REVIEW ARTICLE

Artificial intelligence enabled applications in kidney disease

Sheetal Chaudhuri[1,2] (iD) | Andrew Long[2] | Hanjie Zhang[3] | Caitlin Monaghan[2] |
John W. Larkin[2] | Peter Kotanko[3,4] | Shashi Kalaskar[2] | Jeroen P. Kooman[1] |
Frank M. van der Sande[1] | Franklin W. Maddux[2] (iD) | Len A. Usvyat[2]

[1]Maastricht University Medical Center, Maastricht, The Netherlands

[2]Fresenius Medical Care, Waltham, MA, USA

[3]Renal Research Institute, New York, NY, USA

[4]Icahn School of Medicine at Mount Sinai, New York, NY, USA

Correspondence
Len A. Usvyat, 920 Winter St, Waltham, MA 02451, USA.
Email: len.usvyat@fmc-na.com

Funding information
Fresenius Medical Care provided resources to support this review.

Abstract

Artificial intelligence (AI) is considered as the next natural progression of traditional statistical techniques. Advances in analytical methods and infrastructure enable AI to be applied in health care. While AI applications are relatively common in fields like ophthalmology and cardiology, its use is scarcely reported in nephrology. We present the current status of AI in research toward kidney disease and discuss future pathways for AI. The clinical applications of AI in progression to end-stage kidney disease and dialysis can be broadly subdivided into three main topics: (a) predicting events in the future such as mortality and hospitalization; (b) providing treatment and decision aids such as automating drug prescription; and (c) identifying patterns such as phenotypical clusters and arteriovenous fistula aneurysm. At present, the use of prediction models in treating patients with kidney disease is still in its infancy and further evidence is needed to identify its relative value. Policies and regulations need to be addressed before implementing AI solutions at the point of care in clinics. AI is not anticipated to replace the nephrologists' medical decision-making, but instead assist them in providing optimal personalized care for their patients.

1 | INTRODUCTION

Artificial intelligence (AI) is anticipated to transform health care through advancements in clinical decision support. Rapid advancements in computational power and improvements in statistical techniques ultimately enable AI to be leveraged to identify hidden interactions and patterns within large, complex, multi-level datasets. AI has been suggested as the next natural progression of traditional statistical techniques (eg, logistic regression, linear regression, etc), and these analytical advancements can be applied to the practice of medicine.[1,2] An AI-based "virtual coach" using a diverse set of inputs and algorithms may have the potential to aid in personalized medical guidance for patients.[3] AI medical decision support tools for clinicians may also improve efficiency by optimizing routine workflows and aid them in the process of providing care.[4]

In a recent bibliometric study on the global evolution of AI in health care and medicine, it is shown that clinical applications of AI are relatively common in fields like ophthalmology, oncology, and cardiology.[5] However, the use of AI is scarcely reported in nephrology, despite attributes of large datasets[6] and one of the highest disease burdens.[7] In-center hemodialysis (HD) is typically performed three times per week for 3-5 hours, thus amassing a large volume of clinical data captured in electronic medical records (EMR). These large treatment datasets are ideal for AI applications. With advances in technology, remote treatment monitoring applications allow clinical data to be collected from patients dialyzing at home. Recently, it has also become possible to measure and store beat-to-beat hemodynamic and respiratory values during dialysis treatment.[8] Furthermore, the emerging field of medical grade wearables is anticipated to yield even more robust data in all populations.[9]

The aim of this review is to: (a) provide an overview of the AI application process in a clinical setting; (b) provide brief descriptions of select advanced Machine Learning (ML) algorithms; (c) present the current status of AI in research toward kidney disease and dialysis; and (d) explore future pathways for AI within the discipline of nephrology. This review focusses on the applications of AI in progression to end-stage kidney disease (ESKD) and dialysis omitting the unique acute kidney injury population.

2 | TYPES OF AI

There is no universal definition of AI, but central to most definitions is the ability of a learning system to mimic human behavior. As depicted in Figure 1, AI is an umbrella term that brings together concepts from several fields such as computer science, statistics, algorithmic, ML, information retrieval, and data science at large.[10] ML techniques are very powerful in their ability to detect hidden patterns in large datasets that are otherwise difficult to identify by traditional statistical techniques.

The types of ML techniques that currently exist for building AI applications broadly fall into three families (Figure 2), namely supervised learning (SL), unsupervised Learning (UL), and reinforcement learning (RL). SL and UL are briefly discussed below, although technical details are beyond the scope of this review.[11] Most of the applications of RL are in the fields of board and video games and beyond the scope of this paper.

2.1 | Supervised learning

Supervised learning is the most frequently used type of ML. The objective of SL is to build a predictive model that takes historical input features to predict a specific output. For example, one may want to predict if a patient will miss their next dialysis treatment (binary output Yes/No) or predict how long it would take until a patient will transition to dialysis (continuous output).

Supervised learning can be divided into two categories (classification and regression) depending on the type of the output (Figure 2). In classification, the output belongs to a set of distinct classes (eg, missed treatment vs not missed treatment). In regression, the output is usually a continuous numerical quantity (eg, N days until transitioning to dialysis).

There are many ML algorithms for building predictive models ranging from traditional to more advanced methods. Prediction performance of these models is usually presented as area under the receiver operating characteristic curve (AUROC).[12] The most common traditional SL methods are logistic regression (for classification) and ordinary least squares regression.[13] These traditional methods are popular analysis techniques within health care and hence not discussed here for brevity. Over the past decade, more advanced techniques, such as tree-based methods and deep learning (DL) algorithms, have grown in popularity.

The foundation of tree-based methods is the decision tree, a ML technique for sequentially dividing the samples based on

FIGURE 1 Figure shows the relationship between artificial intelligence (AI), machine learning (ML), and deep learning (DL). ML is a subset of AI and DL is a subset of ML. ML is a sub-discipline of AI that uses training examples of how to perform a specific task without explicit instructions to identify associations for a given outcome measure. DL is a subfield of ML that mimics neural networks to learn [Color figure can be viewed at wileyonlinelibrary.com]

