;18(7):1527-54.
7. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986;323(6088):533-6.
8. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE. 1998;86(11):2278-324.
9. Bae TW, Kim MS, Park JW, Kwon KK, Kim KH. Multilayer Perceptron-Based Real-Time Intradialytic Hypotension Prediction Using Patient Baseline Information and Heart-Rate Variation. Int J Environ Res Public Health. 2022;19(16):10373
10. Cicalese PA, Mobiny A, Shahmoradi Z, Yi X, Mohan C, Van Nguyen H. Kidney Level Lupus Nephritis Classification Using Uncertainty Guided Bayesian Convolutional Neural Networks. IEEE J Biomed Health Inform. 2021;25(2):315-24.
11. Guo Y, Ma J, Xiao L, Fang J, Li G, Zhang L, et al. Identification of key pathways and genes in different types of chronic kidney disease based on WGCNA. Mol Med Rep. 2019;20(3):2245-57.
12. Cao Y, Wang R, Zhang H, Zhai P, Wei J. Genetic Variants in MIR3142HG Contribute to the Predisposition of IgA Nephropathy in a Chinese Han Population. Public Health Genomics. 2022;25(5-6):209-19.
13. Martín-Guerrero JD, Gomez F, Soria-Olivas E, Schmidhuber J, Climente-Martí M, Jiménez-Torres NV. A reinforcement learning approach for individualizing erythropoietin dosages in hemodialysis patients. Expert Syst App. 2009;36(6):9737-42.
14. Vollmer S, Mateen BA, Bohner G, Király FJ, Ghani R, Jonsson P, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ. 2020;368: l6927
15. Bartlett R, Morse A, Stanton R, Wallace N. Consumer-lending discrimination in the FinTech era. J Financ Econ. 2022;143(1): 30-56.
16. Uehlinger DE, Gotch FA, Sheiner LB. A pharmacodynamic model of erythropoietin therapy for uremic anemia. Clin Pharmacol Ther. 1992;51(1):76-89.
17. Martin-Guerrero JD, Camps-Valls G, Soria-Olivas E, Serrano-Lopez AJ, Perez-Ruixo JJ, Jimenez-Torres NV. Dosage individualization of erythropoietin using a profile-dependent support vector regression. IEEE Trans Biomed Eng. 2003;50(10):1136-42.
18. Barbieri C, Molina M, Ponce P, Tothova M, Cattinelli I, Ion Titapiccolo J, et al. An international observational study suggests that artificial intelligence for clinical decision support optimizes anemia management in hemodialysis patients. Kidney Int. 2016;90(2):422-9.
19. Krackov W, Sor M, Razdan R, Zheng H, Kotanko P. Artificial Intelligence Methods for Rapid Vascular Access Aneurysm Classification in Remote or In-Person Settings. Blood Purif. 2021;50(4-5):636-41.
20. Peralta R, Garbelli M, Bellocchio F, Ponce P, Stuard S, Lodigiani M, et al. Development and Validation of a Machine Learning Model Predicting Arteriovenous Fistula Failure in a Large Network of Dialysis Clinics. Int J Environ Res Public Health. 2021;18(23):12355
21. Chao PC, Chiang PY, Kao YH, Tu TY, Yang CY, Tarng DC, et al. A Portable, Wireless Photoplethysomography Sensor for Assessing Health of Arteriovenous Fistula Using Class-Weighted Support Vector Machine. Sensors (Basel). 2018;18(11):3854
22. Niel O, Bastard P, Boussard C, Hogan J, Kwon T, Deschênes G. Artificial intelligence outperforms experienced nephrologists to assess dry weight in pediatric patients on chronic hemodialysis. Pediatric Nephrol. 2018;33:1799-803.
23. Guo X, Zhou W, Yu Y, Cai Y, Zhang Y, Du A, et al. Multiple Laplacian Regularized RBF neural network for assessing dry weight of patients with end-stage renal disease. Front Physiol. 2021;12:2240.
24. Liu YS, Yang CY, Chiu PF, Lin HC, Lo CC, Lai AS, et al. Machine Learning Analysis of Time-Dependent Features for Predicting Adverse Events During Hemodialysis Therapy: Model Development and Validation Study. J Med Internet Res. 2021;23(9):e27098.

25. Thakur SS, Abdul SS, Chiu HS, Roy RB, Huang PY, Malwade S, et al. Artificial-Intelligence-Based Prediction of Clinical Events among Hemodialysis Patients Using Non-Contact Sensor Data. Sensors (Basel). 2018;18(9):2833
26. Wang YF, Hu TM, Wu CC, Yu FC, Fu CM, Lin SH, et al. Prediction of target range of intact parathyroid hormone in hemodialysis patients with artificial neural network. Comput Methods Programs Biomed. 2006;83(2):111-9.
27. Sluyter JD, Raita Y, Hasegawa K, Reid IR, Scragg R, Camargo CA. Prediction of vitamin D deficiency in older adults: the role of machine learning models. J Clin Endocrinol Metab 2022;107(10):2737-47.
28. Rankin S, Han L, Scherzer R, Tenney S, Keating M, Genberg K, et al. A Machine Learning Model for Predicting Mortality within 90 Days of Dialysis Initiation. Kidney360. 2022;3(9):1556-65.
29. Gotta V, Tancev G, Marsenic O, Vogt JE, Pfister M. Identifying key predictors of mortality in young patients on chronic haemodialysis-a machine learning approach. Nephrol Dial Transplant. 2021;36(3):519-28.
30. Mezzatesta S, Torino C, Meo P, Fiumara G, Vilasi A. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. Comput Methods Programs Biomed. 2019;177:9-15.
31. de Gonzalo-Calvo D, Martinez-Camblor P, Bar C, Duarte K, Girerd N, Fellstrom B, et al. Improved cardiovascular risk prediction in patients with end-stage renal disease on hemodialysis using machine learning modeling and circulating microribonucleic acids. Theranostics. 2020;10(19):8665-76.
32. NVIDIA. Real-time Analysis of Massive Continuous Data from a Dialysis Machine to Predict Heart Failure Risk with New Edge AI Platform with NVIDIA: NVIDIA On-Demand; 2022 [Available from: https://www.nvidia.com/en-us/on-demand/session/gtcspring22-s41707/.
33. O'Lone E, Connors M, Masson P, Wu S, Kelly PJ, Gillespie D, et al. Cognition in People With End-Stage Kidney Disease Treated With Hemodialysis: A Systematic Review and Meta-analysis. Am J Kidney Dis. 2016;67(6):925-35.
34. Olczyk P, Kusztal M, Golebiowski T, Letachowicz K, Krajewska M. Cognitive Impairment in End Stage Renal Disease Patients Undergoing Hemodialysis: Markers and Risk Factors. Int J Environ Res Public Health. 2022;19(4):2389
35. Zhang Y, Sheng Q, Fu X, Shi H, Jiao Z. Integrated Prediction Framework for Clinical Scores of Cognitive Functions in ESRD Patients. Comput Intell Neurosci. 2022;2022:8124053.