FIGURE 2 Supervised learning (SL) and unsupervised learning (UL) are the two main categories of machine learning (ML). Deep learning (DL) is a subset of ML. SL algorithms are used to learn the optimal parameters of the predictive model by investigating past examples with known inputs and known outputs. UL algorithms learn about patterns in the input data itself and does not have a known output [Color figure can be viewed at wileyonlinelibrary.com]
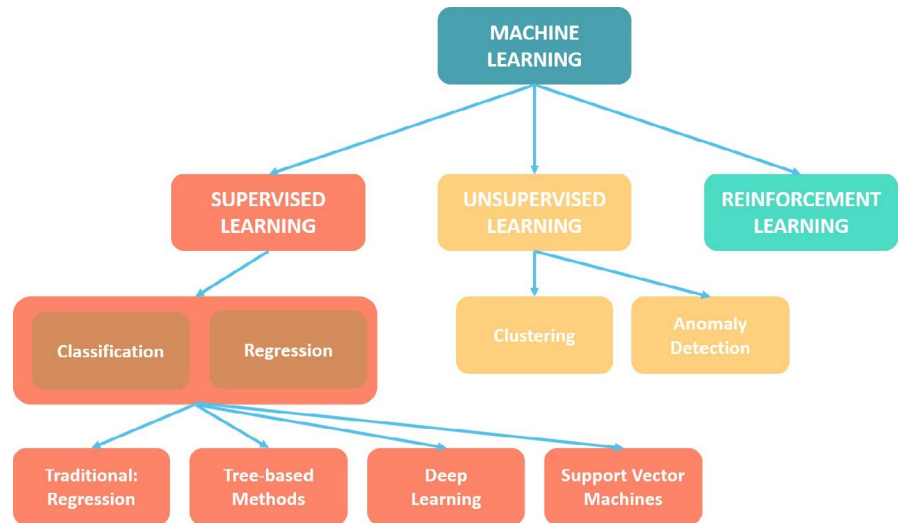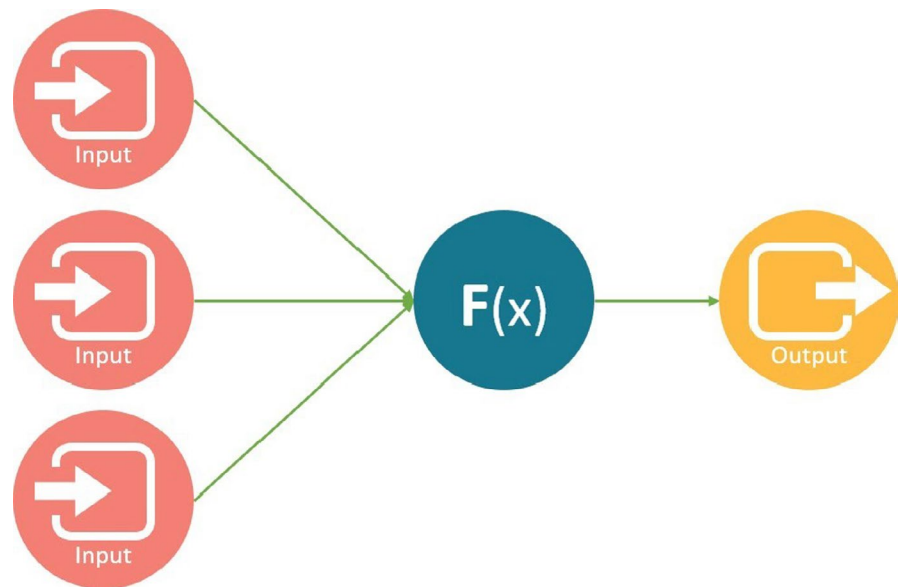
FIGURE 3 A very simple artificial neural network (ANN) with an input layer comprised of three inputs, hidden layer comprised of one neuron, and the output layer. ANN's neuron usually combines input from multiple sources through nonlinear activation functions [Color figure can be viewed at wileyonlinelibrary.com]

determining if a selected feature is greater than, or less than, a threshold determined by the model. At every level of the decision tree, the ML model learns which feature to use, and which threshold is the best. Unfortunately, a single decision tree can memorize the training data, resulting in poor performance on unseen data. As a result, many advanced analytical techniques (eg, random forest and Gradient Boosting Classifier) have been created to improve upon traditional single decision trees, increasing generalization to new data.[14] In random forest methods, multiple decision trees are created using random subsets of samples (ie, by bootstrapping) and random subsets of the input features (ie, bagging). On the other hand, Gradient Boosting methods sequentially add decision trees with few levels of nodes (shallow) that leads to a progressive improvement in model performance. One Gradient Boosting method known as XGBoost is currently one of the top performing models in the ML field.[15]

An extensive bibliography of new SL techniques, their application, and performance compared to traditional techniques is becoming available. Akbilgic et al compared several different ML modeling techniques to predict risk of death in incident dialysis patients.[16] The random forests model outperformed logistic regression with an AUROC of 0.76 compared to an AUROC of 0.68.

Deep learning, which uses artificial neural networks (ANNs), is another SL technique that has grown in popularity in the last decade. ANN began in the 1950s with the MADALINE algorithm,[17] but it was not until recently with advances in computational power that ANN/ DL could be computed in a reasonable time. The name ANN refers to its core functional unit, neuron (Figure 3). ANN's neurons usually receive multiple inputs that are mathematically combined through nonlinear (eg, sigmoidal) activation functions. A simplest neural network is the standard logistic regression. On the other hand, DL consists of stacking multiple layers of these units in the hidden layer (Figure 4). These layers connect to units of an output layer serving as the final output of the model.

The weights of the inputs are the parameters learned in ANN throughout the entire neural network. Given a set of weights, the
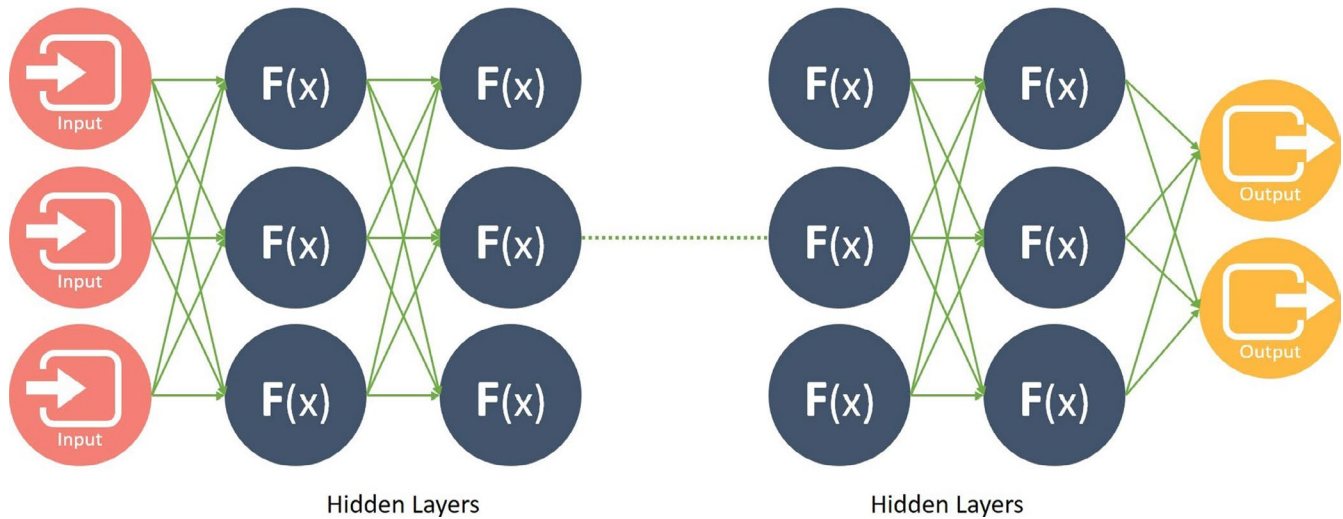
FIGURE 4 Deep Learning Network. Input layer with three inputs, multiple hidden layers of neurons, and two output layers. Higher the number of hidden layers deeper is the network [Color figure can be viewed at wileyonlinelibrary.com]

FIGURE 5 Convolution Neural Network (CNN) is a class of deep learning neural networks that is widely used for image classification. A CNN includes an input layer (image data), multiple hidden layers (convolution to extract features, pooling for subsampling features, and fully connected layer to classify images), and an output layer [Color figure can be viewed at wileyonlinelibrary.com]

training input features are fed forward through the neural network to create a set of predictions. The predictions are then compared to the actual output labels and this difference (ie, the error) is fed backward through the hidden layers. Over several iterations, the network "learns from its mistakes" and optimally adjusts its unit weights to a point where it can accurately predict the outcome.

To optimize these weights, the DL algorithm uses a technique known as backpropagation which was invented in the 1980s.[18] As the number of layers are added to the neural network, the number of weights and connections increase dramatically. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), as shown in Figures 5 and 6, are two variants of ANN that have also been created to reduce the number of weights, resulting in increases in performance, and decreases training time. A CNN is mostly used for image processing and RNN is widely for natural language processing (NLP).[19]

In the medical field, DL[20] (specifically CNN) has been mainly applied for image processing in the fields of radiology, histology, dermatology, and retinopathy, which has been able to demonstrate at or above clinical performance.[21-23] For example, in cardiology, DL has been used to predict outcomes after cardiac arrest.[24]

Support vector machine (SVM) is a form of SL, where the ML algorithm performs complex data transformations on the labeled data and defined output to draw boundaries within the input data. SVMs can be used to solve classification problem as well as a regression problem.[25]

2.2 | Unsupervised learning

In UL, there is no output label, but rather the objective is to learn about patterns in the input data itself. UL techniques usually focus on clustering, dimensionality reduction, or anomaly detection.[26] A commonly used clustering technique is k-means clustering.[27] k-means clustering utilizes an iterative refinement algorithm with assignment step and update step to partition the data into k clusters, and the algorithm aims to minimize the within-cluster variance and maximize the between-cluster variance. It is critical to determine an appropriate number of k clusters when using k-means clustering method. Hierarchical clustering[28] is another commonly used clustering technique that usually creates a hierarchy of clusters from top to bottom. For example, using hierarchical clustering, Liu et al identified clusters of US states based on unhealthy behaviors, preventive measures, and CKD-related outcomes in adults living in cities.[29] They concluded that such information may be of interest to policy makers to understand sociodemographic factors and other risk factors could contribute to the prevalence of CKD.
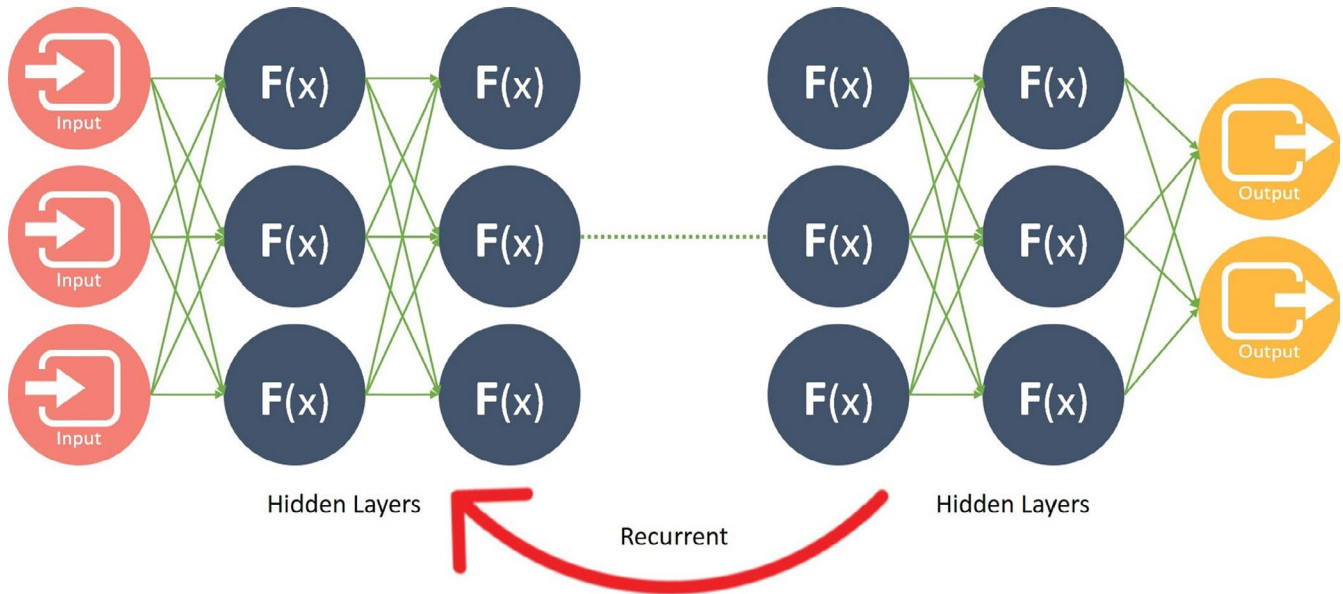
FIGURE 6 Recurrent neural network (RNN). In RNN, the output from the function is fed back in the model in order to minimize error [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Differences between traditional statistical methods vs advanced analytical techniques

Factors	Traditional statistical techniques	Advanced analytical techniques
Training data	Works with Smaller Datasets	Better with Large Datasets
Usability	Exploratory and baseline analysis	Iterative, complex, and ready to be deployed in clinical application
Interpretability	Easily interpretable	Complex techniques can be difficult to interpret
Hardware and training time	Requires simple hardware configuration and less training time	Complex models require powerful computing hardware and more training time
Types of input data	Works well only with categorical and numerical data	Works with all types of data including audio, image, free text
Examples	Logistic and Linear Regression, Generalized Additive Models, single decision tree	Neural Network, complex decision trees with several layers

Table 1 shows a very high level overview of the differences between the traditional statistical techniques and advanced analytical methods.

3 | AI APPLICATION PROCESS

The AI application process in a clinical setting generally consists of a series of stages (Figure 7). For ML, the process begins by defining the problem. This includes understanding the context of the clinical problem at hand and transforming the clinical problem into a relevant ML problem.

The next stage consists of understanding the quality and quantity of the clinical data available and preparation for modeling. Data preparation consists of collection, integration, cleaning, and using clinical knowledge to build predictors (feature engineering) for the ML model. In HD, the enormous amount of EMR data collected at the point of care provide a rich platform to employ ML. ML thrives on processing a huge number of variables combing them in nonlinear interactive ways. This capability allows new kinds of data (eg, free text, images, videos, sound, and temporal data) to be utilized. The volume and complexity of such data add additional challenges in analyzing the data.

With a set of well-engineered features, the predictive models are trained and tuned until acceptable performance is achieved.
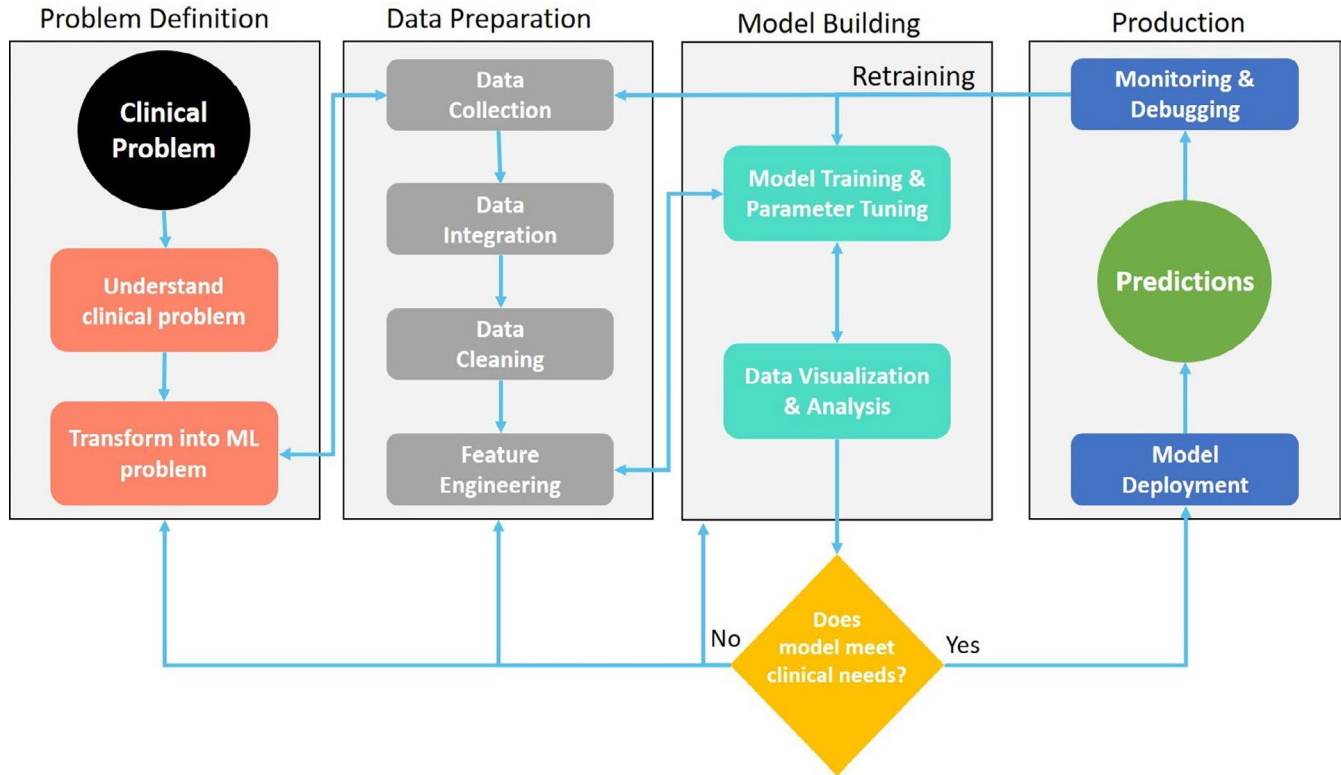
FIGURE 7 Process for application of artificial intelligence with four phases: Problem definition, data preparation, model building, and production [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Key publications of AI applications in kidney disease

Author, year	AI techniques	No of patients	Outcome predicted	Performance	Clinical application use
Akbilgic et al[16] 2019	Random Forest	27 615	Risk of death	AUROC: 0.70-0.76	NA
Goldstein et al[32] 2014	Random Forest	826	Sudden cardiac death	AUROC: 0.78-0.79	NA
Mezzatesta et al[33] 2019	Support Vector Machine	1216	Cardiovascular disease	Accuracy: 92.15%-92.25% AUROC: 0.50-0.74 Precision: 72%-89% Recall: 73%-94%	NA
Chauhan et al[39] 2020	Random Forest	1369	CKD progression	AUROC: 0.77-0.80 PPV: 62% in high-risk group NPV: 92%-96% in low-risk group	NA
Norouzi et al[42] 2016	Artificial Neural Networks	465	CKD progression	MSE: 58.63-64.00 MAE: 4.77-5.93 NMSE: 4.77%-4.88%	NA
Barbieri et al[46] 2016	Artificial Neural Networks	752	Anemia management	MAE: 0.59 g/dL	Yes
Zhang et al[74] 2017	Random Forest	83	Immune fingerprints	AUROC: 0.993 Sensitivity: 98.5% Specificity: 92.6%	NA

Abbreviations: AI, artificial intelligence; AUROC, area under the receiver operating curve; MAE, mean absolute error; MSE, mean square error; NMSE, normalized MSE; NPV, negative predicted value; PPV, positive predicted value.

It is anticipated that some steps of the process have bi-directional arrows because they can result in modifications for previous steps (Figure 7). If the model meets the needs of the clinical problem, the trained model can be deployed in production. During production, it is advisable to monitor the predictions and retrain the model when necessary.

4 | APPLICATIONS IN KIDNEY DISEASE

There are several unmet needs in nephrology and there is a huge potential for use of big data and AI in patients with kidney disease. The applications of AI kidney disease can be broadly subdivided into three main topics: (a) predicting events in the future; (b) treatment and decision aids; and (c) identification of existing, but unrecognized, patterns. Table 2 shows a summary of key AI-related studies that have been published in the field of kidney disease. Currently, only one published study reports the clinical application, although the use of AI in kidney disease is reported more commonly in conference abstracts suggesting that the scientific community has more contributions on the horizon.

4.1 | Predicting the future

4.1.1 | Predicting outcomes

Patients with ESKD have high mortality as well as hospitalization rates.[7] Prediction models can assist with early care planning and triaging resources where there is potential for the greatest clinical benefit. Interventions performed based on predictions would be specific to the outcome and may warrant ad hoc and/or extra evaluations and clinical screenings in addition to routine care. Early mortality and hospitalization prediction models using traditional statistical techniques built on a select number of features have been reported.[30] The AUROC for traditional statistical mortality and hospitalization prediction models usually fall in the range between 0.65 and 0.75.[31] In nephrology, prediction of sudden cardiac death in older HD patients was an early example of employing an advanced ML method where a random forest model yielded a AUROC of 0.79.[32] In another example, Mezzatesta et al used SVM, to predict the risk of ischemic heart disease in dialysis patients with an accuracy of approximately 92%.[33]

Recent studies show a large set of features and their interactions with other features can be employed using advanced ML methods to better estimate potential risk factors preceding mortality and/or hospitalization.[34] For instance, a large dialysis organization (LDO) of Fresenius Medical Care (FMC), an integrated kidney disease care organization, has developed and deployed a predictive model that includes more than 200 variables to identify patients treated with in-center HD who have an increased risk of hospitalization in the next 12 months. The model is built using XGBoost classifier with an AUROC of 0.81.

As part of a pilot study reported in a congress abstract, the LDO has suggested use of the predictive model to assist clinicians with targeting additional interdisciplinary assessments and interventions resulted in a decrease of the average yearly hospital admission rate and average yearly hospital days rate compared to controls in the neighboring region that did not participate in the pilot and receive predictive model reports.[35] Such prediction model appears to have the potential to provide an intelligent method of triaging additional resources in dialysis clinics.

Advanced AI applications are powerful in analyzing vast amounts of clinical data to look for subtle changes in a patient's condition or worsening status for short-term outcomes. Dial disorders/disease exacerbations sometimes exhibit clear symptoms in the days prior to an event; however, the occurrence of minor signals of a worsening condition that do not clearly warrant any immediate intervention or appear unrelated to the cause of the event that takes place soon after can be a clinical challenge. Recent efforts by the LDO reported in a congress abstract led to the development and implementation of a model to predict imminent hospitalizations in ESKD patients who are at risk of getting hospitalized within the next 7 days.[36] The model uses over 1500 variables from a range of data sources (eg, treatment vitals, laboratory measurements, comprehensive assessments, and nursing clinical notes). The unstructured clinical notes are converted into numerical data using NLP techniques, specifically word2vec and CNN. The output of the CNN is then combined with other structured numerical data to train an XGBoost tree classifier. The final model has an AUROC of 0.78. As reported in another congress abstract, this model is currently used by a team of nurses and has improved their workflow significantly.[37] Although the effectiveness of this imminent hospitalization model and subsequent interventions is unknown and being evaluated, its potential to assist clinicians with near real-time insights of risk levels and predictors driving the risk determination is promising and could help them with targeting interventions and transitional care planning before and after hospitalization episodes.

4.1.2 | Predicting chronic kidney disease progression

Chronic kidney disease (CKD) is a growing health crisis across the world.[7] Detecting it early and managing the progression of the disease are critical for positive patient outcomes and controlling health-care costs. Due to challenges in understanding the trajectory of this disease, providing care planning before initiation of dialysis and helping patients make appropriate vascular access and modality choices may be difficult.

Traditional and AI techniques are being developed to predict CKD progression. Tangri et al have developed a traditional regression model for prediction of kidney failure from CKD stages using demographic, clinical, and the most recent clinical data from two independent cohorts of CKD patients stages 3-5.[38] In two other recent studies, random forest models have been developed to generate a prognostic risk score by combining data from EMR and circulating biomarkers, such as plasma tumor necrosis factors and kidney injury molecule-1, to predict CKD progression.[39,40] The AUROC in one of the studies by Chauhan et al was 0.77-0.80. Xiao et al compared several ML methods to predict the risk of proteinuria >1 g/d in CKD patients using demographic data and blood biochemical features.[41] In this case, the traditional logistic regression model outperformed other ML models with AUROC

0.87. They conclude that advanced ML models are best when the amount of data is large, whereas linear models perform better in relatively smaller datasets. On the other hand, Jamshid Norouzzi et al developed an ANN to predict renal failure progression in patients with CKD. The model could accurately (>95%) predict the estimated glomerular filtration rate (eGFR) in 6, 12, and 18 months interval.[42]

As reported in a congress abstract, the Renal Research Institute used data from 28,608 patients with CKD from 2000 to 2011 to construct two linear and spline models that utilize up to 6 months of historic eGFRs, or logarithm of eGFRs (log-eGFRs), for prediction of CKD progression to ESKD.[43] The results of the model were integrated in the CKD Forecaster Tool used at the point of care for nephrologists in clinical decision support system. This helped in-patient education and care planning for the transition from CKD to ESKD. As reported in a congress abstract, nephrologists who used the CKD Forecaster Tool had less patients transitioning to HD with a central venous catheter.[44]

4.2 | Treatment aid

4.2.1 | Treatment and drug prescription

Prescription of drugs, such as erythropoietin, in patients with ESKD by clinicians is both time-consuming and error prone. Automation of part of the prescription process could increase efficiency and improve patient care. Several approaches have been published in the literature to reduce erythropoietin dose and increase the percentage of patients within target.[45] One example is adoption of ANN for anemia management, which was able to increase the percentage of patients in target while reducing hemoglobin variability and erythropoietin dose.[46]

Understanding which drugs are most appropriate for certain patient categories is another area where historic data can guide a decision-making process for the clinicians. For example, informed by results of virtual clinical trials utilizing advanced physiology-based mathematical models of parathyroid gland biology, an LDO of FMC, an integrated kidney disease care organization, has afforded nephrologists working in its clinics the opportunity to prescribe off-label 3× weekly directly observed in-center administration of cinacalcet as an alternative to daily dosing.[47,48] Subsequent observations in currently over 11 000 patients indicate that 3× weekly in-center administration of cinacalcet is noninferior to prescribed daily cinacalcet in controlling parathyroid hormone levels, corroborating the virtual clinical trial results.[49] Although speculative, efforts like this may potentially further optimize and personalize the treatment of secondary hyperparathyroidism, as well as expand the understanding of the debated influence of mineral bone disorder medications on hard outcomes.[50,51] In oncology, several studies show successful predictions of which patients would respond to immunotherapy using AI algorithms.[52,53] Further, ML algorithms have been used to predict which medications would work for which patients with mood disorders.[54]

4.2.2 | Identifying medical errors

Although there are not many references in literature on the use of AI in identifying medical errors in a nephrology setting, it is important to highlight how it can be used. Medical errors are a third leading cause of death in the United States; in 2016, they contributed to more than 251K deaths in the United States alone and accrued $17.8 billion dollars in unnecessary spend.[55,56] Different causes of medical errors exist, such as (a) complexity of the health-care system; (b) system and process design issues; (c) competency, education, and training; and (d) human factors and ergonomics.

Traditional approach to correct medical errors is to create new rules and procedures that need to be utilized in a health-care setting.[57] However, data-driven, AI approaches can also be applied particularly when historic evidence already exists. Most common application of AI in minimizing medical errors is to guide what therapeutic approaches may or may not be ideal for a given patient. Paredes et al[56] explored this in the context of US intensive care units and concluded that ML could aid physicians by providing better predictions about the effect of certain treatments and the likely evolution of sepsis patients.

Further, ML algorithms can assist in guiding decisions where complex, time-dependent, or uncommon medication interactions are at play (such as drug-drug or drug-allergy interactions, therapeutic duplication, etc). Traditional rule-based decision support systems may be insufficient to resolve such issues. AI and technology solutions are likely to be best fitted in these applications.[58] Specific examples of these applications have been successfully demonstrated by prediction algorithms developed at Stanford University.[59] Many technology companies have services that support physicians as they interact with their patients' data that may assist in minimizing medical errors.

Outlier management and outlier detection can also assist with minimizing medication errors. This can be completed through AI-driven algorithms or through Clinical Decision Support models. A team at the Brigham and Women's Hospital evaluated a medication error detection system that uses a probabilistic ML model to identify prescriptions that are outliers based on populations of patients in their EMR system with similar characteristics.[60]

4.3 | Identifying patterns

4.3.1 | Identifying phenotypical patterns

In patients with ESKD, several patterns, such as the malnutrition-inflammation-atherosclerosis syndrome, have been discovered by traditional statistical methods. It has thus increased our pathophysiological understanding and were shown to be strong prognostic indicators. Recently, studies have shown that fluid overload also can be part of a pathophysiologic spectrum including malnutrition and inflammation.[61,62] The concomitant presence of these three risk factors yielded a near six-fold increase in mortality risk.

However, unlike other chronic diseases, pattern detection based on UL techniques has not yet been published in nephrology. In patients with heart failure with preserved ejection fraction, three different phenotypical patterns were identified based on clinical, laboratory, and echocardiographic parameters by agglomerative hierarchical clustering. These clusters differed greatly in mortality risk. In cardiology, the use of UL techniques to detect phenotypical patterns was termed "phenomapping" by the authors.[63] Another example, infection medicine is based on k-means clustering on a cohort of patients with sepsis. Four different phenotypes were observed with a distinct difference in outcome, of which one was characterized by older patients with more chronic illness and renal dysfunction (β phenotype). The highest 28-day mortality (40%) was observed in the δ phenotype, characterized by patients with septic shock and liver dysfunction, as compared to 13% in the β phenotype and 5% in the α phenotype with the lowest risk.[64] Another study identified different metabolic clusters based on k-means clustering including a set of clinical parameters and biomarkers in older adults without diabetes. In the clusters characterized by lower eGFR and albuminuria and the cluster with the highest inflammation, the risk of cardiovascular endpoints was comparable to the diabetic cluster.[65] Whether phenomapping in different diseases has relevance for personalized treatment prescription needs to be addressed in future trials.

4.3.2 | Identifying unknown comorbidities

In addition to making predictions about the future, the power of AI can be utilized to comb through vast amounts of information to uncover hidden patterns in high dimensional data otherwise too complex to identify manually. While an incredible resource of clinical data, EMR consist of both structured and unstructured data, with contributions often added by multiple care providers with different documentation styles and levels of thoroughness. Variations and inconsistencies in a patient's record likely increase with the complexity of their health.

One area of concern involves patient comorbidities. ESKD patients with multiple medical comorbidities face decreased survival likelihoods.[66,67] Prognostic comorbidity indexes indicating patient mortality risk have been used and adapted for renal replacement populations,[68-70] which highlight the critical role that comorbidities play in the complexity of a patient's health picture. In addition to prognostics, comorbidity information is a necessary component involved in medical billing. Medicare's bundled fee-for-service coverage for beneficiaries with ESKD includes payment multipliers for patients with complex health pictures based on specific comorbidities. In order to receive appropriate payment for the extra level of support and services tied to these populations, comorbidities must be properly documented in medical records. In nephrology, one LDO within its integrated kidney disease care organization addressed this clinical need by using ML to find patterns in physician notes common across diseases to identify potential undocumented comorbidities or

to remove comorbidities that are unlikely to exist.[71] Using lab test results, NLP of physician notes, and demographic information, the LDO was able to improve coding over the previous method of randomly chosen manual medical record reviews.

4.3.3 | Image classification for arteriovenous fistula aneurysm and biomarker fingerprints

The Renal Research Institute developed a CNN to automatically classify arteriovenous fistula aneurysm (AVFA) stages. They collected 15-20 seconds panning videos from 30 patients to train a CNN model. CNN was able to automatically classify AVFA stages with >90% classification accuracy. As reported in a congress abstract, using this model in a clinical application will reduce workload for physicians, provide timely AVFA diagnosis, and improve patient care.[72]

Advances in biochemical analytics, such as liquid chromatography-mass spectrometry, provide an unparalleled amount of data from biological samples, giving rise to the rapidly evolving field of metabolomics. A major area of research is to explore if specific compound patterns are correlated with clinical outcomes of interest or if patterns differ between clinical phenotypes. Given the enormous number of metabolites, this question lends itself to the use of AI. Very recently, several groups have successfully applied ML to metabolomics data.[73] Another example of a potential clinical application for AI in peritoneal dialysis was presented by Zhang et al, who used a combination of supervised ML methods to detect specific immune fingerprints allowing rapid detection of causative organisms in peritonitis, potentially facilitating earlier prescription of specific antibiotic treatment.[74] This study demonstrated the power of using advanced analytical model for mining complex biomedical dataset where traditional statistical methods fail to yield satisfactory results.

5 | REFLECTION

Rapid advances in computing, mathematics, and statistics have resulted in the evolution of AI and ML methods. Cloud computing resources might be a more cost-effective way of analyzing large volumes of data and building ML models. Ideally, ML algorithms should be available for use in the community.

Traditional statistical modeling techniques are most appropriate in building simple predictive models, where one has a well-defined problem, good observation set, and established knowledge expertise about the strengths and limitations of the outcomes. Furthermore, traditional techniques learn from data which are static in time, and thus tend to "overfit" to their training, and fare reasonably poorly when they encounter anomalous instances.

However, in an ever-evolving renal care landscape where the problems posed are complex, AI provides several techniques to derive meaningful results. It is very powerful in identifying unknown patterns and anomalies.

Therefore, traditional statistical and advanced AI techniques are both complementary. It is a widely accepted practice to initially build a model using traditional statistical techniques and use that model as a baseline against which AI models are compared against for performance. In a systematic review in the general population, no major differences were found between various advanced AI techniques and traditional statistical modeling techniques in clinical prediction.[75] A prudent approach is to choose a model appropriate for the problem at hand and not necessarily bias oneself to one methodology. Beyond methodologies, it is important to translate modeling results into actionable decision points for patients and care providers. At present, the use of prediction models in dialysis treatment is still in its infancy and further evidence is needed to identify its relative value.

AI techniques can allow for large datasets to be leveraged with minimal efforts. Such techniques have the power to process large volumes of data to identify patterns and features which may impact the outcome. However, outcome selection and follow-up timeframes need to be carefully determined to optimize the performance and potential clinical value. AI model that can predict short-term outcomes may not allow time for interventions to change the course of an event.

AI solutions must follow ethical guidelines and consider at the time of conception whether software programs are medical devices that require formal regulatory pathways and trials.[76,77] Furthermore, before implementation of AI solutions at the point of care, policies and regulations need to be established for delivery of the outputs to clinicians and patients. Models are never 100% accurate, and thus there will be instances where models will predict incorrectly. In such situations, a precedent of accountability needs to be established. AI solutions should be transparent and traceable. It is important that the predictive models use data collected routinely in standard of care or it will likely produce models that are bias by indication. Teams developing and using AI solutions should be aware of this limitation. Thorough evaluation of the input data variables should be conducted as a key step in the selection of outcomes and the process of building predictive models.

While a lot of emphasis is placed in developing powerful and accurate models, more emphasis should be directed toward building an end-to-end team of practitioners in data analytics, data engineering, trainers, care providers, and patients to create effective solutions which would be beneficial for all stakeholders. The effectiveness of the prediction models depends heavily on the ability to use insights to make clinical interventions. On the other hand, interventions need to be thoroughly thought through depending on unique factors driving the clinical outcome and personalized for every patient.

Lastly, AI solutions when implemented at the point of care for nephrologists should be viewed as a clinical decision support tool to extend providers' insights about the patients. AI is not anticipated to replace providers' medical decision-making, but instead assist them in providing optimal personalized care for their patients.

ORCID

Sheetal Chaudhuri 🆔 https://orcid.org/0000-0002-8217-9520
Franklin W. Maddux 🆔 https://orcid.org/0000-0003-1395-2011

REFERENCES

1. Noorbakhsh-Sabet N, Zand R, Zhang Y, et al. Artificial intelligence transforms the future of health care. *Am J Med.* 2019;132:795-801.
2. Usvyat L, Dalrymple LS, Maddux FW. Using technology to inform and deliver precise personalized care to patients with end-stage kidney disease. *Semin Nephrol.* 2018;38:418-425.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25:44-56.
4. Torkamani A, Andersen KG, Steinhubl SR, et al. High-definition medicine. *Cell.* 2017;170:828-843.
5. Tran B, Vu G, Ha G, et al. Global evolution of research in artificial intelligence in health and medicine: a bibliometric study. *J Clin Med.* 2019;8:360.
6. Saez-Rodriguez J, Rinschen MM, Floege J, et al. Big science and big data in nephrology. *Kidney Int.* 2019;95:1326-1337.
7. Xie Y, Bowe B, Mokdad AH, et al. Analysis of the Global Burden of Disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. *Kidney Int.* 2018;94:567-581.
8. Meyring-Wösten A, Zhang H, Ye X, et al. Intradialytic hypoxemia and clinical outcomes in patients on hemodialysis. *Clin J Am Soc Nephrol.* 2016;11:616-625.
9. Wieringa FP, Broers NJH, Kooman JP, et al. Wearable sensors: can they benefit patients with chronic kidney disease? *Expert Rev Med Devices.* 2017;14:505-519.
10. Chaudhuri S, Zhang H, Long A, et al. How artificial intelligence impacts the treatment of kidney disease. *Nephrology News & Issues.* 2019;33. https://www.healio.com/nephrology/kidney-care-community/news/print/nephrology-news-and-issues-%7B44a4c36f-56c7-4f4c-9cbe-68e1ff74f03f%7D/how-artificial-intelligence-impacts-the-treatment-of-kidney-disease. Accessed November 22, 2019.
11. Niel O, Bastard P. Artificial intelligence in nephrology: core concepts, clinical applications, and perspectives. *Am J Kidney Dis.* 2019;74:803-810.

12. Narkhede S. Understanding AUC—ROC Curve. In: *Towards Data Science*; 2018. https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5. Accessed November 22, 2019.

13. Alto V. Understanding the OLS method for simple linear regression. In: *Towards Data Science*; 2019. https://towardsdatascience.com/understanding-the-ols-method-for-simple-linear-regression-e0a4e8f692cc. Accessed November 22, 2019.

14. Chen L. Basic ensemble learning (random forest, adaboost, gradient boosting)—step by step explained. In: *Towards Data Science*; 2019. https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient-boosting-step-by-step-explained-95d49d1e2725. Accessed April 15, 2020.

15. Morde V. XGBoost algorithm: long may she reign. In: *Towards Data Science*; 2019. https://towardsdatascience.com/https-medium-com-vishalmorde-xgboost-algorithm-long-she-may-rein-edd9f99be63d. Accessed April 15, 2020.

16. Akbilgic O, Obi Y, Potukuchi PK, et al. Machine learning to identify dialysis patients at high death risk. *Kidney Int Rep*. 2019;4:1219-1229.

17. Sano T, Tsuchiya S, Suzuki F. A use of Adaline as an automatic method for interpretation of the electrocardiogram and the vectorcardiogram. *Jpn Circ J*. 1969;33:537-544.

18. Kostadinov S. Understanding backpropagation algorithm. In: *Towards Data Science*. 2019. https://towardsdatascience.com/understanding-backpropagation-algorithm-7bb3aa2f95fd. Accessed November 22, 2019.

19. Chan L, Beers K, Yau AA, et al. Natural language processing of electronic health records is superior to billing codes to identify symptom burden in hemodialysis patients. *Kidney Int*. 2020;97:383-392.

20. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med*. 2018;1:18.

21. Jang HJ, Cho KO. Applications of deep learning for the analysis of medical data. *Arch Pharm Res*. 2019;42:492-504.

22. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542:115-118.

23. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402-2410.

24. Kwon J-M, Jeon K-H, Kim HM, et al. Deep-learning-based out-of-hospital cardiac arrest prognostic system to predict clinical outcomes. *Resuscitation*. 2019;139:84-91.

25. Huang S, Cai N, Pacheco PP, et al. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genom Proteom*. 2018;15:41-51.

26. Yu KH, Kohane IS. Framing the challenges of artificial intelligence in medicine. *BMJ Qual Saf*. 2019;28:238-241.

27. Garbade MJ. Understanding K-means clustering in machine learning. In *Towards Data Science*. 2018. https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1. Accessed November 22, 2019.

28. Roman V. Unsupervised machine learning: clustering analysis. In: *Towards Data Science*; 2019. https://towardsdatascience.com/unsupervised-machine-learning-clustering-analysis-d40f2b34ae7e. Accessed November 2019.

29. Liu SH, Li Y, Liu B. Exploratory cluster analysis to identify patterns of chronic kidney disease in the 500 cities project. *Prev Chronic Dis*. 2018;15:E60.

30. Sands JJ, Etheredge GD, Shankar A, et al. Predicting hospitalization and mortality in end-stage renal disease (ESRD) patients using an Index of Coexisting Disease (ICED)-based risk stratification model. *Dis Manag*. 2006;9:224-235.

31. Ramspek CL, Voskamp PW, van Ittersum FJ, et al. Prediction models for the mortality risk in chronic dialysis patients: a systematic review and independent external validation study. *Clin Epidemiol*. 2017;9:451-464.

32. Goldstein BA, Chang TI, Mitani AA, et al. Near-term prediction of sudden cardiac death in older hemodialysis patients using electronic health records. *Clin J Am Soc Nephrol*. 2014;9:82-91.

33. Mezzatesta S, Torino C, Meo PD, et al. A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. *Comput Methods Programs Biomed*. 2019;177:9-15.

34. Kooman J, Usvyat L, Dekker M, et al. Cycles, arrows and turbulence: time patterns in renal disease, a path from epidemiology to personalized medicine? *Blood Purif*. 2019;47:171-184.

35. Chaudhuri S, Rosen S, Larkin J, et al. Effects of predictive modelling risk directed interventions on hospitalizations rates in haemodialysis patients [abstract MO031]. *Nephrol Dial Transplant*. 2020;35:149.

36. Long A, Blanchard T, Willetts J, et al. Using artificial intelligence to help predict imminent hospitalizations in patients with ESRD [abstract FR-PO791]. *J Am Soc Nephrol*. 2018;29:627.

37. Long A, Lindsey A, Willetts J, et al. Reducing hospitalizations with artificial intelligence and clinical decision support: lessons learned [abstract SA-PO967]. *J Am Soc Nephrol*. 2019;30:1014.

38. Tangri N, Grams ME, Levey AS, et al. Multinational assessment of accuracy of equations for predicting risk of kidney failure: a meta-analysis. *JAMA*. 2016;315:164-174.

39. Chauhan K, Nadkarni GN, Fleming F, et al. Initial validation of a machine learning-derived prognostic test (KidneyIntelX) integrating biomarkers and electronic health record data to predict longitudinal kidney outcomes. *Kidney360*. 2020;1:731-739.

40. Chan L, Nadkarni GN, Fleming F, et al. Derivation and validation of a machine learning risk score using biomarker and electronic patient data to predict rapid progression of diabetic kidney disease. *medRxiv*. 2020.

41. Xiao J, Ding R, Xu X, et al. Comparison and development of machine learning tools in the prediction of chronic kidney disease progression. *J Transl Med*. 2019;17:119.

42. Norouzi J, Yadollahpour A, Mirbagheri SA, et al. Predicting renal failure progression in chronic kidney disease using integrated intelligent fuzzy expert system. *Comput Math Methods Med*. 2016;2016:1-9.

43. Han H, Wang Y, Chaudhuri S, et al. Prediction of six months progression to end stage renal disease [abstract TH-PO887]. *J Am Soc Nephrol*. 2016;27:299A.

44. Jiao Y, Kopyt N, Bollu P, et al. Use of kidney disease progression model care planning report associates with lower dialysis catheter rates at the initiation of hemodialysis [abstract SA-PO840]. *J Am Soc Nephrol*. 2019;30:980.

45. Brier ME, Gaweda AE, Aronoff GR. Personalized anemia management and precision medicine in ESA and iron pharmacology in end-stage kidney disease. *Semin Nephrol*. 2018;38:410-417.

46. Barbieri C, Molina M, Ponce P, et al. An international observational study suggests that artificial intelligence for clinical decision support optimizes anemia management in hemodialysis patients. *Kidney Int*. 2016;90:422-429.

47. Schappacher-Tilp G, Cherif A, Fuertinger DH, et al. A mathematical model of parathyroid gland biology. *Physiol Rep*. 2019;7:e14045.

48. Schappacher-Tilp G, Fuertinger DH, Kotanko P. A multi-compartment model capturing the pharmacokinetics of the calcimimetic cinacalcet. *Cell Physiol Biochem*. 2019;53:429-438.

49. Schappacher-Tilp G, Hymes J, Fuertinger D, et al. Thrice weekly vs. daily cinacalcet: virtual clinical trial and its subsequent clinical validation in a large US hemodialysis population [abstract FR-PO130]. *J Am Soc Nephrol*. 2019;30:466.

50. Akizawa T, Kurita N, Mizobuchi M, et al. PTH-dependence of the effectiveness of cinacalcet in hemodialysis patients with secondary hyperparathyroidism. *Sci Rep*. 2016;6:19612.

51. Kalantar-Zadeh K, Shah A, Duong U, et al. Kidney bone disease and mortality in CKD: revisiting the role of vitamin D, calcimimetics, alkaline phosphatase, and minerals. *Kidney Int Suppl*. 2010;S10-S21.

52. Trebeschi S, Drago SG, Birkbak NJ, et al. Predicting response to cancer immunotherapy using non-invasive radiomic biomarkers. *Ann Oncol*. 2019;30:998-1004.

53. Huang C, Clayton EA, Matyunina LV, et al. Machine learning predicts individual cancer patient responses to therapeutic drugs with high accuracy. *Sci Rep*. 2018;8:16444.

54. Institute LHR. *Machine Learning Could Predict Medication Response in Patients with Complex Mood Disorders*. American Association for the Advancement of Science (AAAS); 2018. https://www.eurekalert.org/pub_releases/2018-08/lhri-mlc080818.php. Accessed October 20, 2019.

55. Makary MA, Daniel M. Medical error—the third leading cause of death in the US. *BMJ*. 2016;353:i2139.

56. Paredes M. *Can Artificial Intelligence Help Reduce Human Medical Errors? Two Examples from ICUs in the US and Peru*. Techpolicyinstitute; 2018. https://techpolicyinstitute.org/wp-content/uploads/2018/02/Paredes-Can-Artificial-Intelligence-help-reduce-human-medical-errors-DRAFT.pdf. Accessed November 22, 2019.

57. Kopec D, Kabir MH, Reinharth D, et al. Human errors in medical practice: systematic classification and reduction with automated information systems. *J Med Syst*. 2003;27:297-313.

58. Solomon M. *Beyond the Hype: Real Applications of Artificial Intelligence in Medication Management—Part 1*. Point of Care Partners; 2018. https://www.pocp.com/real-applications-artificial-intelligence-in-medication-management/. Accessed October 20, 2019.

59. Collins N. *Artificial Intelligence Helps Stanford Computer Scientists Predict the Side Effects of Millions of Drug Combinations*. Stanford News; 2018. https://news.stanford.edu/2018/07/10/ai-predicts-drug-pair-side-effects/. Accessed October 20, 2019.

60. Schiff GD, Volk LA, Volodarskaya M, et al. Screening for medication errors using an outlier detection system. *J Am Med Inform Assoc*. 2017;24:281-287.

61. Dekker MJE, Konings C, Canaud B, et al. Interactions between malnutrition, inflammation, and fluid overload and their associations with survival in prevalent hemodialysis patients. *J Ren Nutr*. 2018;28:435-444.

62. Hung S-C, Kuo K-L, Peng C-H, et al. Volume overload correlates with cardiovascular risk factors in patients with chronic kidney disease. *Kidney Int*. 2014;85:703-709.

63. Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation*. 2015;131:269-279.

64. Seymour CW, Kennedy JN, Wang S, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. *JAMA*. 2019;321:2003-2017.

65. Mukamal KJ, Siscovick DS, de Boer IH, et al. Metabolic clusters and outcomes in older adults: the cardiovascular health study. *J Am Geriatr Soc*. 2018;66:289-296.

66. Keane WF, Collins AJ. Influence of co-morbidity on mortality and morbidity in patients treated with hemodialysis. *Am J Kidney Dis*. 1994;24:1010-1018.

67. Khan IH, Catto G, MacLeod AM, et al. Influence of coexisting disease on survival on renal-replacement therapy. *Lancet*. 1993;341:415-418.

68. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40:373-383.

69. Beddhu S, Bruns FJ, Saul M, et al. A simple comorbidity scale predicts clinical outcomes and costs in dialysis patients. *Am J Med*. 2000;108:609-613.

70. Hemmelgarn BR, Manns BJ, Quan H, et al. Adapting the Charlson Comorbidity Index for use in patients with ESRD. *Am J Kidney Dis*. 2003;42:125-132.

71. Blanchard T. Using machine learning to detect errors in medical data. In: *Predictive Analytics World*; 2018. https://www.predictiveanalyticsworld.com/health/2018/agenda/#session52041. Accessed November 22, 2019.

72. Zhang H, Sor M, Koh E, et al. Automatic classification of arteriovenous fistula aneurysms using artificial intelligence [abstract TH-OR142]. *J Am Soc Nephrol*. 2019;30:40.

73. Bahado-Singh RO, Yilmaz A, Bisgin H, et al. Artificial intelligence and the analysis of multi-platform metabolomics data for the detection of intrauterine growth restriction. *PLoS One*. 2019;14:e0214121.

74. Zhang J, Friberg IM, Kift-Morgan A, et al. Machine-learning algorithms define pathogen-specific local immune fingerprints in peritoneal dialysis patients with bacterial infections. *Kidney Int*. 2017;92:179-191.

75. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22.

76. Team RaA. *Ethics Guidelines for Trustworthy AI*. European Commission; 2019. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. Accessed December 4, 2019.

77. US Food and Drug Administration. *Administration UFaD: Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)*. Services DoHaH; 2019. https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf.