Application of digital technology and artificial intelligence in nephrology

Application of Digital Technology and Artificial Intelligence in Nephrology

Sheetal Chaudhuri

Application of Digital Technology and Artificial Intelligence in Nephrology

Cover design: Glori Edewaard
Layout: Tiny Wouters
Production | Printing: ProefschriftMaken

ISBN: 978-94-6469-392-8

Application of Digital Technology and Artificial Intelligence in Nephrology

DISSERTATION

to obtain the degree of Doctor at the Maastricht University,
on the authority of the Rector Magnificus,
Prof.dr. Pamela Habibović
in accordance with the decision of the Board of Deans,
to be defended in public on
Thursday 22nd day of June 2023, at 13:00 hours

by

Sheetal Chaudhuri

Supervisor

Prof. dr. J.P. Kooman

Co-supervisors

Dr. L.A. Usvyat, Fresenius Medical Care, Global Medical Office, Waltham, USA
Dr. J.W. Larkin, Fresenius Medical Care, Global Medical Office, Waltham, USA

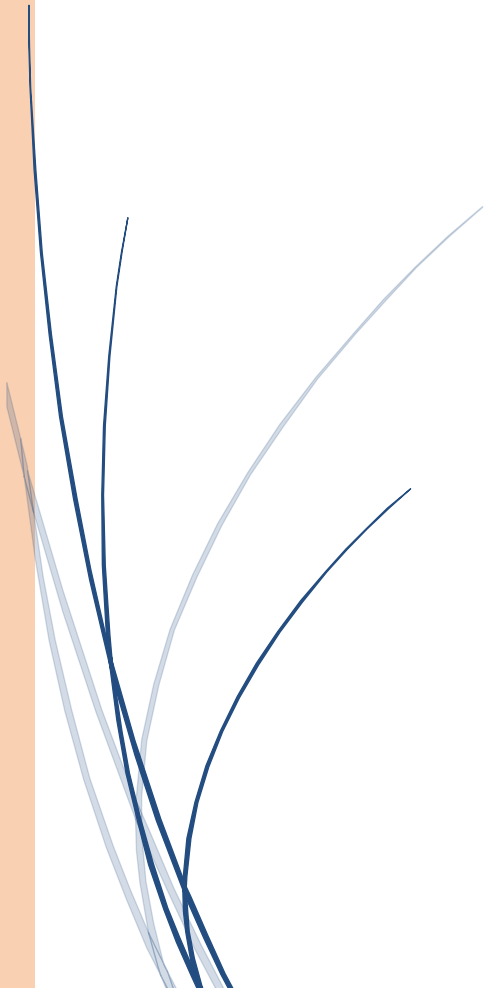
Assessment Committee

Prof. dr. A.A. Kroon (chair)
Prof. dr. A. Boonen
Prof. dr. D. Linz
Prof. dr. M. E. Diaz-Gonzalez de Ferris, University of North Carolina at Chapel Hill, USA
Prof. dr. W. Van Biesen, University Hospital Gent, Belgium

Table of contents

Chapter 1

Introduction



Introduction

Burden of chronic kidney disease

Chronic kidney disease (CKD) is a growing global health concern. CKD is described by the permanent decline in renal function, defined by the accumulation of uremic retention solutes in the blood and/or loss of albumin in the urine.[1,2] CKD is an important non-communicable disease that can be caused by primary abnormalities of the kidney or be secondary to diabetes, vascular diseases and is aggravated by hypertension. In 2017, the global prevalence of CKD had increased by 29.3% compared to 1990 with roughly 700 million cases globally. The increase in mortality from CKD since 1990 was 41.5%. CKD was the 12th leading cause of death globally in 2017.[3]

People who have CKD that progresses to End Stage Kidney Disease (ESKD) require renal replacement therapy (RRT) to remove excess fluid and uremic toxins from the body, to correct and as such to sustain life. RRTs comprise kidney transplantation or dialysis therapies, such as hemodialysis (HD), hemodiafiltration or peritoneal dialysis. Most people are treated with HD for their dialysis therapy. With HD, the blood is perfused through a dialyzer filter that consists of a semipermeable membrane to allow diffusive transport of excess body fluids and retention solutes from blood into dialysate fluids. Conversely, basic anions pass from the dialysate into the blood, neutralizing the acidic imbalance associated with the catabolism of proteins into uremic toxins.[4-6] Excess fluid is removed by ultrafiltration based on pressure difference between the blood and dialysate compartments. HD treatments are typically performed three times per week for 4 hours each session in an outpatient setting, although the length or frequency of the sessions can be individualized according to the needs of the patient. HD treatments can also be performed at home, often with a higher frequency of treatments (e.g. 3-7 days per week) and shorter treatment times.[7] Peritoneal dialysis (PD) is another modality that involves the routine administration of dialysate fluids into the peritoneum via a catheter. This dialysate dwells in the peritoneum to accumulate retention solutes and deliver base predominantly through diffusion. Fluid is removed by osmotic pressures, typically by addition of glucose to the dialysis fluid. After a dwell, the dialysate and excess fluids are flushed out of the body along with the absorbed toxins.[8] PD and home HD are common forms of modalities patients' can choose to perform in a remote setting outside of a dialysis clinic such as their home. Patients may have to change modalities, especially, patients on home modalities often transition to in-center modality due to technique failure. Patients who switched from home to in-center modality have shown significant reduction in quality of life.[9] Monitoring patients on home modality remotely is important but can be a clinical challenge. Remote digital technologies such as applications where patients can remotely report issues or use wearable devices so that clinical teams can track patients may play an important role in

addressing this challenge. Such remote digital technologies can expand clinicians reach through providing more active monitoring and creating more frequent touch points with patients.

The frequency of dialysis yields large amounts of clinical data on the patients, treatments, and outcomes associated with differing care paradigms. For instance, vital signs are commonly measured before, every 30 minutes during, and after each HD treatment performed multiple times per week. Fluid removal and other metrics can be captured from the HD machine every 10 seconds. This robust data could be used to define the risks associated with higher and lower values for vital signs, fluid removal and other negative outcomes among patients who frequently experience hypotension episodes during dialysis and fluid overload, compared to those who are more hemodynamically stable and achieve euvolemia. It is known that removing excess fluid from the body faster can alter cardiac stability and induce intradialytic hypotension due to rapid decline in blood volume, however identification of which patients are at a higher risk is an ongoing clinical dilemma.[10-12] Digital technologies and algorithms may be able to assist in early identification of patients at a higher risk who might need slower fluid removal rates to optimize their cardiac stability during fluid removal.[13]

Although dialysis therapies can help patients who progress to ESKD sustain life, they often have a poor quality of life (QOL) and have some of the poorest outcomes among all chronic diseases. For example, people with ESKD exhibit mortality rates that are higher than people with cancer, diabetes, heart failure, myocardial infarction, or stroke that do not have ESKD.[14-16] Global data from 2013 estimated ESKD is attributable to about 960,000 deaths per year.[17] Because of increase in morbidity burden, there is also higher than average hospitalization and rehospitalization rates in CKD and ESKD. Amongst all Medicare beneficiaries in the United States, 15.6% of people without CKD have a rehospitalization within 30 days; in CKD and ESKD, this rate is 21.5% and 32.0% respectively.[18] Overall, dialysis patients on average have about 1.8 hospitalizations per person year.[18] With large amounts of data collected in Electronic Medical Records (EMR), algorithms can be developed to assess risk of various outcomes (e.g., hospitalization and mortality).

With the Coronavirus Disease (COVID-19) pandemic, patients suffering from ESKD had some of the worst outcomes of any chronic disease.[19] This might be since many of these patients are older and have weakened immune system, in addition to the chronic disease state of ESKD. Receiving HD in a dialysis clinic may also increase the risk of exposing the patient to SARS-COV-2 virus.[20] Studies have shown the excess mortality in ESKD patients during the pandemic was 30% higher among patients on dialysis in US and UK compared to all-cause mortality in the prior years.[21] Studying the trajectories of clinical parameters and identifying high risk patients can likely help in curbing the

spread of the communicable disease and may be able to improve QOL for ESKD patients.

In summary, ESKD patients are unique and may experience unique challenges through their journey as CKD and ESKD patient. Having the ability to detect small signals in from the data already being routinely collected in EMR may have the potential to help in providing proactive care for such patients.

Big data and artificial intelligence in healthcare

Health care is entering a new era where the advent of "big data" brings tremendous opportunities to revolutionize the field. Digital Technologies enable generation, storage, and processing of large volumes of data. In order to be purposeful, data must be analyzed, interpreted, and used to improve patient care. Exploring the associations among different pieces of information derived from large and diverse datasets to enable intelligent and informed decisions is now possible with the emergence of artificial intelligence (AI).[22]

AI is "the science and engineering of making intelligent machines, especially intelligent computer programs" as described by the pioneer Alan Turing in 1950.[23] AI is considered the fourth generation of the industrial revolution, since it is anticipated to transform and impact every possible sector of our lives, just like electricity did in the early 1900s (**Figure 1.1**).

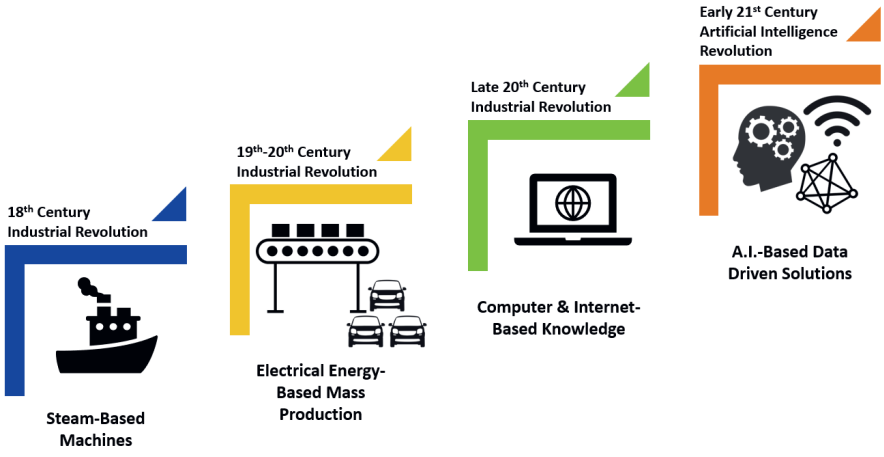


Figure 1.1 Phases of the industrial revolution: First phase of the Industrial Revolution shows the introduction of mechanical production using steam powered engines. Second Phase of the revolution shows the introduction of mass production of goods using electricity. Third Phase shows further automation of production using Information and Technology and the fourth phase of the revolution shows the automation using AI techniques.

AI in kidney care, and health care in general, uses algorithms and software engineering principles to approximate the decisions made by clinicians in the analysis of complex health care data. Traditionally computer-based algorithms in health care include a set of rules encoding expert knowledge on medical decisions. These rules are subsequently applied to draw conclusions about specific clinical scenarios. AI algorithms, however, strive to learn from the data without concrete rules.

AI is an umbrella term that brings together concepts from several fields such as statistics, algorithmics, machine learning, information retrieval, and data science at large. Machine learning (ML), a sub-discipline of AI, is the scientific study of algorithms and statistical models that can learn how to perform a specific task without using explicit instructions by relying on patterns and inference instead. Consequently, ML algorithms are highly "data hungry," often requiring thousands of observations to reach acceptable performance.[24]

The vast amounts of data collected in EMR at the point of care provide a rich platform to employ ML. ML thrives in handling enormous numbers of predictors and can combine them in nonlinear and highly interactive ways. This capacity allows for employing new kinds of data (e.g. text, images, sound, temporal data), whose sheer volume or complexity would previously have made analyzing them untenable.

Deep learning (DL) is another subset of ML that uses multi-layered artificial neural networks (ANN) to create more advanced non-linear feature engineering than traditional ML methods.[25]

In a recent bibliometric study on the global evolution of research of AI in health care and medicine, it is evident that AI techniques and clinical applications of AI are relatively more common in fields like ophthalmology, oncology, and cardiology.[26] It is less prevalent in nephrology, even though the daily disease burden for CKD and ESKD patients is higher compared to the other diseases.[27]

In ophthalmology, multiple AI-based grading algorithms have been developed to screen fundus photographs obtained from diabetic patients and identify who should be referred to the ophthalmologist for treatment. DL models developed had 94% to 98% sensitivity and 93% to 98% specificity across differing models.[28,29]

The first AI based device that was approved and permitted for marketing by the Food and Drug Administration (FDA) was the IDx-DR (IDx Technologies Inc., Coralville, IA), which combines DL image recognition software to analyze retinal images and provide recommendations to refer patients for evaluation of diabetic retinopathy.[30] In oncology, AI has helped with tumor imaging, pathology, and clinical decision making.[31]

AI has also entered the lives of patients in the form of wearable devices to monitor patients remotely and make personalized recommendations. A current health's AI-enabled wearable device measures multiple vital signs and has recently received FDA authorization for patients to use at home.[32] The device can measure pulse, respiration, oxygen saturation, temperature, and mobility. The device provides patients and physicians real-time updates and alerts on vital signs, which empowers patients in their care and allows physicians to address complications promptly. This technology uses ML to analyze continuous vital sign and kinetic data to detect unusual trends. Currently, this device is mostly used by patients with chronic obstructive pulmonary disease (COPD) and heart failure. AI-enabled remote monitoring may help improve hospitalization and mortality in such patients.[33]

Rise of digital technologies in healthcare

Digital technologies include hardware and software applications and services such as telemedicine, mobile phone applications, text messaging, monitoring health via wearable devices and cloud based real time applications and analytics. There have been significant enhancements in all areas of digital technologies for use in the field of health care[34,35]. This is critical in the development of interconnected health systems to aid healthcare professionals and patients in managing health. Digital health is the convergence of all these digital technologies to provide personalized care and enhance effective delivery of health care services. Cloud computing in healthcare offers increased scalability to store big data from EMR and to develop/execute applications in real time[36].

Application of AI and digital technology in nephrology

In the field of nephrology, various initiatives have incorporated ML. One such example is with a ML model developed to predict CKD progression. The model was developed and validated using demographic, clinical, and most recent laboratory data from two independent Canadian cohorts of patients with CKD stages 3 to 5.[37] This CKD progression model has been validated externally and in international populations,[38,39]and is currently used for triage of CKD care in Canada.[40,41] Other studies have developed ML models to accurately predict outcomes and graft survival after kidney transplant.[42,43] ML models have just begun to be applied in health care and they have the potential to improve patient care paradigms when implemented at the point of care in clinical applications in nephrology.

As part of this dissertation, we present several applications related to AI and digital technologies that demonstrate potential and impact in addressing several challenges within Nephrology.

In **chapter 2**, we begin by reviewing AI related concepts and present several different applications of AI in Kidney Disease. We provide an overview of the AI application process in a clinical setting and provide brief descriptions of select advanced ML algorithms. We also present the current state of AI in research towards kidney disease and dialysis and explore future pathways for AI within the discipline of nephrology. The review paper introduces most of the concepts used in the rest of the following chapters of this thesis.

In **chapter 3**, historic patient data from an integrated kidney disease company is leveraged to direct care in quality improvement efforts at its national network of dialysis clinics. The provider has developed and operationally deployed a set of ML models that identify in-center HD patients at an increased risk for multiple all-cause hospitalizations within the next 12 months. The models were used in a pilot called the Dialysis Hospitalization Reduction Program (DHRP) that provides risk directed interdisciplinary team root cause evaluations and personalized interventions to HD patients predicted to be at risk of ≥6 hospital admissions within the next 12 months.

We investigated the impact of the DHRP on clinic-wide hospitalization rates in HD patients. Moreover, considering other analysis supporting the use of behavioral health interventions, we explored profiles of psychosocial barriers that included patient reported outcomes for depression, sleep, and psychological stress status related to one of the DHRP interventions that was explicitly recorded in the EMR.[44]

In **chapter 4**, a ML application is presented to identify patients at risk of having their relative blood volume (RBV) decrease at a rate of at least -6.5% per hour anytime during HD. HD involves removal of fluid from the circulating blood by ultrafiltration and refilling from the extravascular compartments.[45] This process helps preserve blood pressure and tissue perfusion.[46] However, inadequate refilling could lead to a variety of intradialytic symptoms, such as intradialytic hypotension (IDH), fatigue, and cramping.[47,48] IDH can lead to cardiac complications and an increased risk of death.[49-52]

Studies have shown the role of RBV and how adapting the ultrafiltration rate has a positive effect on intradialytic symptoms.[53,54] However, it has been clinically challenging to identify changes in RBV in real time to proactively intervene and reduce potential negative consequences of volume depletion. Hence leveraging advanced technologies to process large volumes of dialysis and machine data in real time and developing prediction models using machine learning (ML) is critical in identifying these signals.

In **chapter 5**, we assess the associations between the level of utilization of the Remote treatment monitoring (RTM) application and hospitalization and technique failure rates to actively identify urgent concerns in PD patients. The modality of PD is suggested to

associate with favorable outcomes compared to in-center HD, however, technique failure is common and adjusted rates for the hospital length of stay tends to be longer in PD.[27,55,56] Remote monitoring systems may improve the care team's ability to actively identify urgent concerns in PD patients and react in a timely manner with diagnostic examinations, interventions, or patient education.[57] A large integrated kidney disease healthcare company started using a RTM application throughout its dialysis organization in the United States since October 2016. The provider constructed and integrated the RTM system into its clinical systems as a quality improvement process. The RTM is an online portal-based treatment record application, whereby patients can create an online account and record the details of individual PD treatments, associated vital signs, and complications. The data from this RTM application are used as part of this study.

In **chapter 6**, we evaluate the trajectories of clinical and laboratory parameters in HD patients due to the COVID-19 pandemic. Dialysis patients appear to be at increased risk for viral transmission with relatively high mortality rates ranging from 11% to 30%.[58-63] During the first half of 2020, there were over 11,200 COVID-19 hospitalizations among Medicare beneficiaries undergoing dialysis in the United States.[64] Various parameters such as pulse, body temperature, C-reactive protein (CRP) and lymphocyte counts at presentation were found to be associated to COVID-19 mortality in kidney failure.[59]

Albeit the clinical presentations in COVID-19 have been somewhat established, the changes in clinical parameters before presentation that characterize disease onset in humans are unknown secondary to a scarcity of longitudinal data available in the general population or collected in registries in the kidney failure population. HD patients have robust routine data collected in EMR affording the opportunity to define the pathophysiological disturbances characterizing the onset and course of COVID-19 in kidney failure patients. The goal of this analysis was to compare trends in clinical and laboratory parameters between HD patients who tested positive or negative for SARS CoV-2. The second goal of this study was to compare clinical trends between survivors and non-survivors who were diagnosed with COVID-19.

In **chapter 7,** we use our findings from **chapter 6** to develop a ML model that predicts the risk of a HD patient having an undetected SARS-CoV-2 infection that is identified after the following 3 or more days. The model was built as part of a healthcare operations effort in response to the COVID-19 outbreak.

In **chapter 8**, we developed ML models to predict mortality in HD patients using a large globally representative dialysis Monitoring Dialysis Outcome (MONDO) database examining various input parameters from the nutritional, inflammatory, hydration, anemia, and mineral metabolism domains. In this effort, we assessed advanced

analytical ML methods as well as traditional statistical methods to highlight the advantages and disadvantages of each modeling technique for mortality prediction.

We conclude the dissertation with a discussion of the ability, implication, and challenges of AI in its applicability in dialysis care.

References

1. Levey AS, de Jong PE, Coresh J, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. Kidney Int 2011;80:17-28.
2. Inker LA, Astor BC, Fox CH, et al. KDOQI US commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD. Am J Kidney Dis 2014;63:713-735.
3. Cockwell P, Fisher LA. The global burden of chronic kidney disease. Lancet 2020;395:662-664.
4. Hoenich NA, Levin R, Ronco C. How do changes in water quality and dialysate composition affect clinical outcomes? Blood Purif 2009;27:11-15.
5. Smith WB, Gibson S, Newman GE, et al. The dynamics of the metabolism of acetate and bicarbonate associated with use of hemodialysates in the ABChD trial: a phase IV, prospective, single center, single blind, randomized, cross-over, two week investigation. BMC Nephrol 2017;18:273.
6. Gottschalk CW, Fellner SK. History of the science of dialysis. Am J Nephrol 1997;17:289-298.
7. Chan CT, Blankestijn PJ, Dember LM, et al. Dialysis initiation, modality choice, access, and prescription: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. Kidney Int 2019;96:37-47.
8. Morelle J, Devuyst O. Water and solute transport across the peritoneal membrane. Curr Opin Nephrol Hypertens 2015;24:434-443.
9. Eneanya ND, Maddux DW, Reviriego-Mendoza MM, et al. Longitudinal patterns of health-related quality of life and dialysis modality: a national cohort study. BMC Nephrol 2019;20:7.
10. Guedes M, Pecoits-Filho R, Leme JEG, et al. Impacts of dialysis adequacy and intradialytic hypotension on changes in dialysis recovery time. BMC Nephrol 2020;21:529.
11. Larkin JW, Reviriego-Mendoza MM, Usvyat LA, et al. To cool, or too cool: Is reducing dialysate temperature the optimal approach to preventing intradialytic hypotension? Semin Dial 2017;30: 501-508.
12. Flythe JE, Xue H, Lynch KE, et al. Association of mortality risk with various definitions of intradialytic hypotension. J Am Soc Nephrol 2015;26:724-734.
13. Kooman JP, Wieringa FP, Han M, et al. Wearable health devices and personal area networks: can they improve outcomes in haemodialysis patients? Nephrol Dial Transplant 2020;35:ii43-ii50.
14. Sesso RC, Lopes AA, Thome FS, et al. Brazilian Chronic Dialysis Census 2014. J Bras Nefrol 2016;38: 54-61.
15. Saran R, Robinson B, Abbott KC, et al. US Renal Data System 2017 Annual Data Report: Epidemiology of Kidney Disease in the United States. Am J Kidney Dis 2018;71:A7.
16. Luyckx VA, Tonelli M, Stanifer JW. The global burden of kidney disease and the sustainable development goals. Bull World Health Organ 2018;96:414-422D.
17. Sharma S, Sarnak MJ. Epidemiology: The global burden of reduced GFR: ESRD, CVD and mortality. Nat Rev Nephrol 2017;13:447-448.
18. USRDS. ESRD Hospitalization. USRDS Annual Data Rport 2021; https://adr.usrds.org/2021/end-stage-renal-disease/5-hospitalization.
19. Weinhandl ED, Wetmore JB, Peng Y, et al. Initial Effects of COVID-19 on Patients with ESKD. J Am Soc Nephrol 2021;32:1444-1453.
20. Robinson BM, Guedes M, Alghonaim M, et al. Worldwide Early Impact of COVID-19 on Dialysis Patients and Staff and Lessons Learned: A DOPPS Roundtable Discussion. Kidney Med 2021;3:619-634.
21. Mahalingasivam V, Su G, Iwagami M, et al. COVID-19 and kidney disease: insights from epidemiology to inform clinical practice. Nat Rev Nephrol 2022;18(8):485-498.
22. 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.
23. TURING AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. Mind 1950;LIX:433-460.
24. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. N Engl J Med 2016;375:1216-1219.

25. Gupta T. Deep Learning: Feedforward Neural Network. Towards Data Science 2017; https://towardsdatascience.com/deep-learning-feedforward-neural-network-26a6705dbdc7.
26. Tran BX, Vu GT, Ha GH, et al. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. J Clin Med 2019;8(3):360.
27. Saran R, Robinson B, Abbott KC, et al. US Renal Data System 2018 Annual Data Report: Epidemiology of Kidney Disease in the United States. Am J Kidney Dis 2019;73:A7-A8.
28. Gargeya R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning. Ophthalmology 2017;124:962-969.
29. 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.
30. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. FDA NEWS RELEASE 2018; https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye.
31. Kann BH, Thompson R, Thomas CR, Jr., et al. Artificial Intelligence in Oncology: Current Applications and Future Directions. Oncology (Williston Park) 2019;33:46-53.
32. Current Health Receives FDA Clearance for its Remote Patient Monitoring Solution, Enables Earlier Intervention and Improved Patient Outcomes. BUSINESS WIRE 2019; https://www.businesswire.com/\news/home/20190206005175/en/Current-Health-Receives-FDA-Clearance-Remote-Patient.
33. Current remote patient monitoring reduces readmissions at Dartford and Gravesham NHS Trust. Health Tech Pulse 2019; https://healthtechpulse.com/2019/01/30/current-remote-patient-monitoring-reduces-readmissions-dartford-and-gravesham-nhs-trust.
34. Yoeli E, Rathauser J, Bhanot SP, et al. Digital Health Support in Treatment for Tuberculosis. N Engl J Med 2019;381:986-987.
35. Rizas KD, Freyer L, Sappler N, et al. Smartphone-based screening for atrial fibrillation: a pragmatic randomized clinical trial. Nat Med 2022;28:1823-1830.
36. Griebel L, Prokosch HU, Kopcke F, et al. A scoping review of cloud computing in healthcare. BMC Med Inform Decis Mak 2015;15:17.
37. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. JAMA 2011;305:1553-1559.
38. Lennartz CS, Pickering JW, Seiler-Mussler S, et al. External Validation of the Kidney Failure Risk Equation and Re-Calibration with Addition of Ultrasound Parameters. Clin J Am Soc Nephrol 2016;11:609-615.
39. 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.
40. Hingwala J, Wojciechowski P, Hiebert B, et al. Risk-Based Triage for Nephrology Referrals Using the Kidney Failure Risk Equation. Can J Kidney Health Dis 2017;4:2054358117722782.
41. Smekal MD, Tam-Tham H, Finlay J, et al. Patient and provider experience and perspectives of a risk-based approach to multidisciplinary chronic kidney disease care: a mixed methods study. BMC Nephrol 2019;20:110.
42. Molnar MZ, Nguyen DV, Chen Y, et al. Predictive Score for Posttransplantation Outcomes. Transplantation 2017;101:1353-1364.
43. Yoo KD, Noh J, Lee H, et al. A Machine Learning Approach Using Survival Statistics to Predict Graft Survival in Kidney Transplant Recipients: A Multicenter Cohort Study. Sci Rep 2017;7:8904.
44. Nien Chen L, Stephanie Johnstone, Felicia N. Speed et al. Associations between Sleep Quality and Quality of Life, Stress, and Depressive Symptoms in Hemodialysis Patients J Am Soc Nephrol 2016;27: 311A.
45. Schneditz D, Roob J, Oswald M, et al. Nature and rate of vascular refilling during hemodialysis and ultrafiltration. Kidney Int 1992;42:1425-1433.
46. Daugirdas JT. Dialysis hypotension: a hemodynamic analysis. Kidney Int 1991;39:233-246.
47. Alvarez L, Brown D, Hu D, et al. Intradialytic Symptoms and Recovery Time in Patients on Thrice-Weekly In-Center Hemodialysis: A Cross-sectional Online Survey. Kidney Med 2020;2:125-130.
48. Kuipers J, Verboom LM, Ipema KJR, et al. The Prevalence of Intradialytic Hypotension in Patients on Conventional Hemodialysis: A Systematic Review with Meta-Analysis. Am J Nephrol 2019;49:497-506.

49. Morfin JA, Fluck RJ, Weinhandl ED, et al. Intensive Hemodialysis and Treatment Complications and Tolerability. Am J Kidney Dis 2016;68:S43-S50.
50. Gul A, Miskulin D, Harford A, et al. Intradialytic hypotension. Curr Opin Nephrol Hypertens 2016;25: 545-550.
51. Sars B, van der Sande FM, Kooman JP. Intradialytic Hypotension: Mechanisms and Outcome. Blood Purif 2020;49:158-167.
52. Keane DF, Raimann JG, Zhang H, et al. The time of onset of intradialytic hypotension during a hemodialysis session associates with clinical parameters and mortality. Kidney Int 2021;99:1408-1417.
53. Steuer RR, Leypoldt JK, Cheung AK, et al. Reducing symptoms during hemodialysis by continuously monitoring the hematocrit. Am J Kidney Dis 1996;27:525-532.
54. Dasselaar JJ, Huisman RM, de Jong PE, et al. Measurement of relative blood volume changes during haemodialysis: merits and limitations. Nephrol Dial Transplant 2005;20:2043-2049.
55. Lukowsky LR, Mehrotra R, Kheifets L, et al. Comparing mortality of peritoneal and hemodialysis patients in the first 2 years of dialysis therapy: a marginal structural model analysis. Clin J Am Soc Nephrol 2013; 8:619-628.
56. Snyder JJ, Foley RN, Gilbertson DT, et al. Body size and outcomes on peritoneal dialysis in the United States. Kidney Int 2003;64:1838-1844.
57. Rosner MH, Lew SQ, Conway P, et al. Perspectives from the Kidney Health Initiative on Advancing Technologies to Facilitate Remote Monitoring of Patient Self-Care in RRT. Clin J Am Soc Nephrol 2017; 12:1900-1909.
58. Corbett RW, Blakey S, Nitsch D, et al. Epidemiology of COVID-19 in an Urban Dialysis Center. J Am Soc Nephrol 2020;31:1815-1823.
59. Hilbrands LB, Duivenvoorden R, Vart P, et al. COVID-19-related mortality in kidney transplant and dialysis patients: results of the ERACODA collaboration. Nephrol Dial Transplant 2020;35:1973-1983.
60. Jager KJ, Kramer A, Chesnaye NC, et al. Results from the ERA-EDTA Registry indicate a high mortality due to COVID-19 in dialysis patients and kidney transplant recipients across Europe. Kidney Int 2020;98(6):1540-1548.
61. Medicine JHUSo. Mortality Analyses. Johns Hopkins Coronavirus Resource Center 2020; (Accessed December 2020) https://coronavirus.jhu.edu/data/mortality.
62. Neumann ME. Latest data show 305 dialysis patient deaths due to COVID-19 in the US. Nephrology News & Issues 2020; (Accessed 22 Apr 2020) https://www.healio.com/nephrology/infection-control/news/online/%7B3a263aa9-ad59-4c3f-aab7-07b8395508e5%7D/latest-data-show-305-dialysis-patient-deaths-due-to-covid-19-in-the-us.
63. Ng JH, Hirsch JS, Wanchoo R, et al. Outcomes of patients with end-stage kidney disease hospitalized with COVID-19. Kidney Int 2020;98(6):1530-1539.
64. USRDS. COVID-19 Supplement. 2020; (Accessed November 2020) https://adr.usrds.org/2020/covid-19-supplement/1-covid-19-supplement.

Chapter 2

Applications of Artificial Intelligence (AI) in Kidney Disease

Sheetal Chaudhuri, Andrew Long, Hanjie Zhang,
Caitlin Monaghan, John W Larkin, Peter Kotanko,
Shashi Kalaskar, Jeroen P Kooman, Frank M van der Sande,
Franklin W Maddux, Len A Usvyat

Semin Dial. 2021;34(1):5-16

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 healthcare. 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 towards 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 3 main topics: (1) predicting events in the future such as mortality and hospitalization (2) providing treatment and decision aids such as automating drug prescription (3) identifying patterns such as phenotypical clusters and arteriovenous fistula aneurysm (AVFA). 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 should not replace the Nephrologists' medical decision making, but instead assist them in providing optimal personalized care for their patients.

Introduction

Artificial intelligence (AI) is anticipated to transform healthcare 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 (e.g. 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 healthcare 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: 1) provide an overview of the AI application process in a clinical setting, 2) provide brief descriptions of select advanced machine learning algorithms, 3) present the current status of AI in research towards kidney disease and dialysis and 4) 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 and dialysis omitting the unique acute kidney injury population.

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 2.1**, AI is an umbrella term that brings together concepts from several fields such as computer science, statistics, algorithmic, machine learning (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.

Figure 2.1 The relationship between AI, 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.

The types of ML techniques that currently exist for building AI applications broadly fall into three families (**Figure 2.2**), 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.

Supervised Learning (SL)

SL 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).

SL can be divided into two categories (classification and regression) depending on the type of the output (**Figure 2.2**). In classification, the output belongs to a set of distinct classes (e.g. missed treatment vs. not missed treatment). In regression, the output is usually a continuous numerical quantity (e.g. N days until transitioning to dialysis).

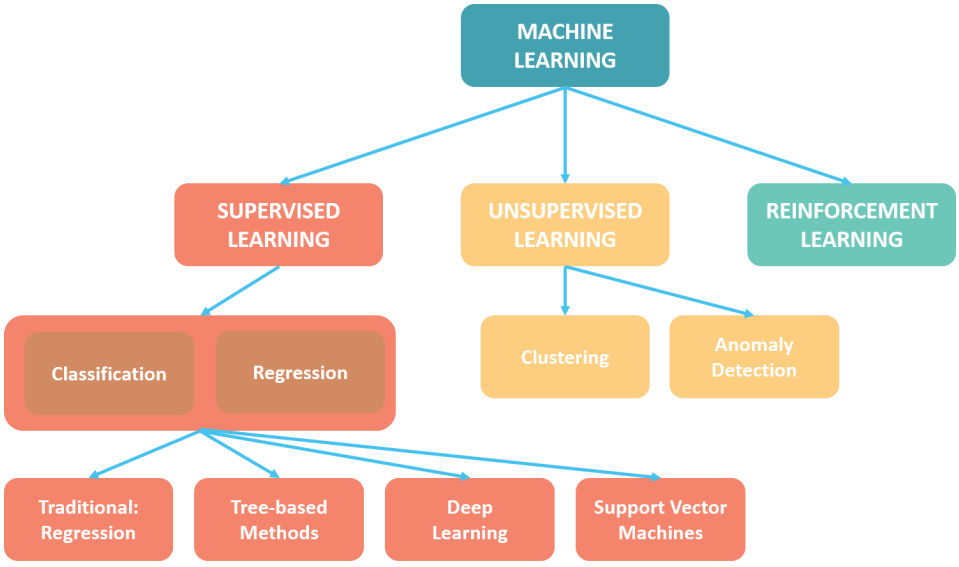


Figure 2.2 Supervised Learning (SL) and Unsupervised Learning (UL) are the two main categories of Machine Learning (ML). Deep Learning (DL) is a subset of ML. SL algorithms are used to learn the optimal parameters of the predictive model by investigating past examples with known inputs and known outputs. UL algorithms learn about patterns in the input data itself and does not have a known output.

There are many ML algorithms for building predictive models ranging from traditional to more advanced methods. Prediction performance of these models are 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 healthcare and hence not discussed here for brevity. Over the past

decade, more advanced techniques such as tree-based methods and deep learning 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 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 (e.g. 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 (i.e. by bootstrapping) and random subsets of the input features (i.e. 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 machine learning field.[15]

An extensive bibliography of new SL techniques, their application, and performance compared to traditional techniques are 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 (DL), which uses artificial neural networks (ANN), 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 wasn't 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, call neuron (**Figure 2.3**). ANN's neurons usually receive multiple inputs that are mathematically combined through non-linear (e.g. sigmoidal) activation functions F(x). 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 2.4**). 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 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 (i.e. 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.

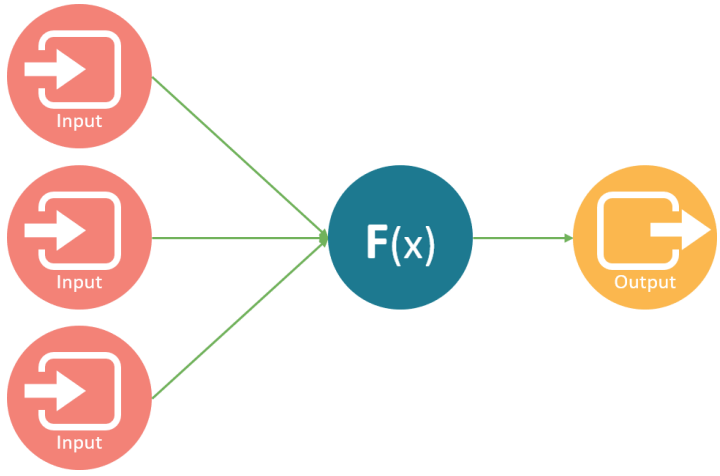


Figure 2.3 A very simple artificial neural network (ANN) with an input layer comprised of 3 inputs, hidden layer comprised of 1 neuron, and the output layer. ANN's neuron usually combines input from multiple sources through non-linear activation functions F(x).

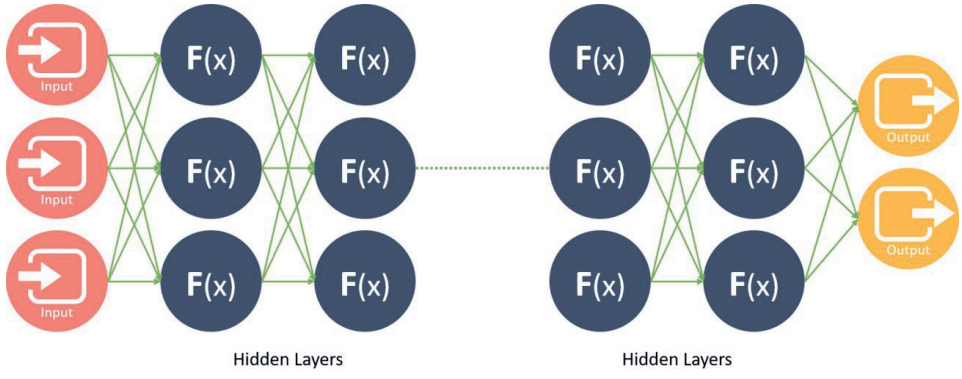


Figure 2.4 Deep Learning Network. Input layer with 3 inputs, multiple hidden layers of neurons and 2 output layers. Higher the number of hidden layers deeper is the network.

To optimize these weights, the DL algorithm uses a technique known as Back Propagation which was invented in the 1980s.[18] As the number of layers are added to the neural network, the number of weights and connections increases dramatically. Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) as shown in **Figure 2.5** and **Figure 2.6** 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]

Figure 2.5 Convolution Neural Network (CNN) is a class of Deep Learning (DL) 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.

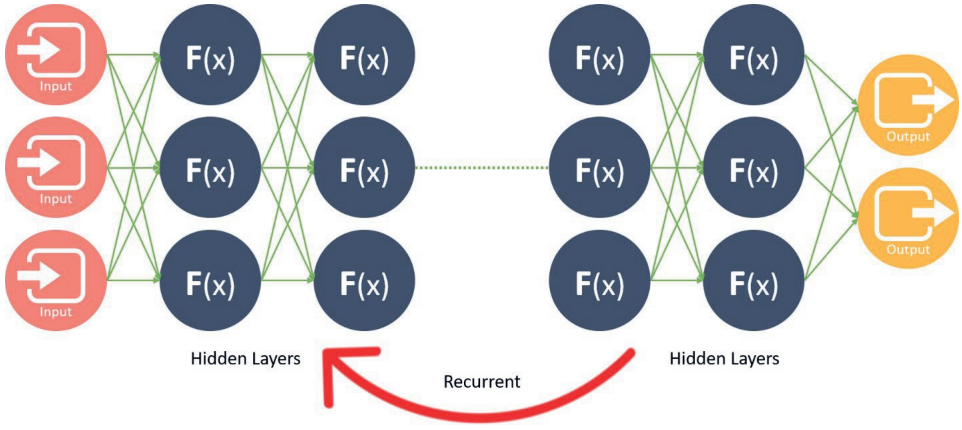


Figure 2.6 Recurrent Neural Network (RNN). In RNN the output from the function are fed back in the model in order to minimize error.

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 supervised learning, 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]

Un-Supervised 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, the algorithm aims to minimize the within-cluster variance and maximize the between-cluster variance. It is critical to determine an appropriate number of clusters k when using k-means clustering method.

Hierarchical clustering[28] is another commonly used clustering technique that 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 socio-demographic factors and other risk factors could contribute to the prevalence of CKD.

Table 2.1. shows a very high level overview of the differences between the traditional statistical techniques and advanced analytical methods.

Table 2.1 Differences between traditional statistical methods versus advanced analytical techniques.

Factors	Traditional statistical techniques	Advanced analytical techniques
Training Data	Works with Smaller Data Sets	Better with Large Data Sets
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

AI Application Process

The AI application process in a clinical setting generally consists of a series of stages (**Figure 2.7**). 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 hemodialysis, the enormous amounts of EMR data collected at the point of care provides 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 (e.g. free text, images, videos, sound, temporal data) to be utilized. The volume and complexity of such data adds additional challenges in analyzing the data.

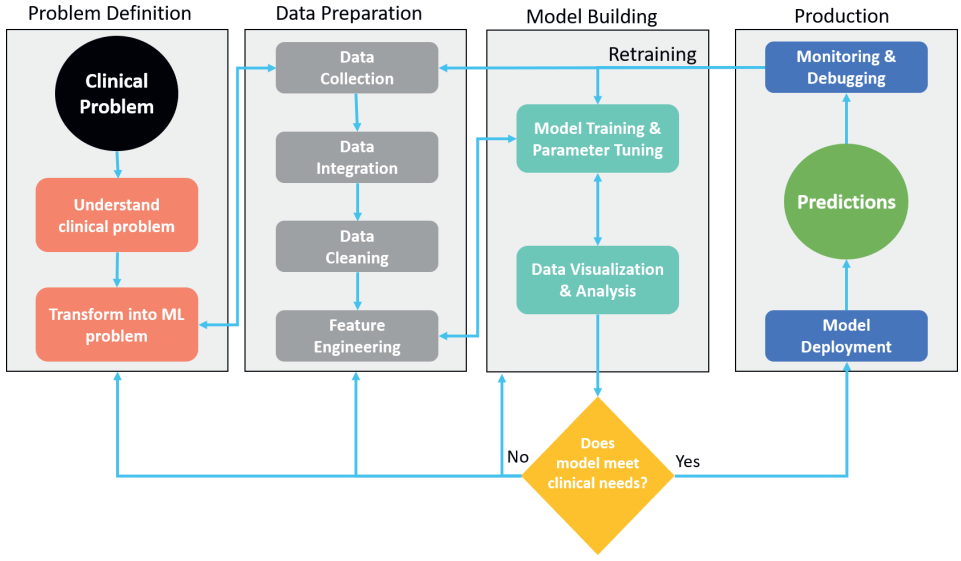


Figure 2.7 Process for Application of AI with 4 phases: Problem Definition, Data Preparation, Model Building and Production.

With a set of well-engineered features, the predictive models are trained and tuned until acceptable performance is achieved. It is anticipated some steps of the process have bi-directional arrows because they can result in modifications for previous steps (**Figure 2.7**). 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.

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 3 main topics: (1) predicting events in the future, (2) treatment and decision aids and (3) identification of existing, but unrecognized, patterns. **Table 2.2** 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 the scientific community has more contributions on the horizon.

Table 2.2 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 BA 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 J 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 C et al.[46] 2016	Artificial neural networks	752	Anemia management	MAE:0.59 g/dL	Yes
Zhang J et al.[74] 2017	Random forest	83	Immune fingerprints	AUROC: 0.993 Sensitivity:98.5% Specificity:92.6%	NA

AUROC: Area Under the Receiver Operating Curve, MSE: Mean Square Error; MAE: Mean Absolute Error; NMSE: Normalized MSE; PPV: Positive Predicted Value; NPV: Negative Predicted Value.

Predicting the future

Predicting outcomes

Patients with end stage kidney disease (ESKD) have high mortality as well as hospitalization rates.[7] Prediction Models can assist with early care planning and triaging resources where there is the 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 hemodialysis (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 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. Dialysis 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 (e.g. 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 are 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.

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 is critical for positive patient outcomes and controlling healthcare 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 to 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/day 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 eGFR in 6, 12 and 18th 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 estimated glomerular filtration rates (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 (CDSS). 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(CVC).[44]

Treatment aid

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 reduced 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 3x 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 3x weekly in-center administration of cinacalcet is non-inferior 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]

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 U.S.; in 2016, they contributed to more than 251K deaths in the U.S. alone and accrued $17.8 billion dollars in unnecessary spend.[55,56] Different causes of medical errors exist such as (a) complexity of the healthcare 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 healthcare 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 U.S. Intensive Care Units(ICU) 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 supports 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]

Identifying patterns

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 showed 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 6-fold increase in mortality risk. However, unlike other chronic diseases, pattern detection based on UL techniques have not yet been published in nephrology. In patients with heart failure with preserved ejection fraction (HFpEF), 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 from 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 end points 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.

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 highlights 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, natural language processing of physician notes, and demographic information, the LDO was able to improve coding over the previous method of randomly chosen manual medical record reviews.

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 sec 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 (LC-MS) 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 data set where traditional statistical methods fail to yield satisfactory results.

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 decisions 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 include 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 into developing powerful and accurate models, more emphasis should be directed towards 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 should not replace providers' medical decision making, but instead assist them in providing optimal personalized care for their patients.

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 BX, Vu GT, Ha GH, 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-Wosten 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.
11. Niel O, Bastard P. Artificial Intelligence in Nephrology: Core Concepts, Clinical Applications, and Perspectives. Am J Kidney Dis 2019;74(6):803-810.
12. Narkhede S. Understanding AUC - ROC Curve. Towards Data Science 2018; (Accessed November 2019): https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5.
13. Alto V. Understanding the OLS method for Simple Linear Regression. Towards Data Science 2019; (Accessed November 2019) : https://towardsdatascience.com/understanding-the-ols- method-for-simple-linear-regression-e0a4e8f692cc.
14. Chen L. Basic Ensemble Learning (Random Forest, AdaBoost, Gradient Boosting)- Step by Step Explained. Towards Data Science 2019; (Accessed April 2020): https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient- boosting-step-by-step-explained-95d49d1e2725.
15. Morde V. XGBoost Algorithm: Long May She Reign. Towards Data Science 2019; (Accessed April 2020): https://towardsdatascience.com/https-medium-com-vishalmorde-xgboost-algorithm-long-she-may-rein-edd9f99be63d.
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. Towards Data Science 2019; (Accessed November 2019): https://towardsdatascience.com/understanding- backpropagation-algorithm-7bb3aa2f95fd.
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 JM, Jeon KH, 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 Genomics Proteomics 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. Towards Data Science 2018; (Accessed November 2019): https://towardsdatascience.com/understanding-k-means- clustering-in-machine-learning-6a6e67336aa1.
28. Roman V. Unsupervised Machine Learning: Clustering Analysis. Towards Data Science 2019; (Accessed November 2019): https://towardsdatascience.com/unsupervised-machine-learning- clustering-analysis-d40f2b34ae7e.
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 P, 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 JP, Usvyat LA, Dekker MJE, 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]. Nephrology Dialysis Transplantation 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: 10.34067/KID.0002252020.
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: 2020.2006.2001.20119552.
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: 6080814.
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:S10-21.
52. Trebeschi S, Drago SG, Birkbak NJ, et al. Predicting Response to Cancer Immunotherapy using Non-invasive Radiomic Biomarkers. Ann Oncol 2019;30(6):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; (Accessed October 2019): https://www.eurekalert.org/pub_releases/2018-08/lhri-mlc080818.php.
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; (Accessed November 2019): https://techpolicyinstitute.org/wp-content/uploads/2018/02/Paredes-Can-Artificial- Intelligence-help-reduce-human-medical-errors-DRAFT.pdf.
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; (Accessed October 2019): https://www.pocp.com/real-applications-artificial-intelligence-in-medication-management/.
59. Collins N. Artificial intelligence helps Stanford computer scientists predict the side effects of millions of drug combinations. Stanford News 2018; (Accessed October 2019): https://news.stanford.edu/ 2018/07/10/ai-predicts-drug-pair-side-effects/.
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 SC, Kuo KL, Peng CH, 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 GR, Edward N, 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. Predictive Analytics World 2018; https://www.predictiveanalyticsworld.com/health/2018/agenda/#session52041.
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; (Accessed December 2019): https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines- trustworthy-ai.
77. Administration UFaD: Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD). In, edited by Services DoHaH, 2019.

Chapter 3

Machine Learning Directed Interventions Associate with Decreased Hospitalization Rates in Hemodialysis Patients

Sheetal Chaudhuri, Hao Han, Len Usvyat, Yue Jiao, David Sweet, Allison Vinson, Stephanie Johnstone Steinberg, Dugan Maddux, Kathleen Belmonte, Jane Brzozowski, Brad Bucci, Peter Kotanko, Yuedong Wang, Jeroen P Kooman, Franklin W Maddux, John Larkin

Int J Med Inform. 2021;153:104541

Abstract

Background
An integrated kidney disease company uses machine learning (ML) models that predict the 12-month risk of an outpatient hemodialysis (HD) patient having multiple hospitalizations to assist with directing personalized interdisciplinary interventions in a Dialysis Hospitalization Reduction Program (DHRP). We investigated the impact of risk directed interventions in the DHRP on clinic-wide hospitalization rates.

Methods
We compared the hospital admission and day rates per-patient-year (ppy) from all HD patients in 54 DHRP and 54 control clinics identified by propensity score matching at baseline in 2015 and at the end of the pilot in 2018. We also used paired t test to compare the between group difference of annual hospitalization rate and hospitalization days rates at baseline and end of the pilot.

Results
The between group difference in annual hospital admission and day rates was similar at baseline (2015) with a mean difference between DHRP versus control clinics of -0.008±0.09 ppy and -0.05±0.96 ppy respectively. The between group difference in hospital admission and day rates became more distinct at the end of follow up (2018) favoring DHRP clinics with the mean difference being -0.155±0.38 ppy and -0.97±2.78 ppy respectively. A paired t-test showed the change in the between group difference in hospital admission and day rates from baseline to the end of the follow up was statistically significant (t-value=2.73, p-value<0.01) and (t-value=2.29, p-value=0.02) respectively.

Conclusions
These findings suggest ML model-based risk-directed interdisciplinary team interventions associate with lower hospitalization rates and hospital day rate in HD patients, compared to controls.

Introduction

End Stage Kidney Disease (ESKD) patients frequently experience emergent complications requiring hospitalization. In the United States, ESKD patients treated by dialysis had on average 1.7 admissions during 2018.[1] In addition to the negative consequences to the patient from the onset/exacerbation of a disease requiring inpatient care, hospitalizations are economically impactful to the healthcare system accounting for about 30% of all ESKD Medicare expenditures.[1] An array of clinical parameters associate with hospitalization events in dialysis patients, including potentially modifiable factors related to the patients' dietary, psychosocial, and other needs.[2-6] However, classifying a dialysis patient's risk for hospitalization, understanding the root cause of likely complications, and providing timely interventions can be challenging based on standard practices.

Most ESKD patients in the United States are treated by outpatient hemodialysis (HD) performed thrice weekly, which amasses large volumes of longitudinal clinical data in electronic medical records (EMR). This robust data collected on HD patients affords opportunities to use Artificial Intelligence (AI) methods to assist with individualized hospitalization risk classifications. AI is an overarching group of advanced analytical techniques bringing together concepts from fields such as computer science, statistics, algorithmics, machine learning (ML), information retrieval, and data science at large.[7] ML techniques are very powerful in their ability to quickly detect hidden patterns in large datasets[8] and have been reported to have the ability to assist with prediction of likely future outcomes for mortality, transplant failure, and other events in ESKD.[9-13]

An integrated kidney disease company has been leveraging historic patient data with advanced analytics to direct care in quality improvement efforts at its national network of dialysis clinics. The provider has developed and operationally deployed a set of ML models that identify in-center HD patients at an increased risk for multiple all-cause hospitalizations within the next 12 months. The models were used in.a pilot called the Dialysis Hospitalization Reduction Program (DHRP) that provides risk directed interdisciplinary team root cause evaluations and personalized interventions to HD patients predicted to be at risk of ≥6 hospital admissions within the next 12 months.

We investigated the impact of the DHRP on clinic-wide hospitalization rates in HD patients. Moreover, in light of other analysis supporting the use of behavioral health interventions, we explored profiles of psychosocial barriers that included patient reported outcomes for depression, sleep, and psychological stress status related to one of the DHRP interventions that was explicitly recorded in the EMR.[14]

Materials and methods

General design

An integrated kidney disease company (Fresenius Medical Care, Waltham, MA, United States) developed a set of ML prediction models for classification of individual dialysis patients at risk for multiple hospitalizations within the next 12 months. These ML models have been used since 2016 to direct interdisciplinary evaluations and interventions in the DHRP quality improvement pilot being performed at select dialysis clinics among a national network (Fresenius Kidney Care, Waltham, MA, United States).

We performed an analysis to evaluate the rolling annual hospital admission and day rates per-patient- year (ppy) in clinics before (2015) and after (2016-2018) the initiation of DHRP and compared rates to matched control clinics not involved in the DHRP.

This analysis was performed under a protocol reviewed by New England Institutional Review Board (Needham Heights, MA, United States; Version 1.0 NEIRB# 17-1305247-1; Revised Version 1.1 NEIRB# 17-1344262-1) who determined this analysis of existing patient data that was de-identified by the investigator was exempt and did not require informed consent. This analysis was conducted in adherence with the Declaration of Helsinki.

Patient population

In this analysis, we included data from HD patients treated in a clinic participating in the DHRP quality improvement pilot implemented across the United States. For the selection of control clinics, we assessed data from the national network and matched clinics with similar attributes that were not involved in the DHRP. Control clinics were matched in a 1:1 ratio to DHRP clinics on the logit of the propensity score for the number of HD patients in the clinic, average age, percentage of male/female, percentage of white/black patients, average albumin, presence of comorbidities (congestive heart failure (CHF), diabetes, ischemic heart disease), and the hospital admission and day rates during the baseline period (2015). We excluded data from all clinics that were managed via the ESKD Seamless Care Organization (ESCO) program, irrespective of their participation in the DHRP.

ML models

ML model was developed to classify an individual's 12-month risk of multiple hospitalizations using historical EMR data from nearly 150,000 in-center HD patients. Overall, there were close to 300 input variables used in the ML models. The input variables included HD patient data on demographics (e.g. age, height, gender, race,

ethnicity), comorbidities, hospitalization history, treatment history, clinical laboratory parameters collected in the EMR, responses from quality of life (QOL) surveys, as well as publicly available data on social, economic, environmental, and geographical factors based on the patient's zip code. (**Figure 3.1**). Many of the treatment, laboratory, and quality of life survey related input variables from the model are described in **Supplemental Table S3.1**, which include the top predictors identified. Additional features were derived from the numeric input variables using coefficient of variations, standard deviations, and monthly linear slopes to identify changes over the month for certain input parameters. Categorical variables such as comorbidities, quality of life and certain socio-economic factors were converted into additional binary variables with a 1 or 0. For development of the ML models, the historic data for all patients were randomly split into 50% training, 20% validation, and 30% test dataset.

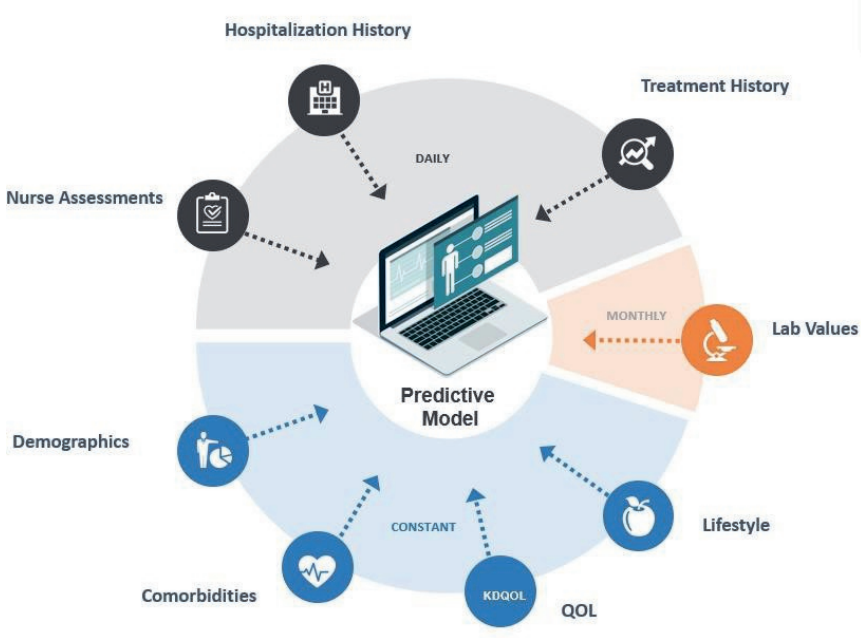


Figure 3.1 Various input sources for developing a machine learning (ML) model to predict patients at risk of hospital admissions in the next 12 months.

The ML model to predict patients having ≥6 hospitalizations within the following 12 months were categorized into two binary classification tasks. One was developed for patients who received in-center HD treatments for at least 120 days, and the other one was developed for patients with less than 120 days of in-center HD treatments. Both models were built using gradient boosting models built using XGboost package.[15]

A 5-fold cross validation on the combination of training and validation data set was used in the model development and the final model took the average of the cross-validation steps to avoid overfitting. Hyperparameters are model-specific internal parameters that are initially set to certain default values to cover general use cases. These parameters must be tuned for the problem at hand to get optimal model performance.[16] The internal parameters such as the maximum tree depth, number of trees and minimum child weight in the XGBoost modeling technique were optimized using exhaustive grid search to iteratively achieve the best AUROC performance.[17]

For each patient, if the models predicted the probability of the patient having at least 6 hospital admissions in the following 12 months; we defined those patients as high-risk. The threshold selected for the model was 0.20. For the ML model directed interventions in the DHRP, the interdisciplinary teams utilized the two high risk models that identified patients at risk of at least 6 hospital admissions in the next 12 months. Top three predictors of the ML model were the QOL Survey Response related to general wellbeing of health, treatment rate accounting treatments missed in the prior year and the average albumin for the patient.

The performance of the model was evaluated using area under the receiver operating characteristic curve (AUROC).[18] The AUROC for the first model for patients with at least 120 days of in-center HD treatments was 0.81, precision of 0.28, recall of 0.57 and F1 score of 0.38 on the test data. The AUROC for the second model for patients less than 120 days of in-center HD treatments was 0.80, precision of 0.29, recall of 0.70 and a F1 score of 0.41 on the test data.

Interventions

Starting in 2016, DHRP clinics received a list of high-risk patients via a secured email and spreadsheet every month. The clinic manager, or a nurse assigned to lead the program, partnered with an interdisciplinary team of social workers, dieticians, and nurses to triage evaluations/interventions to high-risk patients. The interdisciplinary team performed a root cause analysis of risk for hospitalization for each patient identified as high-risk by the ML model. Each team member formulated individualized goals to address any medical instability that could lead to a hospitalization. Also, routine brief meetings were scheduled to discuss team actions taken and next steps for each high-risk patient.

For patients classified to be high-risk by ML model, social workers provided intensive assessment of psychosocial and quality of life barriers, and when appropriate, offered select high-risk patients additional psychosocial evaluations/interventions to target identified barriers through the Social Work Intensive (SWI) program. In the SWI, patient

reported outcome questionnaires were administered before and after the intervention, and social workers charted additional notes on the encounters in the EMR. Dietitians utilized a high-risk assessment workflow looking at weight, nutrition, and access to food and supplements. Nurses assessed high-risk patients focusing on anemia, adequacy, dialysis access, blood pressure, fluid management, prior hospitalizations, glycemic control, and risk of skin ulcers and blood stream infection. Although additional interventions were intended to have been performed by social workers, nurses, and dietitians, the visits were charted in free text fields for routine notes without ability to decipher distinct areas of interventions in the EMR.

Social worker intensive program

The SWI program was one of several components of the DHRP intervention. SWI was a previously established program throughout the provider's national network of clinics, whereby social workers provided additional psychosocial evaluations/interventions (without the ML model) based on identified issues with treatment adherence, achievement of clinical quality targets, and/or other complications. This program demonstrated positive quality of life and hospitalization outcomes.[19,20]

In the DHRP, social workers used the ML model report to direct the SWI behavioral health interventions and dedicated additional time for all patients identified to be at a high risk of hospitalization. High-risk patients with psychosocial concerns were screened with a sleep quality survey (**Supplemental File S3.1**), psychological stress survey (**Supplemental File S3.2**) and the Centre for Epidemiological Studies Depression Scale (CES-D-10) survey [21], all of which are rated on a 1-10 Likert scale. Patients that screened positive for any barriers with sleep, distress, or depressive symptoms were eligible to participate in an 8-week SWI program with an optional 8-week extension (total of 16 weeks) for patients with continuing barriers in self-reported psychosocial outcomes. In the SWI program, the social workers delivered individually tailored cognitive, behavioral, sleep and interpersonal counseling to reduce any patient distress identified by the screening. At the end of the 8-week period, patients were re-screened with the surveys to assess intervention associated improvements in the outcomes. Social workers entered the pre and post screening measurements in the EMR.

Statistical methods

Descriptive statistics

Descriptive statistics were tabulated for demographics and clinical parameters for patients in DHRP clinics and patients in control clinics.

Measuring social worker intervention

We were unable to measure most DHRP related interventions performed by the interdisciplinary team members explicitly due to the limitations in the EMR documentation systems. However, we were able to use social worker evaluation notes and the psychosocial screeners administered in the SWI program as a measure of interventions performed for select high-risk patients. We used the number of assessment and intervention notes recorded by the social workers in the EMR to measure the SWI intervention before and after being identified as high-risk by the ML model. A two-sample T-statistic was used to evaluate the difference between the average number of EMR notes before and after being identified as high-risk. Similarly, T-statistics were used to see if there is a difference in the average depression, stress, and sleep scores before and after the SWI program.

Clinic wide hospital admission and hospitalization days rate

Data from patients at participating clinics were collected and clinic level rolling yearly hospital admission and hospital day rates per patient year (ppy) were calculated at baseline (2015) and 3 years after (2016- 2018) the DHRP program started. Outcomes of the DHRP clinics were compared against control clinics that did not receive the high-risk patient report generated from the ML model. We computed the difference between annual hospital admission and day rates in DHRP and control clinics at baseline (2015) and at the end of follow up (2018). Paired t-tests were used to evaluate the change in the between group difference in the hospital admission and day rates from baseline (2015) to the end of the analysis period (2018).

Results

Clinic and patient characteristics

We used data from all active in-center HD patients in a DHRP and control clinics across the United States over the analysis period from 2016-2018. There were 54 DHRP clinics that had 7767 to 8189 active patients per year during the analysis period. There were 54 control clinics that had 7484 to 7705 active patients each year during the analysis period. The characteristics of patients at DHRP and control clinics is shown for all distinct patients in the clinic groups at baseline during 2015 (**Table 3.1**) and during the follow up period between 2016-2018 (**Table 3.2**).

Table 3.1 Demographic and comorbidities breakdown of patients in DHRP and control clinics at baseline.

	DHRP	Control	Mean Standard Difference
Number of clinics	54	54	
Average number of patients	182	188	-0.067
Average age	61.6	61.0	0.201
Male (%)	57%	56%	0.113
White (%)	61%	60%	0.027
Black (%)	33%	37%	-0.156
CHF (%)	21.5%	23.5%	-0.252
Diabetes (%)	68.1%	68.8%	-0.083
Hypertension (%)	65.5%	70.1%	-0.368
IHD (%)	18.6%	20.1%	-0.138
Average albumin	3.75	3.74	0.275
Hospital admission rate in 2015	1.35	1.35	0
Hospital admission days in 2015	9.26	9.26	0

CHF: Chronic Heart Failure, IHD: Ischemic Heart Disease.

Table 3.2 Demographic and comorbidities breakdown of patients in DHRP and control clinics during follow-up period 2016-2018.

	DHRP	Control
Number of clinics	54	54
Average number of patients	245	251
Average age	61.3	60.8
Male (%)	58%	58%
White (%)	60%	59%
Black (%)	33%	39%
CHF (%)	19.3%	22.8%
Diabetes (%)	69.1%	69.5%
Hypertension (%)	67.6%	73.0%
IHD (%)	18.6%	23.2%
Average albumin	3.7	3.7

CHF: Chronic Heart Failure, IHD: Ischemic Heart Disease.

During the analysis period, 1084 unique patients were identified as high-risk (predicted to have ≥6 hospital admissions in the following 12 months) by the ML model in the DHRP clinics. Within the DHRP clinics, the interdisciplinary teams evaluated and intervened on all high-risk patients. Fourteen percent (14%) of the patients (n=150) predicted to be at a high-risk by the ML model received the behavioral health screening by the social workers. The remaining ten percent (10%) of the total patients identified as high risk (n=111) were enrolled in the SWI program.

Measuring social worker intervention

Among the subset of high-risk patients identified who received behavioral health screening by the social workers (n=150), we found the average number of social work assessment and intervention notes charted in the EMR increased from 1.34±1.02 notes per month in 90 days before to 1.70±1.24 notes in the 90 days after the risk classification (p=0.01).

In the select subset of high-risk patients who received the SWI intervention to address a psychosocial barrier identified (n=111), we found improvements in patient reported outcomes for depressive symptoms, sleep quality, and psychological stress. We observed 62% of high-risk patients enrolled in the SWI program reported a reduction in their depressive symptoms as seen by CESD-10 scores, 64% patients reported an improvement in their sleep quality scores, and 56% patients reported a reduction in their psychological stress scores (**Table 3.3**). After the 8-week SWI intervention, the average CESD-10 scores decreased from 11.4 to 8.12 (CESD-10 scale from 0 to 60 with lower values representing fewer depressive symptoms), sleep scores decreased from 23.2 to 19.7 (sleep screener scale 5 to 50 with lower values representing better sleep quality), and stress scores decreased from 17.0 to 15.9 (stress screener scale 4 to 40 with lower values representing less psychological stress).

Table 3.3 Average Scores Before and After Social Worker Intervention.

Survey	Number of patients	Before intervention Mean±SD (Range)	After intervention Mean±SD (Range)	P-value
CESD-10 score	81	11.42±6.17 (0-27)	8.12 ±6.32 (0-21.5)	<0.0001
Sleep score	72	23.22±10.69 (5-50)	19.68 ±11.51 (5-49)	0.0019
Stress score	67	17.01±10.17 (1-40)	15.85±8.28 (1-35)	0.0055

SD: Standard Deviation

Clinic wide hospitalization rate

Annual rolling hospital admission rates had increasing trends over the analysis period in DHRP and control clinics. Cumulatively, the DHRP clinics exhibited a lower growth in admission and day rates as compared to control clinics (**Figure 3.2A and 3.2B**). At the end of follow up in 2018, the annual admission rate among DHRP clinics was 10% lower and the hospital day rate was 8% lower than control clinics. The DHRP clinics had a 5% and 2% increase in the hospital admission and day rates from 2015 to 2018, respectively. The control clinics showed a 15% and 10% increase in the hospital admission and day rates from 2015 to 2018, respectively.

The between group difference in rolling annual hospital admission and day rates was similar at baseline (2015) with a mean difference between DHRP versus control clinics

of -0.008±0.09 ppy for the admission rate and -0.05±0.96 ppy for the hospital day rate. The between group difference in hospital admission and day rates became more distinct at the end of follow up (2018) with the mean difference in DHRP clinics being -0.155±0.38 ppy and -0.97±2.78 ppy lower than control clinics, respectively. A paired t-rates from baseline to the end of the follow up was statistically significant (t-value=2.73, p-value<0.01 for admission rates; t-value=2.29, p-value=0.02 for hospital day rates).

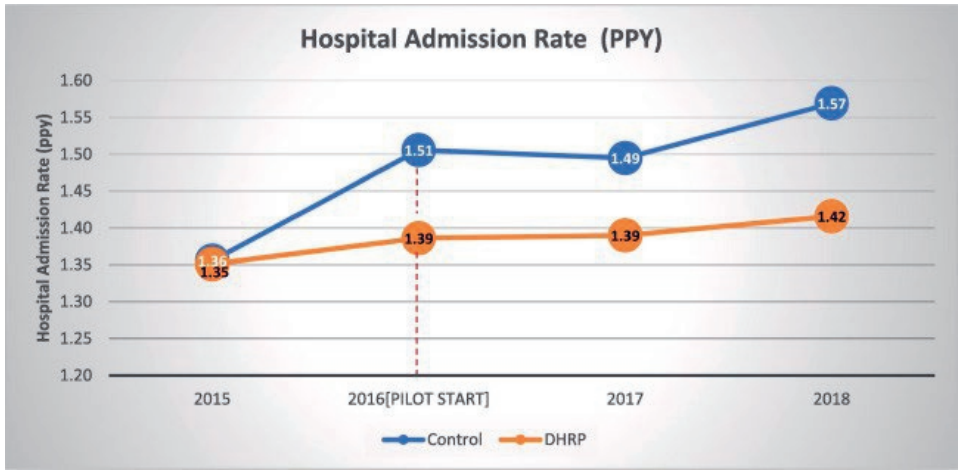


Figure 3.2A Hospitalization rate for DHRP clinics and control clinics at baseline and in follow-up period.

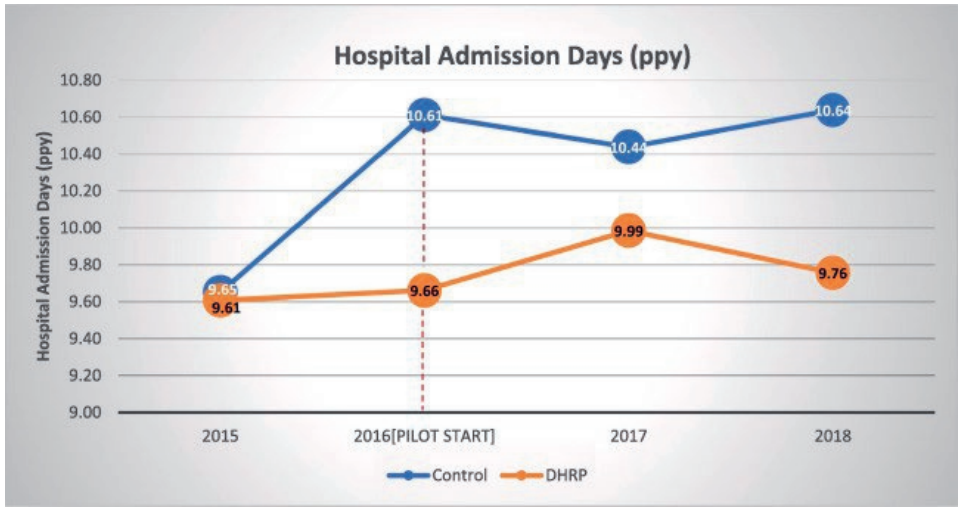


Figure 3.2B Hospitalization days rate for DHRP clinics and control clinics at baseline and in follow-up period.

Discussion

We found ML directed care in outpatient dialysis clinics participating in the DHRP was associated with lower annual rolling all-cause hospitalization rates compared to matched control clinics. DHRP and control clinics exhibited consistent hospitalization rates at baseline in 2015. Although both groups of clinics had increases in hospitalization rates over time, the DHRP clinics had smaller increases compared to control clinics. At the end of follow up in 2018, the between group difference in hospitalization rates was distinct as compared to baseline and showed DHRP clinics had significantly lower hospital admission and day rates versus control clinics. In 2018, this distinction suggests that DHRP was potentially associated with 495 avoided hospital admissions and 2074 avoided hospital days as compared to matched control clinics. It appears monthly ML classifications for patients at a high-risk for multiple hospitalizations (≥6 admissions) in the next year helped direct interdisciplinary team interventions in the DHRP program and lessened the rising trends in the clinics' hospitalization rates over the years. We found 14% of high risk patients were assessed for psychosocial barriers and social workers charted a higher number of notes 90 days after individual patients were identified to be at a high-risk for multiple hospitalizations. Among this subset of high-risk patients, 74% had a psychosocial barrier identified and received tailored behavioral health interventions from the social workers and we observed improvements in patient-reported outcomes for depressive symptoms, sleep quality, and psychological stress. Overall, it appears the ML model was able to aid the DHRP care teams in identifying the right individuals to target for personalized interventions at the right time, overall yielding improvements observable in outcomes at the clinic level.

The risk of hospitalization in ESKD patients is multifactorial and dependent on the etiology of the individual's complication(s). Oftentimes the onset of a disease requiring inpatient care is marked by subtle temporal changes in patient presentations and clinical markers making it a challenge to universally classify hospitalization risk levels during standard of care dialysis practices. To the best of the authors' knowledge, there are no other examples of AI/ML based hospitalization risk models being used to direct care in quality improvement efforts in dialysis. In general, the clinical application of AI in nephrology is scarce with only one report identified in a recent bibliometric study on the global evolution of AI in healthcare.[22,23] The all-cause hospitalization risk models implemented in the DHRP appeared to suitably assist care teams with classifying the subset of individual patients in their clinic at the highest risk. Due to the nature of all-cause risk classification, effective root cause assessment methods to identify complications and target them with effective interventions was essential to yielding the positive results observed.

A recent review article on the factors and barriers related to decreasing admission rates in dialysis patients proposed three key clinical categories driving hospitalization risk that are potentially actionable.[24] These categories include volume control, infection, and psychosocial risks. The DHRP root cause assessments covered these categories, as well as other key areas such as nutrition, dialysis access, glycemic control, and ulcers. While most of these are fundamental areas of focus in dialysis care, psychosocial and behavioral health barriers are often not fully identified in standard care and span outside the field of nephrology.[25,26] High risk patients, in fact, are often seen by the medical team as "too ill" to receive behavioral health intervention and, thus its value is forfeited. Depressive affect and sleep disturbances are known to associate with hospitalization rates and poor outcomes in chronic kidney disease patients.[27-29] The targeted interventions designed to improve mood, sleep quality and perceived stress appear to be possible contributors to the favorable hospitalization rates observed, albeit they were only performed in a small proportion of high-risk patients. Independent analyses are needed to understand the impacts of improvements for each patient reported measure, which include proprietary measures for sleep and psychological stress that have not been validated. Nonetheless, in general it appears the use of ML risk classifications for directing assessments and interventions selected by the patients' care team may help personalize care and improve outcomes.

Although these findings on the application of ML directed care in the DHRP are of importance, there are some limitations to be considered. We only had data to define ML directed assessments specifically by social worker notes charted for concerns of psychosocial barriers, and sleep, depressive symptom, and stress scores were only available on a small subset of patients who received the SWI behavioral health intervention. Although we had some data on select social worker assessments/ interventions, the EMR system of the provider was not structured in a manner that other ML directed assessments/interventions by social workers, dieticians, and nurses could be estimated. Nonetheless, the delivery of personalized interventions in line with needs, preferences, and goals of individuals adds to the novelty of the DHRP and is consistent with the paradigm shifts towards more patient centric care models.[30,31] Having clinicians use their medical discretion to select interventions based on their assessment of each patient is fundamental in healthcare and appeared effective in this experience.

During development of the ML model presented in this paper, we compared advanced ML Techniques such as XGBoost to more traditional techniques such as Generalized Additive Models and found better performance with the advanced ML models. The AUC of the XGBoost model was 0.80-0.81 and the one using GAM was 0.64-0.68 for the different variations of the model.

The successful application of ML directed care in the DHRP holds promise for further development and adoption of ML models that have the potential to provide an intelligent and timely triage of additional resources. Personalized interventions appear prudent to consider in ML directed endeavors, along with a toolbox of assessments and interventions that are anticipated to be effective in improving the quality of life and outcomes of patients. It may be important for the renal care industry to identify ways to increase early access to behavioral health care as patient's increase their risk of hospitalization; using ML modeling might be helpful in directing patients toward this goal. However, ML based risk classification and other clinical decision support tools are never perfect. We believe it is of importance for clinical teams to understand the outputs generated and consider the limitations of the models in the design of directed assessments and interventions.

Conclusions

We found ML directed assessment and personalized interventions in the DHRP, which included behavioral health intervention for some high-risk patients, were associated with lower all-cause hospitalization rates compared to control clinics. The DHRP efforts and findings detail an example of how the clinical application of AI directed care can be successfully conducted and will be of importance for considerations by payors, providers, and clinicians.

References

1. United States Renal Data System. 2018 USRDS annual data report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 2018.
2. Lacson E, Jr., Bruce L, Li NC, et al. Depressive affect and hospitalization risk in incident hemodialysis patients. Clin J Am Soc Nephrol 2014;9:1713-1719.
3. Oquendo LG, Asencio JMM, de Las Nieves CB. Contributing factors for therapeutic diet adherence in patients receiving haemodialysis treatment: an integrative review. J Clin Nurs 2017;26:3893-3905.
4. Adisa O, Jaar BG, Masud T, et al. Association of social worker-assessed psychosocial factors with 30-day hospital readmissions among hemodialysis patients. BMC Nephrol 2018;19:360.
5. Chan KE, Lazarus JM, Wingard RL, et al. Association between repeat hospitalization and early intervention in dialysis patients following hospital discharge. Kidney Int 2009;76:331-341.
6. van Sandwijk MS, Al Arashi D, van de Hare FM, et al. Fatigue, anxiety, depression and quality of life in kidney transplant recipients, haemodialysis patients, patients with a haematological malignancy and healthy controls. Nephrol Dial Transplant 2019;34:833-838.
7. 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.
8. Mitchell T. Machine Learning. McGraw Hill 1997; (Accessed May 2020): http://www.cs.cmu.edu/~tom/mlbook.html.
9. 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.
10. Senanayake S, Barnett A, Graves N, et al. Using machine learning techniques to develop risk prediction models to predict graft failure following kidney transplantation: protocol for a retrospective cohort study. F1000Res 2019;8:1810.
11. Noh J, Yoo KD, Bae W, et al. Prediction of the Mortality Risk in Peritoneal Dialysis Patients using Machine Learning Models: A Nation-wide Prospective Cohort in Korea. Sci Rep 2020;10:7470.
12. Mezzatesta S, Torino C, Meo P, 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.
13. Gabutti L, Vadilonga D, Mombelli G, et al. Artificial neural networks improve the prediction of Kt/V, follow-up dietary protein intake and hypotension risk in haemodialysis patients. Nephrol Dial Transplant 2004;19:1204-1211.
14. Nien Chen L, Stephanie Johnstone, Felicia N. Speed et al. Associations between Sleep Quality and Quality of Life, Stress, and Depressive Symptoms in Hemodialysis Patients J Am Soc Nephrol 2016;27: Page 311A.
15. Chen T GC. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining San Francisco, California, USA, Association for Computing Machinery 2016:785–794.
16. Dernoncourt F, Nemati S, Kassis EB, et al. Hyperparameter Selection. Secondary Analysis of Electronic Health Records: Cham (CH), 2016:419-427.
17. Boto Ferreira M, Costa Pinto D, Maurer Herter M, et al. Using artificial intelligence to overcome over-indebtedness and fight poverty. J Bus Res 2020.
18. Hoo ZH, Candlish J, Teare D. What is an ROC curve? Emerg Med J 2017;34:357-359.
19. Stephanie J, Nien-Chen Li, Franklin Maddux, Eduardo Lacson. Social Worker Driven Program to Reduce Hemodialysis Therapy Non-Adherence. J Neph Social Work 2014:37:41.
20. Nien Chen L, Felicia N. Speed, Marta Reviriego Mendoza et al. Improved Sleep Quality Is Associated with Reduced Hospitalization Rate and Increased Treatment Adherence in Hemodialysis Patients. J Am Soc Nephrol 2016;27:311A.
21. Bradley KL, Bagnell AL, Brannen CL. Factorial validity of the Center for Epidemiological Studies Depression 10 in adolescents. Issues Ment Health Nurs 2010;31:408-412.

22. Tran BX, Vu GT, Ha GH, et al. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. J Clin Med 2019;8:360.
23. 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.
24. Golestaneh L. Decreasing hospitalizations in patients on hemodialysis: Time for a paradigm shift. Semin Dial 2018;31:278-288.
25. Bringsvor HB, Skaug K, Langeland E, et al. Symptom burden and self-management in persons with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2018;13:365-373.
26. Chaiard J, Weaver TE. Update on Research and Practices in Major Sleep Disorders: Part II- Insomnia, Willis-Ekbom Disease (Restless Leg Syndrome), and Narcolepsy. J Nurs Scholarsh 2019;51:624-633.
27. Yamamoto R, Shinzawa M, Isaka Y, et al. Sleep Quality and Sleep Duration with CKD are Associated with Progression to ESKD. Clin J Am Soc Nephrol 2018;13:1825-1832.
28. Iliescu EA, Coo H, McMurray MH, et al. Quality of sleep and health-related quality of life in haemodialysis patients. Nephrol Dial Transplant 2003;18:126-132.
29. Chan L, Tummalapalli SL, Ferrandino R, et al. The Effect of Depression in Chronic Hemodialysis Patients on Inpatient Hospitalization Outcomes. Blood Purif 2017;43:226-234.
30. Chao SH. Patient-centric care management. J Manag Care Pharm 2006;12:S10-13.
31. Constand MK, MacDermid JC, Dal Bello-Haas V, et al. Scoping review of patient-centered care approaches in healthcare. BMC Health Serv Res 2014;14:271.
32. Chaudhuri S, Rosen S,Larkin J et al. Effects of Predictive Modelling Risk directed Interventions on Hospitalizations Rates in Haemodialysis Patients [Abstract MO031]. Nephrology Dialysis Transplantation 2020;35:149.
33. Chaudhuri S, Han H, Usvyat L, et al. Machine Learning Directed Interventions Associate with Decreased Hospitalization Rates in Hemodialysis Patients. Int J Med Inform. 2021;153:104541.

Suppplemental material

Supplemental file 3.1

Patient Name:	**Date of Pre-Screening:** **Date of Post Screening:**

Sleep Screener

Sleep problem identified	How bothered are you by it? Rate 1-10 (1=not at all) (10 = very bothered)	Sleep problem identified	How bothered are you by it? Rate 1-10 (1=not at all) (10 = very bothered)
1. Difficulty *falling* asleep	1 2 3 4 5 6 7 8 9 10	**4. *Interrupted* sleep** (awakening, then falling asleep again on and off through the night)	1 2 3 4 5 6 7 8 9 10
2. Difficulty *staying* asleep	1 2 3 4 5 6 7 8 9 10	**5. *Restless legs* that awaken you** *Restless legs syndrome is the uncomfortable feeling in legs or an urge to keep moving your legs	1 2 3 4 5 6 7 8 9 10
3. Difficulty *awakening*	1 2 3 4 5 6 7 8 9 10		

Are you currently taking any medications For Sleep? For Anxiety? For Nerves? For Stress? For Depression? For Pain? ❑Yes ❑No

If yes to any of the above, list below:

Name Name of Medication	Taken for what symptoms?	How often do you take them?

Has patient ever been referred for a sleep evaluation? ❑Yes ❑No
Has patient ever participated in a sleep evaluation? ❑Yes ❑No

Comments:

Supplemental file 3.2

Stress Screener

Ask patient: **"When you feel stressed or worried, what do you find yourself worrying about most lately?"**	
Indicate theme of the stressor and describe what the patient worries about *(ask if pt can tell you more about what they worry will happen)*	**How bothered are you by this? Rate 1-10** (1=not at all 10 = the most bothered)
❑ Financial/Insurance ❑ Family/Relationships ❑ Health Symptoms ❑ Loss/Grief ❑ Other Describe:	1 2 3 4 5 6 7 8 9 10
❑ Financial/Insurance ❑ Family/Relationships ❑ Health Symptoms ❑ Loss/Grief ❑ Other Describe:	1 2 3 4 5 6 7 8 9 10
❑ Financial/Insurance ❑ Family/Relationships ❑ Health Symptoms ❑ Loss/Grief ❑ Other Describe:	1 2 3 4 5 6 7 8 9 10
❑ Financial/Insurance ❑ Family/Relationships ❑ Health Symptoms ❑ Loss/Grief ❑ Other Describe:	1 2 3 4 5 6 7 8 9 10

Table S3.1 Important treatment, laboratory and quality of life related input variables.

	N	Mean (±Std Dev)
Pre treatment systolic blood pressure [mmHg]	157467	148.92(±18.63)
Pre treatment diastolic blood pressure [mmHg]	157464	77.88(±11.69)
Pre treatment weight [kg]	157467	85.42(±23.9)
Pre Temperature [F]	157467	97.43(±0.42)
Treatment time [min]	157468	223.9(±32.67)
Inter dialytic weight gain [kg]	156647	2.41(±0.96)
Ultrafiltration rate [mL/hr/kg]	157357	8.08(±3.02)
Albumin [g/dL]	155226	3.81(±0.38)
Bicarbonate [mEq/L]	155103	23.25(±2.23)
Calcium [mg/dL]	155421	9(±0.57)
Ferritin [ng/mL]	153321	981.03(±513.9)
Hemoglobin [g/dL]	156561	10.69(±0.96)
Lymphocytes [%]	113741	20.18(±7.25)
Neutrophils [%]	113741	65.99(±8.41)
Phosphorous [mg/dL]	155464	5.25(±1.25)
Potassium [mEq/L]	155505	4.74(±0.51)
Serum sodium[mEq/L]	151071	138.13(±2.74)
Neutrophil lymphocyte ratio [NLR]	113741	4.3(±3.76)
ekTV	149496	1.49(±0.27)
Creatinine [mg/dL]	154302	8.41(±2.97)
Count of comorbidities	157468	21.02(±15.85)
Age [yrs]	157468	61.62(±14.45)
Vintage [yrs]	157468	3.31(±3.84)
QOL Ans 01: General wellbeing of health [ranked percentile]	139910	44.19(±19.12)
QOL Ans 02: Issues with moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf [ranked percentile]	139917	51.08(±32.58)
QOL Ans 03: Issues with climbing several flights of stairs [ranked percentile]	139914	42.33(±33.29)
QOL Ans 04: Accomplished less than you would like due to physical health [ranked percentile]	139899	47.55(±39.02)
QOL Ans 05: Were limited in the kind of work or other activities due to physical health [ranked percentile]	139889	42.83(±38.97)
QOL Ans 06: Accomplished less than you would like due to emotional health [ranked percentile]	139898	71.07(±35.39)
QOL Ans 07: Didn't do work or other activities as carefully as usual due to emotional health [ranked percentile]	139873	77.29(±31.84)
QOL Ans 08: How much did pain interfere with your normal work? [ranked percentile]	139895	67(±25.65)
QOL Ans 09: Have you felt calm and peaceful? [ranked percentile]	139903	68.69(±20.8)
QOL Ans 10: Did you have a lot of energy? [ranked percentile]	139905	48.12(±22.92)
QOL Ans 11: Have you felt downhearted and blue? [ranked percentile]	139888	77.69(±20.29)
QOL Ans 12: How much of the time has your physical health or emotional problems interfered with your social activities? [ranked percentile]	139894	71.69(±24.1)
QOL Ans 13: My kidney disease interferes too much with my life [ranked percentile]	139833	44.91(±29.02)

Table S3.1 (contiued)

	N	Mean (±Std Dev)
QOL Ans 14: Too much of my time is spent dealing with my kidney disease [ranked percentile]	139827	49.56(±29.1)
QOL Ans 15: I feel frustrated dealing with my kidney disease [ranked percentile]	139822	56.1(±30.02)
QOL Ans 16: I feel like a burden on my family [ranked percentile]	139821	64.91(±30.62)
QOL Ans 17: Soreness in your muscles? [ranked percentile]	139788	72.39(±23.51)
QOL Ans 18: Chest pain? [ranked percentile]	139793	92.17(±13.94)
QOL Ans 19: Cramps? [ranked percentile]	139790	75.58(±21.48)
QOL Ans 20: Itchy skin? [ranked percentile]	139790	75.02(±23.25)
QOL Ans 21: Dry skin? [ranked percentile]	139795	72.12(±23.64)
QOL Ans 22: Shortness of breath? [ranked percentile]	139808	82.86(±20.32)
QOL Ans 23: Faintness or dizziness? [ranked percentile]	139802	86.67(±17.12)
QOL Ans 24: Lack of appetite? [ranked percentile]	139804	84.42(±19.31)
QOL Ans 25: Washed out or drained? [ranked percentile]	139775	70.85(±23.17)
QOL Ans 26: Numbness in hands or feet? [ranked percentile]	139792	75.09(±24.79)
QOL Ans 27: Nausea or upset stomach? [ranked percentile]	139787	83.41(±20.16)
QOL Ans 28A: (Hemodialysis patient only) Problems with your access site? [ranked percentile]	138719	90(±16)
QOL Ans 28B: (Peritoneal dialysis patient only) Problems with your catheter site? [ranked percentile]	6721	84.3(±25.58)
QOL Ans 29: Fluid restriction? [ranked percentile]	139771	72.7(±25.07)
QOL Ans 30: Dietary restriction? [ranked percentile]	139777	76.64(±22.38)
QOL Ans 31: Your ability to work around the house? [ranked percentile]	139758	73.61(±24.97)
QOL Ans 32: Your ability to travel? [ranked percentile]	139751	67.99(±28.66)
QOL Ans 33: Being dependent on doctors and other medical staff? [ranked percentile]	139759	79.02(±23.59)
QOL Ans 34: Stress or worries caused by kidney disease? [ranked percentile]	139759	76.9(±24.14)
QOL Ans 35: Your sex life? [ranked percentile]	138405	80.78(±26.94)
QOL Ans 36: Your personal appearance? [ranked percentile]	139674	84.43(±21.39)
Physical component score	139703	37.93(±9.07)
Treatment rate	157468	0.42(±0.08)
Number of monthly foot checks	157468	1.14(±1.02)

QOL: Quality of Life.

Chapter 4

Real-Time Prediction of Intradialytic Relative Blood Volume: A Proof-of- Concept for Integrated Cloud Computing Infrastructure

Sheetal Chaudhuri, Hao Han, Caitlin Monaghan, John Larkin, Peter Waguespack, Brian Shulman, Zuwen Kuang, Srikanth Bellamkonda, Jane Brzozowski, Jeffrey Hymes, Mike Black, Peter Kotanko, Jeroen P Kooman, Franklin W Maddux, Len Usvyat

BMC Nephrol 2021;22:274

Abstract

Background
Inadequate refilling from extravascular compartments during hemodialysis can lead to intradialytic symptoms, such as hypotension, nausea, vomiting, and cramping/myalgia. Relative blood volume (RBV) plays an important role in adapting the ultrafiltration rate which in turn has a positive effect on intradialytic symptoms. It has been clinically challenging to identify changes RBV in real time to proactively intervene and reduce potential negative consequences of volume depletion. Leveraging advanced technologies to process large volumes of dialysis and machine data in real time and developing prediction models using machine learning (ML) is critical in identifying these signals.

Method
We conducted a proof-of-concept analysis to retrospectively assess near real-time dialysis treatment data from in-center patients in six clinics using Optical Sensing Device (OSD), during December 2018 to August 2019. The goal of this analysis was to use real-time OSD data to predict if a patient's relative blood volume (RBV) decreases at a rate of at least -6.5% per hour within the next 15 minutes during a dialysis treatment, based on 10-second windows of data in the previous 15 minutes. A dashboard application was constructed to demonstrate how reporting structures may be developed to alert clinicians in real time of at-risk cases. Data was derived from three sources: (1) OSDs, (2) hemodialysis machines, and (3) patient electronic health records.

Results
Treatment data from 616 in-center dialysis patients in the six clinics was curated into a big data store and fed into a Machine Learning (ML) model developed and deployed within the cloud. The threshold for classifying observations as positive or negative was set at 0.08. Precision for the model at this threshold was 0.33 and recall was 0.94. The area under the receiver operating curve (AUROC) for the ML model was 0.89 using test data.

Conclusion
The findings from our proof-of concept analysis demonstrate the design of a cloud-based framework that can be used for making real-time predictions of events during dialysis treatments. Making real- time predictions has the potential to assist clinicians at the point of care during hemodialysis.

Introduction

Hemodialysis (HD) involves removal of fluid from the circulating blood by ultrafiltration and refilling from the extravascular compartments.[1] This process helps preserve blood pressure and tissue perfusion.[2] However, inadequate refilling could lead to a variety of intradialytic symptoms, such as intradialytic hypotension (IDH), fatigue, and cramping.[3,4] IDH can lead to cardiac complications and an increased risk of death.[5-8]

Studies have shown the role of relative blood volume (RBV) and how adapting the ultrafiltration rate has a positive effect on intradialytic symptoms.[9,10] However, it has been clinically challenging to identify changes in RBV in real time to proactively intervene and reduce potential negative consequences of volume depletion. Hence leveraging advanced technologies to process large volumes of dialysis and machine data in real time and developing prediction models using machine learning (ML) is critical in identifying these signals.

A network of dialysis clinics routinely captured hematocrit, oxygen saturation, and intravascular blood volume during dialysis using Optical Sensing Device (OSD) device.[11-13] The OSD provides clinicians with the ability to have near real-time monitoring of the patient's clinical status during HD. During dialysis treatments, data is collected every ten seconds, which is required to be stored, curated, and analyzed timely interventions. There is a dearth of knowledge about utilizing this machine data for monitoring treatment level parameters and personalizing care for HD patients. This may be secondary to traditional storage and computing resources being unable to handle the processing of such large data stores.

Big data technologies and cloud-based services are novel tools that can provide the necessary infrastructure to support such near real-time applications. Big data is a field that incorporates ways to analyze, systematically extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional software.[14] Cloud technology moves big data processing off local computers and onto shared web services, allowing for greater optimization of resources and faster processing as a result. Cloud platforms provides a secure, efficient, and reliable way to process and analyze data.

We conducted a retrospective analysis to assess dialysis treatment data from 2019. This analysis was used to develop a proof-of-concept that cloud infrastructure can be used in clinical care and provide necessary data to consider if implementation in the future is warranted. The model developed in this proof-of-concept was not utilized in clinical practice.

We created a ML application to identify patients at risk of having their RBV decrease at a rate of at least -6.5% per hour anytime during HD. A dashboard application was constructed demonstrate how reporting structures may be developed to alert clinicians in real time of at-risk cases.

Methods and design

General design

For this proof-of-concept analysis, we used data from adult patients treated at six clinics (Fresenius Kidney Care, Waltham, MA, United States) that universally used OSD during HD as a standard of care between December 2018 through August 2019. In these six clinics, there was hardware previously setup to transfer data from the OSD device to a secure Internet of Things (IoT) private server on Amazon Web Services (AWS; Amazon Web Services, Inc., Seattle, WA, United States) using IoT software.[15,16] The AWS server was compliant with the Health Insurance Portability and Accountability Act (HIPAA).[15] Amazon Web Services (AWS), Microsoft Azure, and Google's cloud platforms are the most broadly adopted web services platform in the world.[17-19]

The goal of this analysis was to use historic OSD data to build a prediction model that can actively classify patients at risk of having their RBV decrease at a rate of at least -6.5% per hour within the next 15 minutes of HD throughout the entire treatment. Also, we aimed to construct a dashboard to that could be considered for delivery of alerts for patients predicted at risk.

This analysis was performed under a protocol that was approved by New England Institutional Review Board under a waiver of informed consent per title 45 of the United States Code of Federal Regulations part 46.116(f) (Needham Heights, MA, United States; NEIRB# 17-1311567-1). The analysis was conducted in adherence with the Declaration of Helsinki.

Patient population

We included data from patients who were greater than or equal to 18 years of age and females were not known to be pregnant.

Optical sensing device

The OSD (Crit-Line®, Bad Homburg, Germany) profiles patient's intradialytic status to assist clinicians monitor the treatment assessment and intervention during hemodialysis.[20] By monitoring blood volume percent changes, caregivers can adjust

treatment as necessary to maximize fluid removal and prevent common intradialytic symptoms, such as IDH, nausea, vomiting, and cramping[9,21-24], as well as minimize the risk of worse outcomes.[12,25]

Per the manufacture's specifications for RBV thresholds[20], when the rate of change in RBV, based on the latest 15 minutes of data, is decreasing less than -3% per hour, the ultrafiltration rate might be increased without immediate risk of intradialytic symptoms. In this case the patient's plasma refill rate is occurring at the same or a greater rate than the ultrafiltration rate. When the rate of change in RBV, based on the latest 15 minutes of data, is decreasing between -3% and -6.5% per hour, it indicates a suitable compromise between a high ultrafiltration rate and the prevention of intradialytic symptoms. When the rate of change in RBV, based on the latest 15 minutes of data, is greater than -6.5% per hour, there is a rapid decrease in RBV and bears a higher risk for intradialytic symptoms, such as lightheadedness, nausea, vomiting, cramping, or hypotension. Prior studies have shown reductions in intradialytic complications with ultrafiltration based on RBV targets in relatively consistent ranges[9,21-24], and that ultrafiltration performed targeting RBV deceases between -3% and - 6.5% per hour associates with better patient outcomes.[12,25]

Model data and features

The ML model was trained and tested on a static set of historical observations from our system. Data was derived from three sources: (1) OSDs, (2) HD machines, and (3) patient electronic health records.

OSD data and treatment data from the 2008 T® dialysis machines were collected every 10 seconds during dialysis treatments. OSD data included variables like blood volume alert level, RBV, changes in hematocrit, hemoglobin, oxygen saturation, minimum oxygen saturation, and oxygen alert level.

Dialysis machine data included variables such as systolic blood pressure, diastolic blood pressure, mean arterial blood pressure, pulse, delivered equilibrated (e)Kt/V, average small molecular clearance [Kecn], projected single pool (sp) Kt/V, first plasma serum sodium, body volume, blood flow rate, conductivity, dialysate flow rate, intervention performed on the machine, arterial pressure, dialysate temperature, venous pressure, ultrafiltration rate, blood volume processed, ultrafiltration goal, ultrafiltration volume removed, and remaining time on dialysis in minutes (RTD).

Patient demographic information such as age, height, access type, and clinic ID were referenced from the on-premises clinical data warehouse. Patient measures in the clinic on the day of treatment included pre-dialysis/post-dialysis weight and the type of dialyzer used in treatment.

OSD and 2008T® dialysis machine data from five separate time windows: 1, 5, 10, 15 minutes, and since-start-of-treatment windows were used to derive additional intra-treatment features using average, minimum, maximum and standard deviation for each time window. The final dataset spanned 751,354 treatment records and 493 input variables including features for average, minimum, maximum, and standard deviation for the continuous variables at each time point.

Predictive model

The model was built using the AWS SageMaker [26] development platform. The curated final dataset of 751,354 treatment records was randomly split into training data (80%), validation data (10%) and test data (10%). The target variable was a binary indicator of patients who experienced a decrease in RBV at a rate of at least -6.5% per hour during a dialysis treatment within the next 15 minutes.

Figure 4.1 shows the ascertainment period and the prediction period for the model.

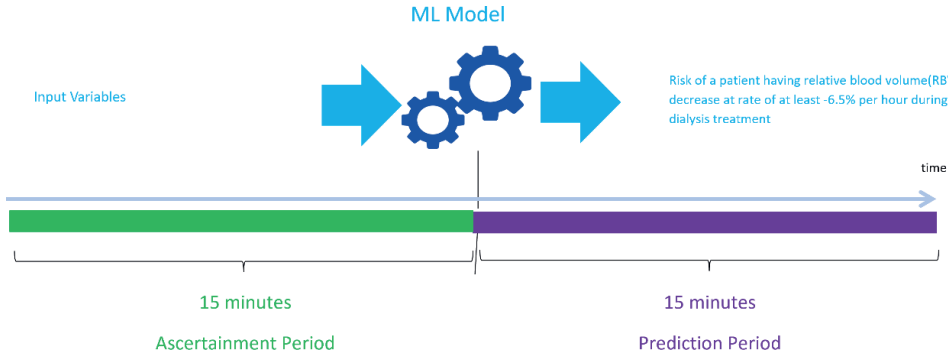


Figure 4.1 Ascertainment period and prediction period.

The data showed a 22% prevalence within observations in which RBV decreased at rate of at least 6.5% per hour during a dialysis treatment. Given the imbalanced nature of our data, we limited our ML model selection to algorithms that deal well with such data, including support vector and random forest families.[27,28] The final model was trained using a ML tool known as an extreme gradient boosting (XGBoost) algorithm.[29] Hyperparameters are model-specific internal parameters that are initially set to certain default values to cover general use cases. These parameters must be tuned for the problem at hand to get optimal model performance.[30] After model selection, the Bayesian optimization strategy implemented by Amazon SageMaker was used to tune the model hyperparameters to maximize the AUROC.[31] In the Bayesian tuning strategy ML algorithms performance is modeled as a sample of a Gaussian process.[32,33] This

allows information from prior iterations to inform the next parameters to try to optimize model performance, balancing both exploration of values not yet used with exploitation of the best-known results.

The predicted probability output by the model was converted to a binary prediction to predict positive and negative cases of RBV decreasing at rate of at least -6.5% per hour during a dialysis treatment. The cut-off threshold for the binary prediction was set to 0.08, so if the prediction score was above 0.08, then the patient was flagged to be at risk of decreasing RBV. The threshold was set by evaluating the results of the training and validation data.

Feature importance from the gradient boosting algorithm was used to derive top features (variables) that were considered highly predictive of the outcome. The feature importance is calculated using the gain method, or the relative contribution of the corresponding feature for each tree in the model. The method works by averaging the training loss reduction caused by feature utilization for each split in the decision tree.[34]

Conceptional analysis design

Figure 4.2 shows a general design of the analysis setup. Conceptually, the analysis consisted of three main components: (1) Hardware and devices needed to monitor patients, (2) Cloud-based Service for real-time data analysis and communication, and (3) On-premises secure Data Warehouse to reference patient-protected information needed for data analysis.

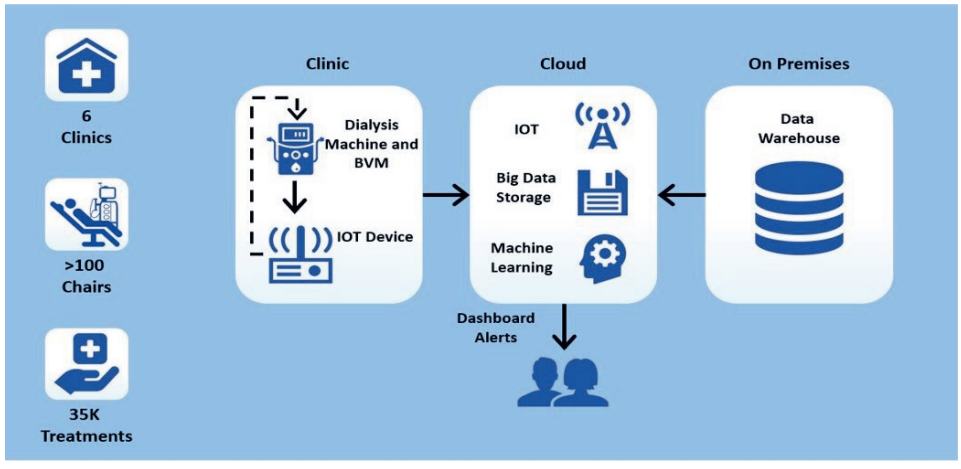


Figure 4.2 Analysis design.

In the cloud, IoT software processed incoming data from the clinic. The data was then curated into a big data store within the cloud. The big data store referenced on-premises data warehouse to securely extract patient-protected information and other clinical data and then to securely feed into the ML model. These multiple sources of data were made available to the Machine Learning Engine (MLE) which was also hosted in the cloud. The MLE would then make a prediction based on the data and generate an alert to the clinicians and nurses for identified at-risk cases. The entire analysis pipeline had to be optimized to ensure low latency (*i.e.* ensure timeliness of near real-time prediction). This optimization process is beyond the scope of the current discussion.

Cloud computing infrastructure flow for generating real time dashboard

The data flow for the entire modeling pipeline within AWS is shown in **Figure 4.3**. Green arrows in Figure 4.3 show how the data flows from the clinics using OSD, the dialysis machines, and the warehouse into the cloud to train a model and provide data to the endpoint interface. The orange arrows show how the data flows in real time from the clinic using OSD and the data warehouse. Each new message in the cloud data store triggers a function, which creates 493 different features used in the trained model. These features are then provided as an input data parameter to the endpoint interface to generate a prediction and store the results in another data store. The prediction results are then used in a dashboard to generate a proof-of-concept clinical user interface.

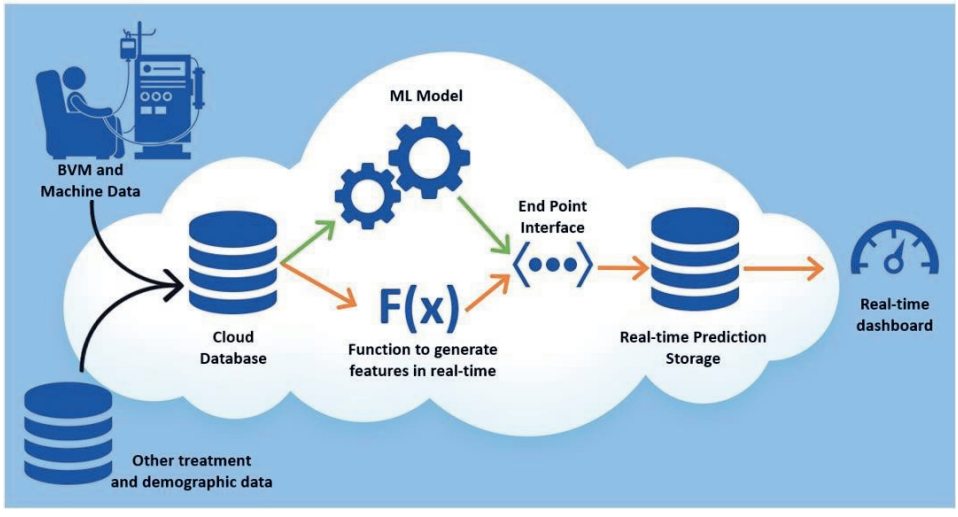


Figure 4.3 Data flow for real-time prediction using the AWS cloud environment.

Analysis of ML model performance

Performance of ML model was measured by the area under the receiver operating curve (AUROC) in the training, validation, and testing datasets, as well as the recall and precision in the testing datasets. AUROC measures the rate of true and false positives classified by the prediction model across probability thresholds. **Table 4.1** shows the definition of true/false positive and negative predictions classified by the model in the assessment of performance in the testing dataset.

Table 4.1 Definition of true/false positive and negative predictions classified by the model in the assessment of performance in the testing dataset.

True positives	Patients correctly classified as having a risk of their relative blood volume (RBV) decrease at a rate of at least -6.5% per hour within the next 15 minutes by the model.
False positives	Patients incorrectly classified as having a risk of their RBV decrease at a rate of at least -6.5% per hour within the next 15 minutes by the model.
True negatives	Patients correctly classified as not having a risk of their RBV decrease at a rate of at least -6.5% per hour within the next 15 minutes by the model.
False negatives	Patients incorrectly classified as not having a risk of their RBV decrease at a rate of at least -6.5% per hour within the next 15 minutes by the model.

Recall (sensitivity) measures the rate of true positives classified by the model at a specified threshold and is calculated as follows:

$$\underline{R}ecall = number\ of\ true\ positives\ /\ (number\ of\ true\ positives + number\ of\ false\ negatives)$$

Precision measures the positive predictive value for the model at a specified threshold and is calculated as follows:

$$Precision = number\ of\ true\ positives\ /\ (number\ of\ true\ positives + number\ of\ false\ positives)$$

Similarly, Specificity for the model is defined as:

$$Specificity = number\ of\ true\ negatives\ /\ (number\ of\ true\ negatives + number\ of\ false\ positives)$$

And the Negative Predictive Value (NPV) is defined as:

$$NPV = number\ of\ true\ negatives\ /\ (number\ of\ true\ negatives + number\ of\ false\ negatives)$$

AUROC, recall, precision, specificity and NPV metrics yield scores on a scale of 0 (lowest) to 1 (highest). A model performing at chance would yield an AUROC of 0.5. The cut-off threshold for classifying predictions was selected to optimize recall and precision according to the use case.

Results

Patient characteristics

We obtained data from 616 adult in-center HD patients that were treated in six clinics to build the prediction model. Patient demographics are shown in **Table 4.2**. The descriptive statistics of numeric input variables used to train the model are shown in **Table 4.3**.

Table 4.2 Demographics of patients at the start of the study period (entire cohort).

Patient characteristics	Value
Number of patients	616
Average age	64.5(SD: ±14.83)
Male	57.8%
Black	27.5%
White	68.6%
Hispanic	22.1%
Congestive heart failure	24.7%
Diabetes	39.1%
Hypertension	77.5%
Ischemic heart disease	24.5%
Average albumin [g/dL]	3.8(SD: ±0.38)

ML model performance and feature importance

The resulting predictive model was tested on 10% (75072 records) of the treatment data from all 616 patients, which was withheld during training.

Table 4.3 Descriptive statistics of numeric input variables (Training data).

Variable	N	Mean ± SD
Diastolic blood pressure [mmHg]	562012	70.13±14.42
Mean arterial blood pressure [mmHg]	562012	96.72±18.13
Mean pulse [bpm]	562012	71.83±11.08
Systolic blood pressure [mmHg]	562012	133.75±24.27
Delivered equilibrated (E)Kt V	415498	0.62±0.24
Mean Kecn	539626	252.38±33.25
Projected single pool (sp)Kt V	533226	0.78±0.54
First plasma Na [mEq/L]	537453	140.24±3.78
Body volume [L]	533226	34.84±8.46
Critline relative blood volume alert [%]	601210	-12.42±3.95
Relative blood volume (RBV) [%]	601210	-4.16±6.87
Changes in hematocrit [%]	601210	26.83±15.42
Hemoglobin [g/dL]	601210	8.96±5.33
Oxygen saturation [%]	601210	69.39±39.25
Minimum oxygen saturation [%]	601210	88.37±16.84
Oxygen alert level [%]	601210	68.29±34.64
Blood flow rate [mL/min.]	601210	348.99±142.23
Conductivity [mS/cm]	601210	13.7±1.09
Dialysate flow rate [mL/min.]	601210	643.03±182.94
Monitor temp [°C]	601210	36.49±0.88
Arterial pressure [mmHg]	601210	-162.43±77.12
Dialysate temperature [°C]	601210	32.86±52.96
Venous pressure [mmHg]	601210	156.84±71.85
Ultrafiltration rate [mL/Hr]	601210	550.6±350.28
Blood volume processed [L]	601210	420.41±271.47
Remaining time on dialysis [min]	601210	95.46±71.16
Ultrafiltration goal [mL]	601210	2506.46±1042.49
Ultrafiltration volume removed [mL]	601210	1280.32±999.72
Age [yrs]	560899	66.57±14.35
Height [cm]	535433	168.1±11
Most recent post-dialysis weight [kg]	552865	80.12±24.05
Most recent pre-dialysis weight [kg]	552865	82.08±24.55
Average 30 days post dialysis weight [kg]	552865	80.06±24.08
Average 30 days pre dialysis weight [kg]	552865	82.03+/-24.59

SD: Standard Deviation.

Using a low threshold of 0.08, the model had a recall rate of 0.94, meaning the model was able to capture 94% of the observations that had a decrease in RBV at a rate of at least -6.5% per hour within the next 15 minutes. The precision of the model was 0.33. The specificity for the model was 0.52 and the NPV was 0.97. The AUROC (**Figure 4.4**) for the final hyperparameter tuned model was 0.89. The red dot on the figure shows the true positive rate and the false positive rate at a threshold of 0.08.

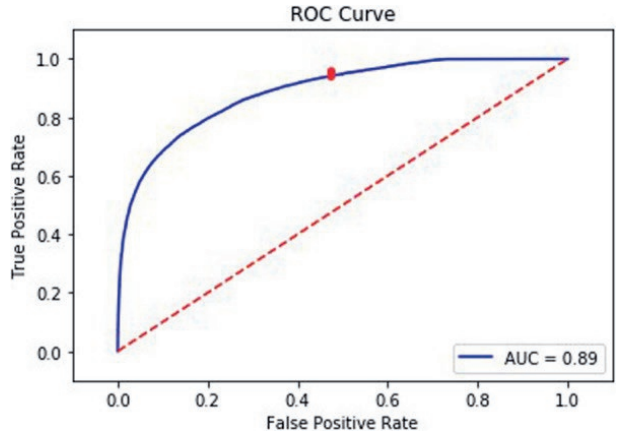


Figure 4.4 Area under the receiver operating curve for the prediction model.

Figure 4.5 shows a list of the top 10 features from the tuned model that were most predictive of a patient experiencing their RBV decrease at a rate of at least -6.5% per hour during a dialysis treatment in the next 15 minutes. It shows how valuable each feature was to the model in predicting the outcome. Higher value of the feature implies it is more important in calculating the outcome of the model.

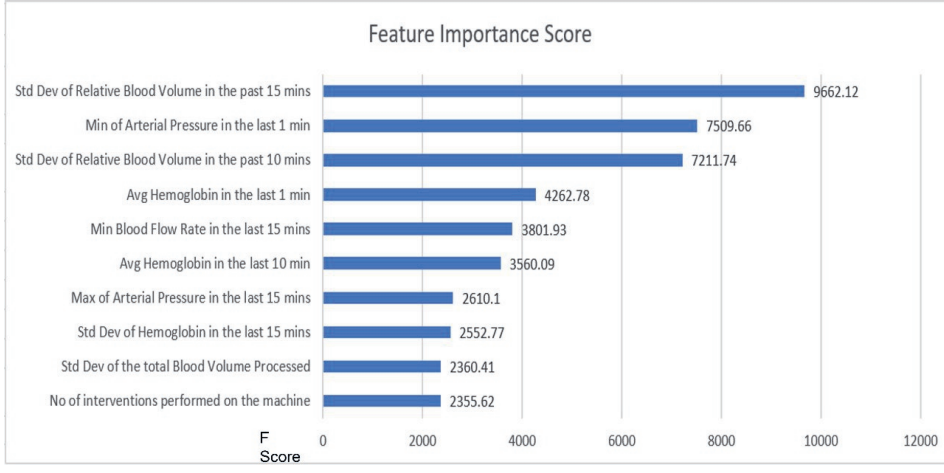


Figure 4.5 Top 10 features from prediction model and the feature importance score.

Proof-of-concept dashboard

Figure 4.6 shows the proof-of-concept dashboard for a patient during dialysis treatment. The patient goes through various stages of having a risk of RBV decreasing at rate of at least -6.5% per hour (that is entering Profile C as shown in the figure). The probability of the prediction of profile C generated from the model is above 80% before the actual occurrence of the event denoted in red under the Profile header. The RBV % at this point drops below -6.5% as shown under the Blood Volume % header. This dashboard illustrates that the model was able to predict the occurrence of the event before it happened.

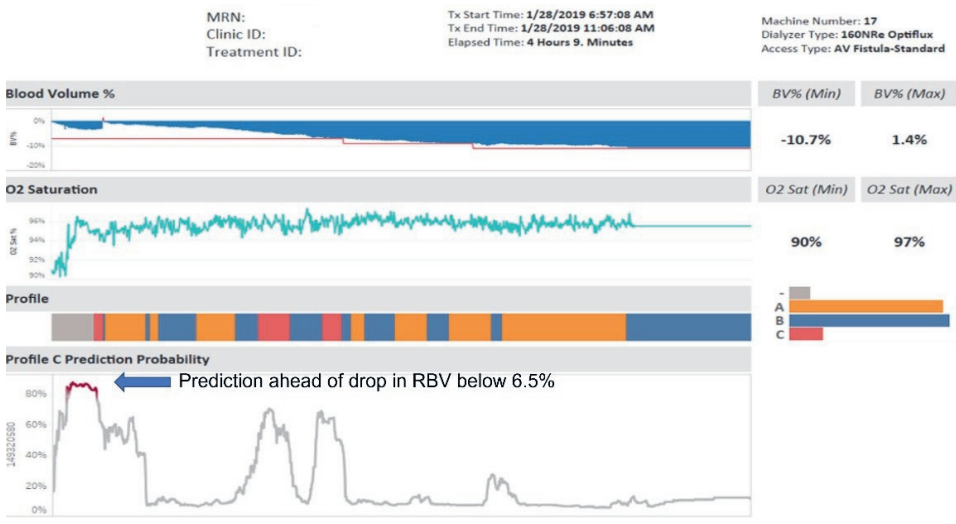


Figure 4.6 Proof-of-concept dashboard for monitoring risk during dialysis treatment.

Discussion

The findings from our proof-of-concept analysis suggest the potential for real-time reporting and prediction of treatment blood volume profiles that are associated with an increased risk of intra-dialytic symptoms and would subsequently be amenable to intervention. Furthermore, the architectural framework demonstrated in this paper can be used for making real-time predictions of other events during dialysis treatments; and as such this analysis serves as a proof-of-concept.

Making real-time predictions can help clinicians and nurses to provide proactive support at the point of care during dialysis treatment. A practical implication for the present would be that, if nurses and clinicians are alerted to the risk of a drop in the blood volume 15 mins prior to the RBV decreasing at rate of at least -6.5% per hour during a dialysis treatment, they would have ample time to intervene and adjust the ultrafiltration rate in order to prevent that patient from entering the risk zone for intradialytic symptoms like IDH.[20]

Prior studies have been attempted to monitor hematocrit and reduce intradialytic symptoms, however, they were not used in standard practice because of the difficulty in interpreting the OSD outputs updated every 10 seconds.[9,35] The ML model presented in this analysis enhances the findings and delivers them in a comprehendible way. The top predictors of a RBV decreasing at a rate of at least -6.5% per hour were shown to include the variability in RBV in the prior 10 and 15 mins, minimum arterial pressure in the prior 1 min, mean hemoglobin in the prior 1 and 10 min(s), and minimum blood flow rate in the prior 15 mins, as well as other metrics related to atrial pressure, hemoglobin, and total blood volume processed (**Figure 4.5**). The feature importance of these parameters appears to be identifying combinations of minor signals providing early signs of issues with ultrafiltration (e.g. peristaltic pump being starved of flow due to higher resistance in the access circuit). Ultimately, this model may have the potential to support the clinicians by classifying risk levels in near real-time. This analysis also adds onto the proof-of-concept analysis from Barbieri et al., where they developed an artificial neural network model predicting session-specific Kt/V, fluid volume removal, heart rate, and BP based on patient characteristics, historic hemodynamic responses, and dialysis-related prescriptions.[36,37]

Cloud Computing Resources provide seamless tools to build, analyze, and integrate real-time predictive models without investing in many hardware and software resources on premise. This allows for a secure and cost-effective way of building predictive models when resources are limited. These applications can also be scaled on-demand, where support can be expanded from tens to hundreds of clinics seamlessly.

Along with the disease burden, inadequate dialysis process may play a role in the pathophysiology of cardiac injury, cognitive impairment, and brain injury in HD patients.[38-40] Large amounts of data collected from the dialysis machines to build and deploy ML models can be used in personalizing dialysis treatments for HD patients. Optimizing dialysate temperature, monitoring access flows, modeling retention solute clearance and electrolyte profiling, and predicting IDH are other examples of how machine data can be utilized to personalize treatments for patients. Successful applications of analyzing and modeling large amounts of clinical data from the

machines will require technology and a framework like what has been presented in this paper.

This paper provides an important proof-of-concept for the application of a ML-based model in the prevention of intradialytic complications. However, it should be stressed that while the decline in RBV during dialysis is an important risk factor for IDH, the critical decline in RBV and the level at which the patient experiences IDH also differs significantly between patients.[7,21] IDH is an important risk factor for mortality, as well as for ischemia of vital organs, such as heart and brain, which may lead to long-term organ damage. Therefore, methods to reduce the risk for this complication are of vital importance [25, 40-42]. Other factors, such as an impairment in vascular reactivity or the cardiovascular status of the patient play an important role in the sensitivity of the patient to a decline in RBV. Moreover, there is a possibility of misclassification of patients at risk, where the model predicts that the RBV will decrease at a rate of at least -6.5% per hour during a dialysis treatment whereas it does not; hence the clinical intervention should be designed in such a way that it does not have an adverse impact on the treatment or the patient. In this respect, it is also important that profiles with a small decline in RBV may carry the risk of adverse outcomes, possibly because of its relationship with fluid overload [25]. Therefore, the results of the model should always be interpreted in the context of the patient.

The goal of this proof- of-concept project was to demonstrate the architecture of how machine data can be utilized in real time. The goal of the dashboard if implemented is to capture as many patients as possible who would have an adverse intradialytic event or have the risk of dropping RBV at the rate of at least -6.5% per hour. Hence, the focus was on sensitivity rather than specificity or precision when determining the threshold used to evaluate model performance. However, in a real-world implementation, the optimal threshold can be selected to minimize either false positives or false negatives, which will depend on the intervention and reporting demands.

This architecture also demonstrates the capabilities of a cloud-based framework in handling the large amounts of patient and treatment data collected from dialysis machines. ML models can be utilized for personalizing care in dialysis patients in real time. However, there will be instances when the ML model will predict incorrectly, so teams developing interventions using ML models need to be aware of this limitation. This proof-of-concept could also be used for predicting low or differing ranges of RBV. The clinical team responsible for designing interventions will need to interpret RBV targets and adjust ultrafiltration rate in a personalized manner considering each patient's unique history of intradialytic complications. Also, the true performance of the ML model can only be demonstrated after conducting randomized clinical trials. The cloud-based framework should allow scaling of this proof-of-concept analysis;

however, this has not been tested in real world application. Models deployed at point of care could also be used to receive feedback from the nurses and clinicians to serve as refined input to retrain the model.

Conclusion

This proof-of-concept analysis demonstrated the potential of the creation and deployment of a real- time predictive model based on patient and dialysis treatment data. The mechanics for triggering a model endpoint based on real-time message capture and to produce real-time reporting that includes treatment metrics coupled with model inferences were successfully implemented. The challenge will be to scale for large amounts of data and to design appropriate interventions.

References

1. Schneditz D, Roob J, Oswald M, et al. Nature and rate of vascular refilling during hemodialysis and ultrafiltration. Kidney Int 1992;42:1425-1433.
2. Daugirdas JT. Dialysis hypotension: a hemodynamic analysis. Kidney Int 1991;39:233-246.
3. Alvarez L, Brown D, Hu D, et al. Intradialytic Symptoms and Recovery Time in Patients on Thrice-Weekly In-Center Hemodialysis: A Cross-sectional Online Survey. Kidney Med 2020;2:125-130.
4. Kuipers J, Verboom LM, Ipema KJR, et al. The Prevalence of Intradialytic Hypotension in Patients on Conventional Hemodialysis: A Systematic Review with Meta-Analysis. Am J Nephrol 2019;49:497-506.
5. Morfin JA, Fluck RJ, Weinhandl ED, et al. Intensive Hemodialysis and Treatment Complications and Tolerability. Am J Kidney Dis 2016;68:S43-S50.
6. Gul A, Miskulin D, Harford A, et al. Intradialytic hypotension. Curr Opin Nephrol Hypertens 2016;25: 545-550.
7. Sars B, van der Sande FM, Kooman JP. Intradialytic Hypotension: Mechanisms and Outcome. Blood Purif 2020; 49:158-167.
8. Keane DF, Raimann JG, Zhang H, et al. The time of onset of intradialytic hypotension during a hemodialysis session associates with clinical parameters and mortality. Kidney Int 2021;99:1408-1417.
9. Steuer RR, Leypoldt JK, Cheung AK, et al. Reducing symptoms during hemodialysis by continuously monitoring the hematocrit. Am J Kidney Dis 1996;27:525-532.
10. Dasselaar JJ, Huisman RM, de Jong PE, et al. Measurement of relative blood volume changes during haemodialysis: merits and limitations. Nephrol Dial Transplant 2005;20:2043-2049.
11. Balter P, Artemyev M, Zabetakis P. Methods and challenges for the practical application of Crit-Line monitor utilization in patients on hemodialysis. Blood Purif 2015;39:21-24.
12. Sibbel SP, Ficociello LH, Black M, et al. Effects of Crit-Line(R) monitor use on patient outcomes and epoetin alfa dosing following onset of hemodialysis: a propensity score- matched study. Blood Purif 2014;37:249-257.
13. Yoshida I, Ando K, Ando Y, et al. A new device to monitor blood volume in hemodialysis patients. Ther Apher Dial 2010;14:560-565.
14. Ristevski B, Chen M. Big Data Analytics in Medicine and Healthcare. J Integr Bioinform 2018;
15. Amazon. AWS HIPAA Eligible Services. Amazon 2020; (Accessed March 2020): https://aws.amazon.com/compliance/hipaa-eligible-services-reference/.
16. Enterprise HP. Edge Device. Hewlett Packard Enterprise 2020; (Accessed March 2020):https:// www.hpe.com/us/en/servers/edgeline-systems.html.
17. Amazon. What is AWS. Amazon 2020; (Accessed March 2020): https://aws.amazon.com/what-is-aws/.
18. Microsoft. Microsoft Azure. Microsoft 2020; (Accessed March 2020): https://azure.microsoft.com/en-us/
19. Google. Google Cloud. Google 2020; (Accessed June 2020): https://cloud.google.com/.
20. America FMCN. Crit-Line IV Monitor. 2020; (Accessed May 2020): https://fmcna.com/products/fluid-management/crit-line-iv-monitor/.
21. Barth C, Boer W, Garzoni D, et al. Characteristics of hypotension-prone haemodialysis patients: is there a critical relative blood volume? Nephrol Dial Transplant 2003;18:1353- 1360.
22. Rodriguez HJ, Domenici R, Diroll A, et al. Assessment of dry weight by monitoring changes in blood volume during hemodialysis using Crit-Line. Kidney Int 2005;68:854-861.
23. Santoro A, Mancini E, Paolini F, et al. Blood volume regulation during hemodialysis. Am J Kidney Dis 1998;32:739-748.
24. Santoro A, Mancini E, Basile C, et al. Blood volume controlled hemodialysis in hypotension- prone patients: a randomized, multicenter controlled trial. Kidney Int 2002;62:1034-1045.
25. Preciado P, Zhang H, Thijssen S, et al. All-cause mortality in relation to changes in relative blood volume during hemodialysis. Nephrol Dial Transplant 2019;34:1401-1408.
26. Amazon. SageMaker. Amazon 2020; (Accessed March 2020): https://aws.amazon.com/sagemaker/.
27. Sarica A, Cerasa A, Quattrone A. Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. Front Aging Neurosci 2017; 9: 329.
28. Noble WS. What is a support vector machine? Nat Biotechnol 2006; 24: 1565-1567.

29. Chen T GC. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining San Francisco, California, USA, Association for Computing Machinery 2016:785–794.
30. Dernoncourt F, Nemati S, Kassis EB, et al. Hyperparameter Selection. Secondary Analysis of Electronic Health Records: Cham (CH), 2016:419-427.
31. Amazon. How Hyperparameter Tuning Works. Amazon 2021; (Accessed June 2021): https://docs.aws.amazon.com/sagemaker/latest/dg/automatic-model-tuning-how-it- works.html.
32. Chao Qin YZ, Fangxun Bao, Caiming Zhang, Peide Liu, Peipei Liu. XGBoost Optimized by Adaptive Particle Swarm Optimization for Credit Scoring". Mathematical Problems in Engineering 2021; 2021, Article ID 6655510 (Accessed May 2021): https://doi.org/10.1155/2021/6655510.
33. J. Snoek HL, and R. P. Adams. Practical bayesian optimization of machine learning algorithms. Advances in Neural Information Processing Systems, vol 4 2012; 4 (Accessed May 2021): https://arxiv.org/ abs/1206.2944.
34. Tjeng Wawan Cenggoro B, ArifBudiarto, JamesBaurley,TeddySuparyanto, BensPardamean. Features Importance in Classification Models for Colorectal Cancer Cases Phenotype in Indonesia. Procedia Computer Science 2019; Volume 157, 2019 (Accessed May 2021): 313-320 https://www. sciencedirect.com/science/article/pii/S1877050919310907.
35. Al-Hilali N, Al-Humoud HM, Ninan VT, et al. Profiled hemodialysis reduces intradialytic symptoms. Transplant Proc 2004; 36: 1827-1828.
36. Barbieri C, Cattinelli I, Neri L, et al. Development of an Artificial Intelligence Model to Guide the Management of Blood Pressure, Fluid Volume, and Dialysis Dose in End-Stage Kidney Disease Patients: Proof of Concept and First Clinical Assessment. Kidney Dis (Basel) 2019; 5: 28-33.
37. Kriegeskorte N, Golan T. Neural network models and deep learning. Curr Biol 2019;29:R231-R236.
38. McIntyre CW. Effects of hemodialysis on cardiac function. Kidney Int 2009;76:371-375.
39. Eldehni MT, Odudu A, McIntyre CW. Brain white matter microstructure in end-stage kidney disease, cognitive impairment, and circulatory stress. Hemodial Int 2019;23:356-365.
40. McIntyre C, Crowley L. Dying to Feel Better: The Central Role of Dialysis-Induced Tissue Hypoxia. Clin J Am Soc Nephrol 2016;11:549-551.
41. Flythe JE, Xue H, Lynch KE, et al. Association of mortality risk with various definitions of intradialytic hypotension. J Am Soc Nephrol 2015;26:724-734.
42. Assimon MM, Wang L, Flythe JE. Cumulative Exposure to Frequent Intradialytic Hypotension Associates With New-Onset Dementia Among Elderly Hemodialysis Patients. Kidney Int Rep 2019;4:603-606.
43. 2018 Common Rule, Title 45, Code of Federal Regulations, Part 46, Protection of Human Subjects. Department of Health and Human Services 19 June 2018: (Accessed 13 Jul 2021):https://www.hhs.gov/ohrp/regulations-and-policy/regulations/2045-cfr-2046/revised- common-rule-regulatory-text/index.html#2046.2116(f).

Chapter 5

Remote Treatment Monitoring on Hospitalization and Technique Failure Rates in Peritoneal Dialysis Patients

Sheetal Chaudhuri, Hao Han, Carlos Muchiutti,
Jessica Ryter, Marta Reviriego-Mendoza, Dugan Maddux,
John W Larkin, Len A Usvyat, Dinesh Chatoth,
Jeroen P Kooman, Franklin W Maddux

Kidney360 2020;1(3):191-202

Abstract

Background

An integrated kidney disease healthcare company implemented a peritoneal dialysis (PD) remote treatment monitoring (RTM) application in 2016. We assessed if RTM utilization associates with hospitalization and technique failure rates.

Methods

We used data from adult PD patients (age ≥18 years) treated from Oct 2016 through May 2019 who registered online for the RTM. Patients were classified by RTM use during a 30-day baseline after registration. Groups were: non-users (never entered data), moderate users (entered 1-to-15 treatments), and frequent users (entered >15 treatments). We compared hospital admission/day and sustained technique failure (required >6 consecutive weeks of hemodialysis) rates over 3, 6, 9, and 12 months of follow-up using Poisson and Cox models adjusted for patient/clinical characteristics.

Results

Among 6,343 patients, 64.5% were non-users, 10.6% were moderate-users, and 24.9% were frequent- users. Incidence rate of hospital admission was 22% (incidence rate ratio (IRR)=0.775; p=0.002), 24% (IRR=0.762; p<0.001), 23% (IRR=0.768; p=<0.001), and 26% (IRR=0.737; p=<0.001) lower in frequent-users after 3, 6, 9, and 12 months respectively versus non-users. Incidence rate of hospital days was 38% (IRR=0.618; p=0.013), 35% (IRR=0.654; p=0.001), 34% (IRR= 0.657; p=<0.001), and 32% (IRR=0.680; p<0.001) lower in frequent-users after 3, 6, 9, and 12 months versus non-users. Sustained technique failure risk at 3, 6, 9, and 12 months was 33% (hazard ratio (HR)=0.671; p=0.020), 31% (HR=0.686; p=0.003), 31% (HR=0.687; p=0.001), and 27% (HR=0.726; p=0.001) lower in frequent-users versus non-users. Among a sub-group of survivors of the 12-month follow-up, sustained technique failure risk was 26% (HR=0.736; p=0.023) and 21% (HR=0.793; p=0.054) lower after 9 and 12 months in frequent-users versus non-users.

Conclusions

Our findings suggest frequent use of a RTM application associates with less hospital admissions, shorter hospital length of stay, and lower technique failure rates. Adoption of RTM applications may have the potential to improve timely identification/ intervention of complications.

Introduction

The modality of peritoneal dialysis (PD) is suggested to associate with favorable outcomes compared to in-center hemodialysis (HD), however, technique failure is common and adjusted rates for the hospital length of stay tends to be longer in PD.[1-3] In the United States, home dialysis patients typically have monthly clinic visits where their clinical status, assessment of treatment quality and adherence is assessed based on written treatment records and self-reported complications.[4,5] This monitoring process poses limits to the clinicians view of complications and clinical needs, which can impede timely medical decisions.

Remote monitoring systems may improve the care teams ability to actively identify urgent concerns in PD patients and react in a timely manner with diagnostic examinations, interventions, or patient education.[6] Remote monitoring systems broadly include an array of technologies and processes in the areas of telemedicine/telehealth such as telephonic/video assessments, connected health sensors, and health record portals.[6] Prior reports of various types of remote monitoring systems tested in the PD population suggest they associate with improvements in patient satisfaction, quality of life, nutrition, exit site infection rates, peritonitis rates, and hospitalization rates.[7-13] However, patient groups included in the evaluations of remote monitoring systems to date are small and the findings may not be generalizable.

A large integrated kidney disease healthcare company started using a remote treatment monitoring (RTM) application throughout its dialysis organization in the United States since October 2016. The provider constructed and integrated the RTM system into its clinical systems as a quality improvement process. The RTM is an online portal-based treatment record application, whereby patients can create an online account and record the details of individual PD treatments, associated vital signs, and complications. We aimed to assess the associations between the level of utilization of the RTM application and hospitalization and technique failure rates.

Materials and methods

General design

We performed a retrospective analysis to assess the longitudinal associations between the frequency of use of the RTM in PD patients and outcomes after deployment in a dialysis organization in the United States. This analysis was performed under a protocol that was reviewed by New England Independent Review Board who determined it was an exempt assessment of existing patient data from a quality improvement process,

which was anonymized and did not require informed consent per title 45 of the United States Code of Federal Regulations part 46.102 (Needham Heights, MA, United States; NEIRB# 1- 9652-1). The analysis was conducted in adherence with the Declaration of Helsinki.

Patient population

We used data from adult PD patients (age ≥18 years) treated anytime during 01 Oct 2016 to 31 May 2019 at the dialysis organization (Fresenius Kidney Care, Waltham, MA, United States) of a large integrated kidney disease healthcare company (Fresenius Medical Care, Bad Homburg, Germany). We included data from all PD patients who: 1) registered online and created a RTM account on, or before 31 May 2018, 2) were treated continuously with PD for at least 30 days after registration, and 3) were not hospitalized within 30 days after registration. We excluded patients with: 1) a body mass index (BMI) >65 Kg/m^2, or 2) missing data for any covariates used for the adjustment of the analysis (refer to the variables section of the materials and methods).

RTM application

The RTM named the "PatientHub" is a health record portal technology available to all PD patients treated by the dialysis organization in the United States who have internet access. The RTM can be used through either a secure personal website portal, or mobile device application. After the creation of a personal RTM account and profile, patients can: 1) view their dialysis orders, laboratory results, concomitant medications, and supply orders, and 2) document their daily PD treatment data, vital signs, and complications. A schematic of the prior monitoring process and RTM process is shown in **Figure 5.1A and 5.1B**.

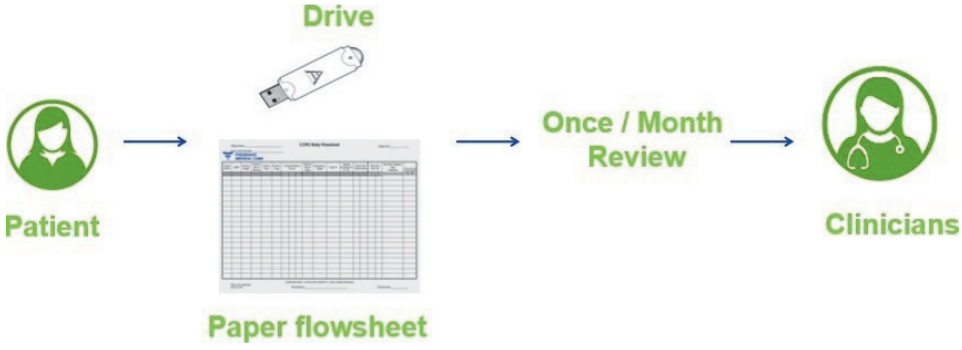


Figure 5.1A Schematic of how the paper flowsheets were submitted manually for review by clinicians once a month.

Figure 5.1B Schematic of how electronic flowsheets are submitted via PatientHub RTM for daily review by clinicians.

The daily PD treatment data documented in the RTM includes the treatment date, total ultrafiltration from cycler, dialysate type, bag size and number of bags used, medications added to the dialysate, as well as details on any manual exchanges performed. The daily clinical data documented includes weight, blood pressure, pulse, temperature, blood glucose, and confirmation of routine exit site care. The data on complications documented daily includes drain/fill, PD fluid, or exit site issues. Once submitted by the patient, the data is displayed in the electronic medical record (EMR) for the care team to review. The care team is instructed to review patient RTM entries on at least a daily basis on business days, but the timing of the daily review is at the discretion of the clinician. A preview of the PatientHub RTM application is shown in **Figure 5.2**.

Variables

The dependent variables were hospital admission counts per patient year (PPY), hospital days PPY, and sustained technique failure counts (PPY) from 30 days after RTM registration to 3, 6, 9, and 12 months of follow-up. Sustained technique failure was defined as PD complications that required patients to receive >6 consecutive weeks of treatment with HD.

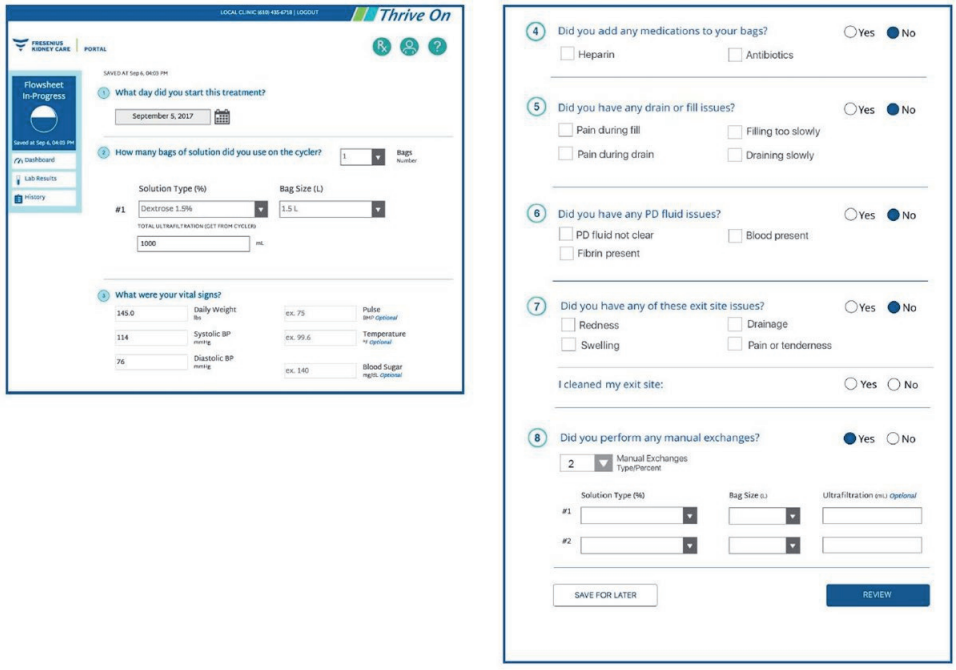


Figure 5.2 Preview of the PatientHub RTM application.

The independent variables were the frequency of RTM use during the baseline period 30 days after online registration. We defined non-users as patients who never documented any treatment record in the RTM within 30 days of registration, moderate-users as patients who documented 1 to 15 treatment records in the RTM within 30 days of registration, and frequent-users as patients who documented more than 15 treatment records in the RTM within 30 days of registration. These cut points for classification of the frequency of baseline RTM use were chosen based on the distribution of data; most patients tended to use the RTM more than 15 times, or never within 30 days of registration.

Covariates used for the description of the baseline patient characteristics and for adjustment of statistical models included age, sex, race, ethnicity, dialysis vintage, alcohol use, urbanicity of residence, education level, congestive heart failure, diabetes, ischemic heart disease, albumin, residual kidney function (RKF), and weekly Kt/V. The most recent categorical variables to the date of registration were recorded. The mean value of continuous variables during 30 days from RTM registration was computed and recorded, with exception of RKF that was determined from the most recent value.

Statistical methods

Adjusted Poisson regression models were constructed to assess the associations in hospital admission and day rates at the 3, 6, 9, and 12 months of follow-up periods for non-users (reference) versus moderate-users and frequent-users. The number of days patients were actively receiving dialysis during the follow-up period (patient exposure days) was used to calculate the hospital admission and day rates PPY.

An adjusted Cox regression model was constructed to assess the associations in sustained technique failure rates at the 3, 6, 9, and 12 months of follow-up periods for non-users versus moderate-users and frequent-users. A Kaplan Meier curve with log-rank tests was used to assess the associations in the time on PD modality over the 12 months follow-up for non-users versus moderate-users and frequent-users.

For hospitalization and technique failure rates, we censored data on patients who died, received a transplant, or where discharged from the providers clinic network at end of each respective 3-month period of the 12-month follow-up.

Like the previous analysis, adjusted Cox regression methods were used in a sub-analysis of survivors of the 12-month follow-up period to assess the associations in technique failure rates and the time on PD modality in a group of patients who had equivalent opportunities to experience a technique failure. This sub-analysis of survivors excluded all attrition (patients who died, received a transplant, or where discharged from the providers clinic network).

Results

Patient characteristics

In a population of 36,577 PD patients treated at a large dialysis provider during the analysis period, 11,079 adult patients were treated by PD for 30 days without being hospitalized during baseline and completed the online registration creating a RTM profile. Among this cohort, we included data from 6,343 patients at 931 clinics, and excluded 4,736 patients due to incomplete/missing data on covariates used in the adjustment of the analysis (**Figure 5.3; Table 5.1**). Among eligible patients (n=6,343), 64.5% never entered treatment data (non-users), 10.6% entered 1 to 15 treatment records (moderate- users), and 24.9% entered more than 15 treatment records (frequent-users) during a 30 day baseline period after registration. Patients mean age ranged from 54 to 58 years old between groups (**Table 5.2**). Frequent-users tended to more often be of a white race, non-Hispanic ethnicity, educated, and had a shorter dialysis vintage. The within group trends of RTM usage were fairly sustained over the

12-month follow-up, as compared to baseline, yet frequent-users had decreases in the mean number of entries (**Figure 5.4**). On the 12th month of follow up, frequent-users on average documented 10 treatments in the RTM.

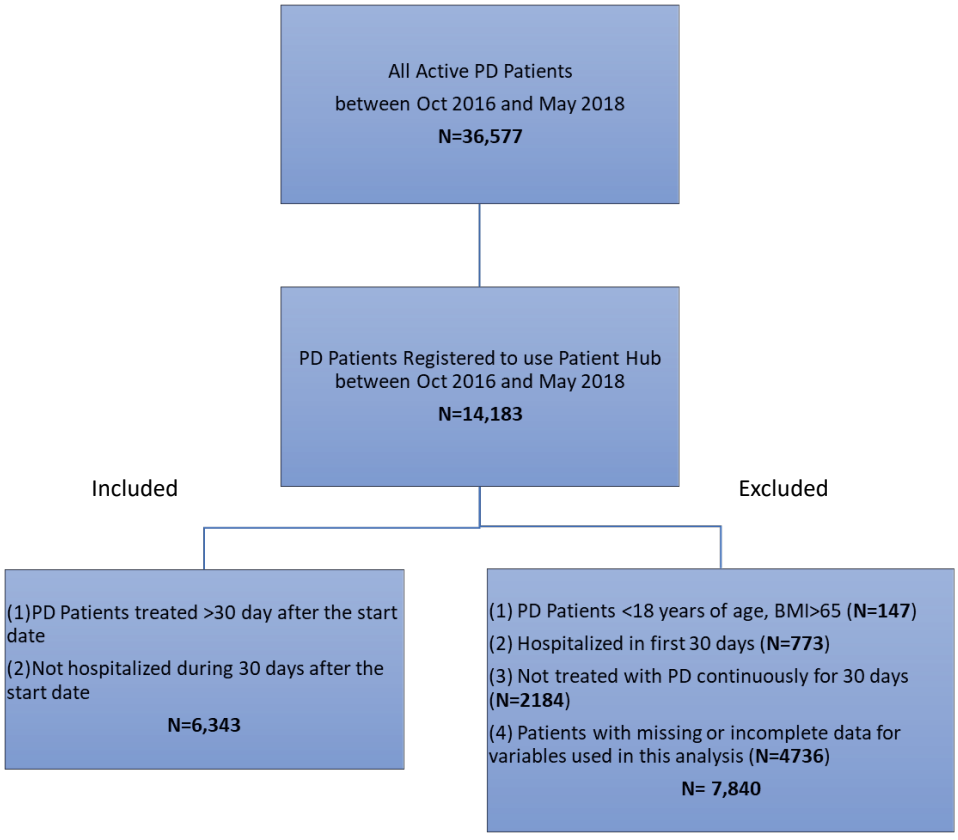


Figure 5.3 Patient flow diagram.

Table 5.1 Comparison of baseline characteristics among participants included versus patients excluded due to only missing data for adjustment of event analysis.

Cohort	Included	Excluded due to missing variables for adjustment	Included versus excluded p-value
Patient count	6343	4736	
Demographics			
Age (yrs)	56.9±15.2	57.2±14.7	0.310
Males (%)	57%	57%	0.933
Black race (%)	23%	25%	0.003
White race (%)	72%	68%	<0.001
Other race (%)	5%	7%	0.008
Hispanic ethnicity (%)	10%	11%	0.022
Dialysis vintage (days)	690	472	<0.001
Alcohol use (%)	57%	57%	0.934
Urbanicity			
Metropolitan (%)	81%	83%	0.028
Micropolitan (%)	11%	10%	0.146
Rural (%)	8%	7%	0.131
Education			
College or beyond (%)	58%	58%	0.972
High school or equivalency (%)	33%	32%	0.654
Less than high school or equivalency (%)	9%	10%	0.507
Comorbidities			
Congestive heart failure (%)	11%	11%	0.666
Diabetes (%)	54%	54%	0.983
schemic heart disease (%)	13%	11%	0.028
Laboratory			
Albumin (g/dL)	3.55±0.45	3.58±0.47	0.002
Residual kidney function (mL/min)	4.26±3.32	4.67±4.69	<0.001
Weekly Kt/V	2.43±1.05	2.39±1.04	0.113

Comparisons between groups were made using t-tests.

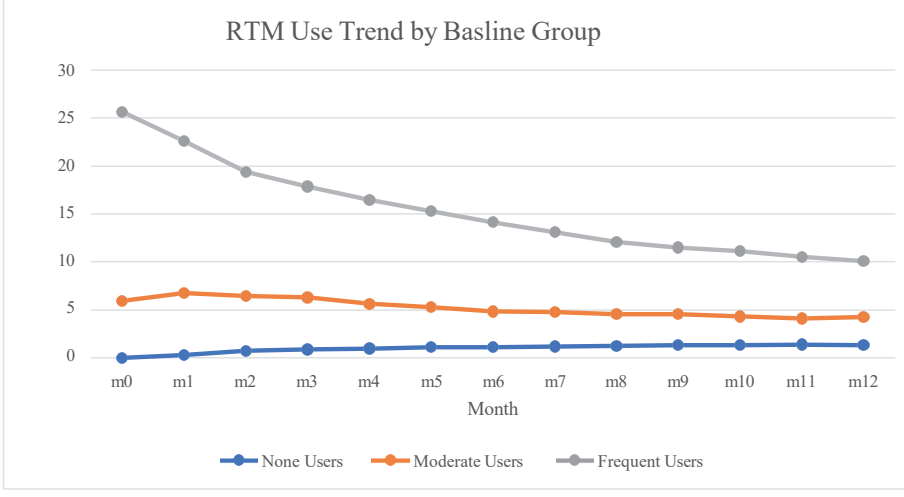


Figure 5.4 Trend in the mean number of RTM entries in each month of the follow-up period by baseline RTM use group category.

Longitudinal hospitalization rates associated with RTM use

Higher RTM usage in the 30 days following the start date was found to be associated with progressively lower unadjusted hospital admission and day rates in the follow up periods (**Figures 5.5a and Figure 5.5b**). A Poisson analysis showed the incidence rate of hospital admission was 22%, 24%, 23% and 26% lower in frequent users of the RTM application after 3, 6, 9, and 12 months of follow-up, as compared to non- users (**Table 5.3**). The incidence rate of a greater hospital length of stay in days was found to be 38%, 35%, 34% and 32% lower in frequent-users of the RTM application after 3, 6, 9, and 12 months of follow-up versus non-users (**Table 5.4**). Albeit qualitative differences were observed favoring moderate-use of the RTM, there were not significant differences compared with non-users.

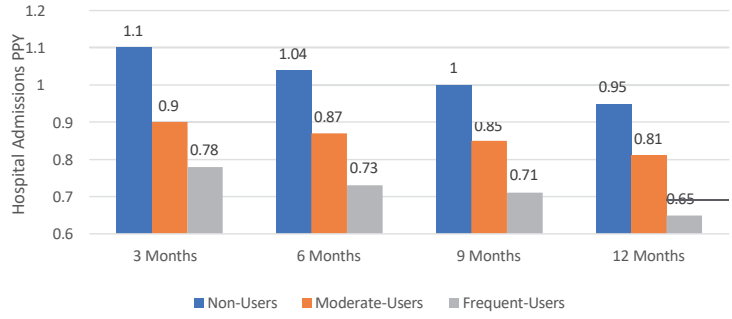


Figure 5.5A Hospital admission rate by baseline RTM use after 3, 6, 9, 12 months of follow-up.

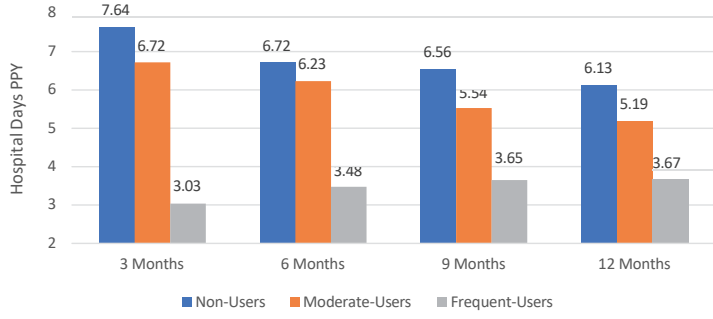


Figure 5.5B Hospital day rate by baseline RTM use after 3, 6, 9, 12 months of follow-up.

Table 5.3 Associations in adjusted hospital admission rates by baseline remote treatment monitoring (RTM) use.

	3 Month admission count		6 Month admission count		9 Month admission count		12 Month admission count	
Parameter	IRR Estimate	95% CI	IRR Estimate	95% CI	IRR Estimate	95% CI	IRR Estimate	95% CI
Moderate-users	0.907	0.736-1.120	0.922	0.775-1.098	0.937	0.795-1.105	0.933	0.803-1.085
Frequent-users	0.775	0.662-0.906	0.762	0.669-0.870	0.768	0.678-0.871	0.737	0.657-0.829
Age	1.001	0.996-1.005	1.001	0.997-1.005	1.001	0.997-1.004	1.001	0.998-1.004
Females	1.070	0.934-1.226	1.131	1.009-1.268	1.057	0.947-1.178	1.062	0.961-1.172
Black race	0.989	0.849-1.153	0.970	0.854-1.102	0.890	0.787-1.006	0.878	0.785-0.983
Other race	0.770	0.559-1.059	0.714	0.543-0.939	0.665	0.511-0.865	0.598	0.465-0.770
Hispanic ethnicity	1.007	0.813-1.247	0.968	0.808-1.160	0.870	0.737-1.038	0.874	0.744-1.026
Comorbidity: Congestive heart failure	1.302	1.094-1.548	1.351	1.170-1.560	1.302	1.133-1.498	1.228	1.078-1.399
Comorbidity: Diabetes	1.244	1.090-1.417	1.237	1.109-1.379	1.215	1.096-1.348	1.239	1.126-1.362
Comorbidity: Ischemic heart disease	1.229	1.037-1.456	1.158	1.003-1.338	1.157	1.008-1.328	1.153	1.015-1.308
Albumin	0.508	0.444-0.581	0.513	0.458-0.574	0.516	0.463-0.575	0.518	0.469-0.573
Residual kidney function	0.982	0.953-1.012	0.984	0.959-1.009	0.977	0.954-1.001	0.978	0.957-1.000
Weekly Kt/V	0.912	0.802-1.038	0.878	0.783-0.985	0.897	0.806-0.999	0.912	0.829-1.004
Dialysis vintage	1.000	1-1.0001	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000
Education: College or beyond	0.902	0.731-1.114	1.226	0.820-1.178	1.006	0.845-1.196	1.138	0.873-1.200
Education: High school or equivalency	0.988	0.795-1.227	0.963	0.892-1.295	1.088	0.910-1.300	0.999	0.921-1.278
Alcohol dependency: Declined to answer	0.695	0.237-2.032	0.983	0.270-1.655	0.941	0.457-1.938	1.023	0.597-2.016
Alcohol dependency: No	0.977	0.846-1.129	1.075	0.835-1.060	0.973	0.868-1.091	1.084	0.842-1.037
Urbanicity: Metropolitan	1.201	0.951-1.517	0.669	1.006-1.494	1.162	0.965-1.400	1.097	0.960-1.347
Urbanicity: Micropolitan	0.908	0.673-1.227	0.940	0.749-1.237	1.004	0.796-1.267	0.934	0.808-1.234

Reference groups: Non-users, males, white race; Education: no high school; Alcohol dependency: yes, Urbanicity: Rural. IRR = Incidence Rate Ratio. Hospital admission rates are adjusted by all paraeters listed in the table.

Table 5.4 Associations in adjusted hospital day rates by baseline Remote Treatment Monitoring (RTM) use.

Parameter	3 Month day count		6 Month day count		9 Month day count		12 Month day count	
	IRR Estimate	95% CI	IRR Estimate	95% CI	IRR Estimate	95% CI	IRR Estimate	95% CI
Moderate-users	0.893	0.559-1.426	0.951	0.692-1.307	0.887	0.663-1.185	0.883	0.677-1.152
Frequent- users	0.618	0.424-0.902	0.654	0.505-0.846	0.657	0.524-0.824	0.680	0.554-0.834
Age	1.005	0.994-1.015	1.005	0.998-1.012	1.005	0.999-1.011	1.006	1.001-1.012
Females	1.055	0.781-1.427	1.132	0.913-1.404	0.992	0.829-1.186	0.979	0.830-1.154
Black race	0.646	0.443-0.944	0.767	0.598-0.985	0.734	0.588-0.915	0.733	0.599-0.898
White race	0.537	0.244-1.183	0.567	0.329-0.976	0.492	0.298-0.812	0.448	0.277-0.723
Hispanic ethnicity	1.034	0.647-1.651	0.971	0.695-1.359	0.886	0.656-1.196	0.895	0.680-1.178
Comorbidity: Congestive heart failure	1.310	0.886-1.935	1.331	1.017-1.743	1.311	1.032-1.667	1.249	1.001-1.559
Comorbidity: Diabetes	1.391	1.032-1.873	1.452	1.180-1.787	1.340	1.119-1.607	1.305	1.107-1.538
Comorbidity: Ischemic heart disease	1.110	0.754-1.632	1.077	0.823-1.410	1.067	0.840-1.354	1.068	0.859-1.328
Albumin	0.519	0.383-0.705	0.517	0.418-0.460	0.480	0.398-0.578	0.482	0.406-0.571
Residual kidney function	0.940	0.878-1.005	0.968	0.919-1.018	0.942	0.909-0.976	0.947	0.916-0.979
Weekly Kt/V	0.962	0.752-1.231	0.861	0.685-1.080	1.006	0.923-1.098	0.994	0.903-1.095
Dialysis vintage	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000
Education: College or beyond	1.293	0.778-2.151	1.293	0.903-1.851	1.310	0.951-1.803	1.303	0.976-1.740
Education: High school or equivalency	1.146	0.674-1.946	1.217	0.838-1.763	1.283	0.922-1.783	1.232	0.914-1.661
Alcohol dependency: Declined to answer	0.354	0.011-10.943	0.393	0.041-3.774	0.535	0.098-2.916	0.934	0.288-3.036
Alcohol dependency: No	0.890	0.646-1.226	0.899	0.720-1.122	0.986	0.808-1.204	0.979	0.817-1.175
Urbanicity: Metropolitan	1.257	0.727-2.173	1.287	0.878-1.884	1.265	0.903-1.773	1.157	0.859-1.559
Urbanicity: Micropolitan	1.110	0.565-2.175	1.153	0.722-1.840	1.186	0.789-1.784	1.164	0.812-1.669

Reference groups: Non-users, males, other race; Education: no high school; Alcohol dependency: yes, Urbanicity: Rural. IRR = Incidence Rate Ratio. Hospital days rates are adjusted by all parameters listed in the table.

Longitudinal technique failure rates associated with RTM use

We observed higher RTM usage in the 30 days following the start date was associated with lower rates of sustained PD technique failure (i.e. required >6 consecutive weeks of treatment with HD) in all follow up periods (**Figure 5.6**). A Cox analysis showed the adjusted risk of sustained PD technique failure at 3, 6, 9, and 12 months of follow-up was 33%, 31%, 31% and 27% lower in frequent-users of the RTM versus non-users (**Table 5.5**). Kaplan–Meier estimate for PD duration days without sustained PD technique failure identified frequent-users remained on PD longer compared to non-users (log-rank test frequent-users p=0.024; **Figure 5.7A**). No significant differences were found in sustained technique failure rates in moderate-users compared to non-users.

A Cox analysis of a subgroup patients who survived the entire 12 month follow-up period and continued to be treated in the providers clinics confirmed higher RTM usage was associated with 26% and 21% lower adjusted risk of sustained PD technique failure in frequent-users versus non-users at the 9 and 12 month follow up periods (**Table 5.6**). However, we did not find significant differences between moderate- users and non-users of the RTM. In this subgroup of survivors, we also observed the adjusted risk of sustained technique failure was 51% and 42% lower in moderate-users of the RTM at 3 and 6 months of follow-up compared to non-users, yet no significant differences were found at later timepoints. The Kaplan–Meier estimate of PD duration days without sustained PD technique failure showed that frequent-users of the RTM in the survivor subgroup remained on a PD modality longer than non-users (log-rank test frequent-users p<0.001; **Figure 5.7B**).

Table 5.5 Associations in adjusted sustained PD technique failure rates by baseline Remote Treatment Monitoring (RTM) use.

	3-month PD technique failure		6-month PD technique failure		9-month PD technique failure		12-month PD technique failure	
Parameter	Hazard Ratio	95% CI	Hazard Ratio	95% CI	Hazard Ratio	95% CI	Hazard Ratio	95% CI
Moderate-users	0.657	0.403-1.071	0.763	0.544-1.072	0.852	0.644-1.127	0.869	0.673-1.123
Frequent-users	0.671	0.481-0.938	0.686	0.536-0.878	0.687	0.555-0.850	0.726	0.598-0.881
Age	1.000	0.991-1.009	1.001	0.994-1.008	0.997	0.991-1.003	0.997	0.991-1.002
Females	0.898	0.674-1.197	0.912	0.735-1.131	0.825	0.683-0.996	0.824	0.693-0.979
Black race	0.777	0.556-1.088	0.917	0.723-1.163	0.908	0.739-1.116	0.908	0.752-1.096
Other race	0.390	0.160-0.953	0.516	0.289-0.923	0.612	0.385-0.972	0.560	0.360-0.869
Hispanic ethnicity	0.782	0.489-1.250	0.880	0.625-1.240	0.967	0.738-1.285	0.966	0.745-1.252
Comorbidity: Congestive heart failure	1.267	0.865-1.856	1.240	0.935-1.645	1.188	0.924-1.527	1.131	0.894-1.432
Comorbidity: Diabetes	1.104	0.843-1.447	1.105	0.906-1.349	1.064	0.896-1.263	1.109	0.947-1.298
Comorbidity: Ischemic heart disease	0.918	0.622-1.355	1.009	0.762-1.336	1.062	0.833-1.355	0.988	0.786-1.241
Albumin	0.640	0.479-0.854	0.640	0.516-0.794	0.693	0.574-0.837	0.681	0.572-0.810
Residual kidney function	1.033	0.989-1.078	1.016	0.974-1.061	1.009	0.970-1.049	1.013	0.978-1.050
Weekly Kt/V	0.580	0.436-0.771	0.578	0.456-0.733	0.670	0.547-0.822	0.673	0.559-0.810
Dialysis vintage	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000
Urbanicity: Metropolitan	0.803	0.531-1.214	0.932	0.671-1.295	0.898	0.676-1.193	0.915	0.703-1.193
Urbanicity: Micropolitan	0.628	0.354-1.116	0.743	0.478-1.153	0.745	0.511-1.086	0.807	0.572-1.139
Education: College or beyond	0.668	0.445-1.003	0.850	0.614-1.177	0.941	0.706-1.255	0.946	0.728-1.229
Education: High school or equivalency	0.744	0.488-1.134	0.929	0.664-1.301	0.991	0.736-1.334	0.963	0.734-1.265
Alcohol dependency: Declined to answer	0.000	0.000-1.18E+187	0.000	0.000-1.87E+135	0.000	0.000-1.02E+191	0.000	0.000-3.05E+172
Alcohol dependency: No	0.837	0.629-1.114	0.823	1.016	0.852	0.709+1.023	0.811	0.686-0.957

Reference groups: non-users, males, white race; Education: no high school, Alcohol dependency: yes, Urbanicity: Rural. Sustained PD technique failure rates are adjusted by all parameters listed in the table.

Table 5.6 Associations in adjusted sustained PD technique failure rates by baseline Remote Rreatment Monitoring (RTM) use among survivors of the 12 month follow-up.

Parameter	3 Month technique failure		6 Month technique failure		9 Month technique failure		12 Month technique failure	
	Hazard Ratio	95% CI	Hazard Ratio	95% CI	Hazard Ratio	95% CI	Hazard Ratio	95% CI
Moderate-users	0.487	0.236-1.003	0.578	0.351-0.952	0.811	0.566-1.163	0.825	0.594-1.145
Frequent-users	0.749	0.494-1.134	0.744	0.546-1.015	0.736	0.566-0.958	0.793	0.627-1.004
Age	0.996	0.984-1.008	0.998	0.989-1.007	0.993	0.985-1.000	0.991	0.984-0.997
Females	0.890	0.617-1.284	0.836	0.633-1.103	0.737	0.582-0.933	0.759	0.612-0.941
Black race	0.718	0.465-1.109	0.842	0.620-1.143	0.858	0.664-1.109	0.840	0.664-1.063
Other race	0.354	0.112-1.120	0.455	0.213-0.972	0.513	0.280-0.41	0.461	0.258-0.824
Hispanic ethnicity	0.888	0.513-1.538	0.808	0.524-1.247	0.960	0.681-1.354	0.998	0.733-1.360
Comorbidity: Congestive heart failure	1.107	0.646-1.897	1.044	0.700-1.556	1.092	0.783-1.523	0.990	0.722-1.357
Comorbidity: Diabetes	1.003	0.711-1.415	1.157	0.896-1.495	1.076	0.869-1.333	1.111	0.913-1.351
Comorbidity: Ischemic heart disease	0.926	0.554-1.547	0.977	0.674-1.415	1.076	0.792-1.463	1.077	0.812-1.428
Albumin	0.556	0.382-0.807	0.587	0.444-0.776	0.650	0.512-0.824	0.641	0.516-0.796
Residual kidney function	1.033	0.985-1.083	1.021	0.972-1.071	1.016	0.973-1.061	1.020	0.982-1.060
Weekly Kt/V	0.574	0.405-0.815	0.572	0.428-0.766	0.698	0.550-0.886	0.676	0.544-0.841
Dialysis vintage	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000	1.000	1.000-1.000
Urbanicity: Metropolitan	0.702	0.419-1.175	0.844	0.556-1.281	0.793	0.561-1.120	0.800	0.580-1.102
Urbanicity: Micropolitan	0.463	0.217-0.988	0.695	0.399-1.209	0.712	0.452-1.120	0.766	0.507-1.157
Education: College or beyond	0.694	0.413-1.169	0.926	0.602-1.423	1.011	0.702-1.455	1.052	0.754-1.467
Education: High school or equivalency	0.777	0.452-1.335	1.065	0.685-1.658	1.073	0.737-1.562	1.062	0.752-1.499
Alcohol dependency: Declined to answer	0.000	0.000->1E+10	0.000	0.000->1E+10	0.000	0.000->1E+10	0.000	0.000->1E+10
Alcohol dependency: No	0.841	0.582-1.215	0.862	0.656-1.135	0.846	0.674-1.063	0.782	0.637-0.960

Reference groups: non-users, males, white race; Education: no high school, Alcohol dependency: yes, Urbanicity: Rural. Sustained PD technique failure rates are adjusted by all parameters listed in the table.

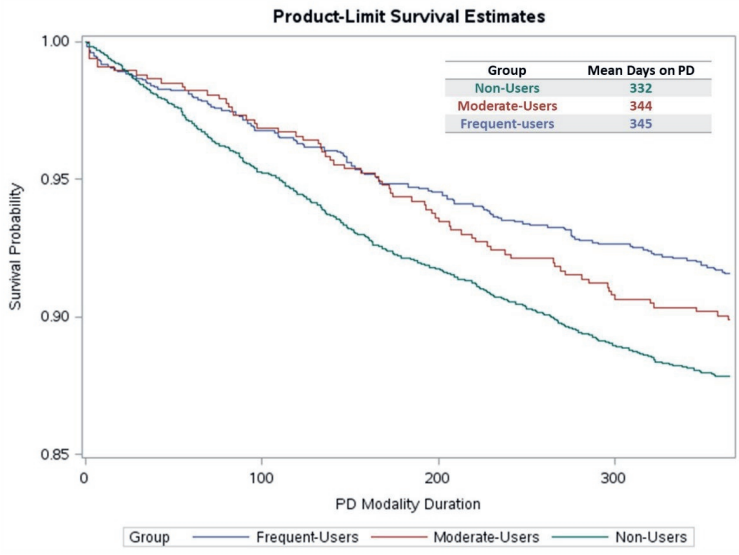


Figure 5.7A Kaplan Meier curve plot to assess the associations in the time on PD modality over the 12 months follow-up for non-users, moderate-users and frequent-users. Frequent-users of the RTM remained on a PD modality 13 days longer in comparison to the non-users group.

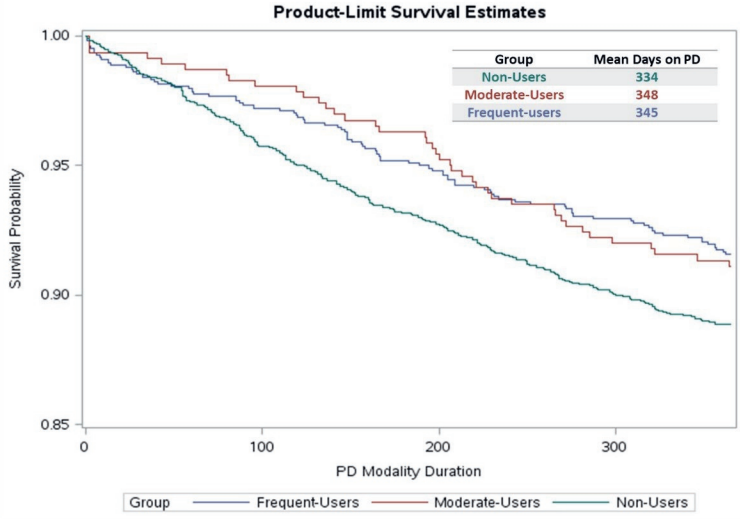


Figure 5.7B Kaplan-Meier plot for survivors of the 12-month follow-up period to assess the associations in technique failure rates and the time on PD modality in a group of patients who had equivalent opportunities to experience a technique failure. Frequent-users of the RTM in the survivor subgroup remained on a PD modality 11 days longer than non-users.

Discussion

In a large population of PD patients who registered online for the PatientHub RTM application, we found higher RTM use was associated with lower hospitalization and sustained technique failure rates.

Hospital admission and day rates were observed to be 22% and 38% lower within 3 months, respectively, and continued to decrease in frequent-users of the RTM over the 12 months of follow-up compared to non-users. The risk of sustained PD technique failure (i.e. required >6 consecutive weeks of HD) was about 30% lower during follow-up for frequent-users versus non-users of the RTM. Assessment of a subgroup of survivors of the 12-month follow-up period found consistent trends, but significant differences were only observed at the 9- and 12-month follow-up periods in frequent-users versus non-users. Consistent trends were seen with respect to moderate-users of the RTM versus non-users, albeit outcomes did not significantly differ with exception of sustained PD technique failure rates in the subgroup of survivors at the 3- and 6-month follow-up periods. These results further substantiate prior findings suggesting use of other RTM systems in PD patients may reduce hospital rates[14,15], and reveal frequent RTM use may also have the potential to increase sustained use of PD as a modality.

The associations between RTM use in PD patients and hard outcomes has been reported in a limited number of small cross-sectional analyses. A study of 63 patients who had a RTM system incorporated in their PD cycler found that the incidence of hospital admissions was 39% (IRR=0.61; 95% confidence interval (CI) 0.39 to 0.95) lower and hospital days was 54% (IRR=0.46; 95% CI 0.23 to 0.92) lower compared to 63 matched patients who did not have a RTM system in their cycler.[14] Another study of 269 PD patients who received an intervention of daily RTM of blood pressure and weight coupled with video conferencing telehealth care showed that the adjusted risk of hospital admission was 46% (OR=0.54; 95% CI 0.33 to 0.89) lower and hospital days was 54% (OR=0.46; 95% CI 0.26 to 0.81) lower compared to before the intervention.[15] It has been estimated that adoption of RTM systems to monitor PD treatments may also yield some economic benefits to the healthcare system in various countries.[13,15]

The influence of RTM on technique failure rates has not been reported previously in PD, yet the improvements we found are consistent with observations from RTM in the home HD population.[16] Notably, technique failure rates in all groups for our analysis were lower than many reports in the literature[17,18], which could be in part due to our definition of a sustained technique failure event that required >6 consecutive weeks of HD and appropriately did not count technique failures as composite outcomes including death. These findings could also be in part representative of patients included in our analysis being a highly select healthier population compared to the overall PD

population. Despite the differences in our technique failure rates with some reports, they are relatively consistent with technique failure rates reported in Japan and Asia that exclude mortality events from the definition.[18]

RTM has been suggested to have the potential to improve patient-care team communications, timely interventions, and patient outcomes in PD for more than a decade.[19] Despite this, the dialysis population has been known to have barriers to access the internet along with inadequate understanding of online systems.[20-22] It is estimated that about 35% to 90% of dialysis patients use the internet.[20-22] In our analysis we included 17% of patients from the overall active PD population based on inclusionary restrictions to have created an online account, age ≥18 years, and BMI <65. Prior reports suggest similar adoption with 18% of PD patients in Columbia using a different remote monitoring system incorporated via connected health sensors in the cycler.[14] Given PD patients are typically younger, it would be expected that there might be a larger proportion of PD patients with access to the internet. If this is a correct assumption, RTM might have the potential to be used in a larger proportion of the PD population, and as smartphone and computer technology advances and becomes more universally affordable, it could become an option for treatment monitoring in most of the patients.

Our analysis design that included a group of PD patients who universally had internet access to create a RTM account, along with temporal assessments of outcomes adjusted for confounding variables related to demographics, urbanicity, education, comorbidities, and laboratories, adds to the strength of these findings. The relative age of patients did not differ in PD patients who were non-users versus frequent- users, which was expected given the inclusion of only patients with internet access and the ability to use online applications. Most of the analysis population (65%) never used the RTM to enter clinical information within 30 days of creating an account online. RTM use groups tended to have a sustained higher or lower pattern of RTM use over time compared to baseline, albeit the frequent-user group had some temporal decreases in entries. Both non-users and moderate-users of the RTM may be specific patient types that could be encouraged and trained to become more active in their care monitoring, and if frequent use is adopted as an adjunct, it may have the potential to yield improvements in outcomes.

Given non-users of RTM were more commonly of a black race, Hispanic, with less education, and a longer dialysis vintage compared to frequent-users, it may be import to target interventions to increase RTM use in patients with these attributes who initially elect to capture their treatment data using an RTM. However, further studies would be needed to test if targeting interventions based on patient profiles would be effective.

The RTM evaluated in this study is an application-based patient and clinician portal that we qualitatively believe required relatively minimal resources to construct, deploy, and maintain, as compared to modem-based systems integrated into cyclers. The flexible ability of the RTM to allow patients to enter records when it works the best for them during/around their treatment, combined with the ability for providers to review patients' entries around their daily workflows may be advantageous attributes of this type of connected health technology. These characteristics of the application-based RTM platform appear to allow it to be rapidly beneficial for the smaller percentage of patients with internet access who adopt RTM and allows for scalability over time.

Although this analysis has many strengths, there are some limitations including the inclusion a sub- group of the PD population who had access to the internet and registered for the RTM online.

Therefore, these findings are not anticipated to be generalizable to PD patients without access to the internet. We excluded patients with missing data on covariates to provided groups with equivalent adjustments for comparisons, yet this made our population have a higher representation of patients with a White race and longer dialysis vintage, among other distinctions. Neither the RTM, nor EMR, captured data on interventions performed due to findings from RTM entries, so we are not able to assess if interventions are being performed in a more timely manner before monthly clinic visits. Also, we cannot rule out that the favorable associations of higher RTM use might be due to patients being more engaged and having a higher health literacy. However, given the inclusion requirements of the design and the adjustments for education, these are not clear confounders. It is possible that more use of the RTM could be a driver influencing engagement and literacy, thereby teaching the patient to become more of a partner in their care. However, this concept would require further investigations.

Conclusions

Our findings suggest frequent use of a RTM application associates with less hospital admissions, shorter hospital length of stay, and lower rates of sustained technique failure requiring HD exposure for more than 6 weeks. It appears prudent for PD care teams and providers to consider adopting RTM applications to better engage patients in their care, recognize and manage potential complications in a timely manner, improve the sustainability on the modality, and improve patient outcomes.

References

1. Saran R, Robinson B, Abbott KC, et al. US Renal Data System 2018 Annual Data Report: Epidemiology of Kidney Disease in the United States. Am J Kidney Dis 2019;73:A7-A8.
2. Lukowsky LR, Mehrotra R, Kheifets L, et al. Comparing mortality of peritoneal and hemodialysis patients in the first 2 years of dialysis therapy: a marginal structural model analysis. Clin J Am Soc Nephrol 2013; 8:619-628.
3. Snyder JJ, Foley RN, Gilbertson DT, et al. Body size and outcomes on peritoneal dialysis in the United States. Kidney Int 2003;64:1838-1844.
4. Mehrotra R, Devuyst O, Davies SJ, et al. The Current State of Peritoneal Dialysis. J Am Soc Nephrol 2016; 27:3238-3252.
5. Tong M, Wang Y, Ni J, et al. Clinical features of patients treated by peritoneal dialysis for over a decade. Am J Clin Exp Urol 2017;5:49-54.
6. Rosner MH, Lew SQ, Conway P, et al. Perspectives from the Kidney Health Initiative on Advancing Technologies to Facilitate Remote Monitoring of Patient Self-Care in RRT. Clin J Am Soc Nephrol 2017; 12:1900-1909.
7. Nayak A, Karopadi A, Antony S, et al. Use of a peritoneal dialysis remote monitoring system in India. Perit Dial Int 2012;32:200-204.
8. Nayak A, Antony S, Nayak KS. Remote monitoring of peritoneal dialysis in special locations. Contrib Nephrol 2012;178:79-82.
9. Nayak KS, Ronco C, Karopadi AN, et al. Telemedicine and Remote Monitoring: Supporting the Patient on Peritoneal Dialysis. Perit Dial Int 2016;36:362-366.
10. Gallar P, Vigil A, Rodriguez I, et al. Two-year experience with telemedicine in the follow-up of patients in home peritoneal dialysis. J Telemed Telecare 2007;13:288-292.
11. Nakamoto H. Telemedicine system for patients on continuous ambulatory peritoneal dialysis. Perit Dial Int 2007;27 Suppl 2:S21-26.
12. Cargill A, Watson AR. Telecare support for patients undergoing chronic peritoneal dialysis. Perit Dial Int 2003;23:91-94.
13. Makhija D, Alscher MD, Becker S, et al. Remote Monitoring of Automated Peritoneal Dialysis Patients: Assessing Clinical and Economic Value. Telemed J E Health 2018;24:315-323.
14. Sanabria M, Buitrago G, Lindholm B, et al. Remote Patient Monitoring Program in Automated Peritoneal Dialysis: Impact on Hospitalizations. Perit Dial Int 2019;39:472-478.
15. Lew SQ, Sikka N, Thompson C, et al. Impact of remote biometric monitoring on cost and hospitalization outcomes in peritoneal dialysis. J Telemed Telecare 2018:1357633X18784417.
16. Weinhandl ED, Collins AJ. Relative risk of home hemodialysis attrition in patients using a telehealth platform. Hemodial Int 2018;22:318-327.
17. Perl J, Davies SJ, Lambie M, et al. The Peritoneal Dialysis Outcomes and Practice Patterns Study (PDOPPS): Unifying Efforts to Inform Practice and Improve Global Outcomes in Peritoneal Dialysis. Perit Dial Int 2016;36:297-307.
18. Nakamoto H, Kawaguchi Y, Suzuki H. Is technique survival on peritoneal dialysis better in Japan? Perit Dial Int 2006;26:136-143.
19. Chand DH, Bednarz D. Daily remote peritoneal dialysis monitoring: an adjunct to enhance patient care. Perit Dial Int 2008;28:533-537.
20. Schatell D, Wise M, Klicko K, et al. In-center hemodialysis patients' use of the internet in the United States: a national survey. Am J Kidney Dis 2006;48:285-291.
21. Bonner A, Gillespie K, Campbell KL, et al. Evaluating the prevalence and opportunity for technology use in chronic kidney disease patients: a cross-sectional study. BMC Nephrol 2018;19:28.
22. Lew SQ, Sikka N. Are patients prepared to use telemedicine in home peritoneal dialysis programs? Perit Dial Int 2013;33:714-715.

Chapter 6

Trajectories of Clinical and Laboratory Characteristics Associated with COVID- 19 in Hemodialysis Patients by Survival

Sheetal Chaudhuri, Rachel Lasky, Yue Jiao, John Larkin, Caitlin Monaghan, Anke Winter, Luca Neri, Peter Kotanko, Jeffrey Hymes, Sangho Lee, Yuedong Wang, Jeroen P Kooman, Franklin Maddux, Len Usvyat

Hemodialysis International. 2022; 26: 94–107

Abstract

Introduction

The clinical impact of COVID-19 has not been established in the dialysis population. We evaluated the trajectories of clinical and laboratory parameters in hemodialysis (HD) patients.

Methods

We used data from adult HD patients treated at an integrated kidney disease company who received a RT-PCR test to investigate suspicion of a SARS-CoV-2 infection between 01 May and 01 Sep 2020. Nonparametric smoothing splines were used to fit data for individual trajectories and estimate the mean change over time in patients testing positive or negative for SARS-CoV-2 and those who survived or died within 30 days of first suspicion or positive test date. For each clinical parameter of interest, the difference in average daily changes between COVID-19 positive versus negative group and COVID-19 survivor versus non-survivor group was estimated by fitting a linear mixed effects model based on measurements in the 14 days before (i.e., day -14 to day 0) day 0.

Results

There were 12,836 HD patients with a suspicion of COVID-19 who received RT-PCR testing (8,895 SARS- CoV-2 positive). We observed significantly different trends ($p<0.05$) in pre-HD systolic blood pressure (SBP), pre-HD pulse rate, body temperature, ferritin, neutrophils, lymphocytes, albumin, and interdialytic weight gain (IDWG) between COVID-19 positive and negative patient. For COVID-19 positive group, we observed significantly different clinical trends ($p<0.05$) in pre-HD pulse rate, lymphocytes, neutrophils and albumin between survivors and non-survivors. We also observed that, in the group of survivors, most clinical parameters returned to pre-COVID-19 levels within 60-90 days.

Conclusion

We observed unique temporal trends in various clinical and laboratory parameters among HD patients who tested positive versus negative for SARS-CoV-2 infection and those who survived the infection versus those who died. These trends can help to define the physiological disturbances that characterize the onset and course of COVID-19 in HD patients.

Introduction

The Coronavirus Disease (COVID-19) pandemic has greatly affected the dialysis community. Dialysis patients appear to be at increased risk for viral transmission with relatively high mortality rates ranging from 11% to 30%.[1-6] During the first half of 2020, there were over 11,200 COVID-19 hospitalizations among Medicare beneficiaries undergoing dialysis in the United States.[7] Various parameters such as pulse, body temperature, C-reactive protein (CRP) and lymphocyte counts at presentation were found to be associated to COVID-19 mortality in kidney failure.[4] However, the incubation time has not been clearly defined and patients may be infected with Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) and potentially infectious weeks before presentation with symptoms. Early detection of changes in physiological parameters have been suggested to aid the identification/prediction of patients at risk for COVID-19.[8] Apart from early detection, trends in clinical and laboratory parameters may also have prognostic significance. For example, distinct differences in trajectories of clinical and laboratory parameters before the start of kidney replacement therapy have been shown for patients who survived versus died during the first year on hemodialysis (HD).[9]

Albeit the clinical presentations in COVID-19 have been somewhat established[4], the changes in clinical parameters before presentation that characterize disease onset in humans are unknown secondary to a scarcity of longitudinal data available in the general population or collected in registries in the kidney failure population. HD patients have robust routine data collected in Electronic Health Records (EHRs) affording the opportunity to define the pathophysiological disturbances characterizing the onset and course of COVID-19 in kidney failure patients. The goal of this analysis was to compare trends in clinical and laboratory parameters between HD patients who tested positive or negative for SARS CoV-2. The second goal of this study was to compare clinical trends between survivors and non-survivors who were diagnosed with COVID-19.

Methods

General design

We used data from HD patients treated at a dialysis network in the United States of an integrated kidney disease company (Fresenius Medical Care, Waltham, MA, United States) between May and November 2020. HD patients who were suspected to have a SARS-CoV-2 infection at the outpatient dialysis clinics universally received reverse transcription polymerase chain reaction (RT-PCR) testing to diagnosis COVID-19. For

this analysis, we retrospectively evaluated the trends in clinical and laboratory parameters 90 days before the date of suspicion for SARS-CoV-2 infection among HD patients who were diagnosed RT-PCR COVID-19 positive versus those that were negative. Among HD patients with RT-PCR confirmed COVID-19, we also assessed at the trajectories for those who survived versus died within 30 days after suspicion of SARS-CoV-2 infection.

This analysis was performed under a protocol reviewed by New England Institutional Review Board (Needham Heights, MA, United States; Version 1.0 NEIRB# 17-1376378-1) who determined this analysis of existing patient data that was de-identified by the investigator was exempt and did not require informed consent. This analysis was conducted in adherence with the Declaration of Helsinki.

Patient population

We included data from adult (age ≥18 years) HD patients who received RT-PCR testing to investigate suspicion of a SARS-CoV-2 infection between 01 May and 01 September 2020. We required patients to have a minimum follow-up period of 90 days; follow-up data was captured through 30 November 2020 as applicable. Suspicion of SARS-CoV-2 infection was determined at presentation by active signs and symptoms of a flu-like illness. We excluded data from patients under investigation for SARS-CoV-2 that did not have a documented RT-PCR result, which included asymptomatic patients who were exposed to someone with known COVID-19 and were monitored for symptoms, as well as patients who were diagnosed with COVID-19 outside the outpatient clinic.

We used the first reported suspicion date to define the day 0 among COVID-19 positive patients. In limited cases without a documented suspicion date, we used the RT-PCR positive test date to define the day 0. For a control group, we identified patients who had one or more negative COVID-19 test result without any positive result during the analysis period. We used the first COVID-19 negative test date to define the day 0 since the date of suspicion for negative patients was not recorded in the provider's EHR. Patients with invalid/inconclusive test results were excluded from the analysis.

Statistical methods

We computed mean daily values for an array of clinical variables across the 90 days before day 0 for COVID-19 positive and COVID-19 negative groups. We reported findings from *a priori* selection of variables that appeared to have notable changes before COVID-19; these included pre-HD systolic blood pressure (SBP), pre-HD pulse, pre-HD body temperature, lymphocytes, neutrophils, ferritin, albumin, interdialytic weight gain (IDWG), and creatinine.

All data was collected during provision of standard medical care for HD patients. Data on SBP, pulse, body temperature, IDWG was collected on a per HD treatment basis. Laboratories were collected monthly with exception of ferritin that was collected on a quarterly basis.

Nonparametric smoothing splines[10] were constructed to fit data for individual trajectories and estimate the mean change over time since first suspicion/COVID-19 positive or negative date. Among COVID-19 positive patients, we stratified data for those who survived or died within 30 days of first day 0; trajectories were plotted 90 days after day 0 in survivors and up to 30 days after day 0 in COVID-19 in patients who died.

For each clinical parameter of interest, the difference in average daily changes between COVID-19 positive versus negative group and COVID-19 survivor versus non-survivor group was estimated by fitting a linear mixed effects model based on measurements in the 14 days before (i.e., day -14 to day 0) day 0. The analysis used all available data without any imputation.

Average value of clinical and laboratory parameters on day 0 were compared using unpaired t-test. Average value of clinical and laboratory parameters on day 0 between survivors and non-survivors was also compared using unpaired t-test.

Analyses were performed using SAS version 9.4 (SAS, Cary, NC, USA). Smoothing splines and visualizations were conducted using R version 3.5.2 (R Foundation, Vienna, Austria).

Results

Characteristics of HD patients who received COVID-19 testing

There were 12,836 HD patients with a suspicion of COVID-19 who received RT-PCR testing (8,895 COVID- 19 positive and 3,941 COVID-19 negative patients) between 01 May and 01 Sep 2020. The demographics and comorbidities for the two groups of patients are shown in **Table 6.1.** There was a slightly lower proportion of patients with a white race and higher proportion of patients with diabetes in the COVID- 19 positive group compared to the negative group. The mean number of days between the suspicion date and the positive test date in COVID-19 positive group was 5.6 days.

Table 6.1 Demographics and comorbidities of patients who tested COVID-19 positive and negative.

Parameter	COVID-19 positive	COVID-19 negative
Total number of patients	8,895	3,941
Male	54%	55%
White	37%	43%
Mean age as of first symptom date (std dev)	61.8 (14.2)	60.3 (14.7)
Mean vintage as of first symptom date (std dev)	4.0 (4.1)	3.8 (3.8)
Diabetes	69%	66%
Congestive heart failure	21%	22%
Ischemic heart disease	24%	24%

Trajectories of vital signs before COVID-19 testing

We observed the COVID-19 positive group had decreases in pre-HD SBP weeks before day 0; the SBP was around 5 mmHg lower at the suspicion date versus 14 days prior (**Figure 6.1**). Contrary to this, the COVID-19 negative group had trends for increases in pre-HD SBP in the weeks before testing. The linear mixed effects model estimated the daily change in pre-HD SBP during the 14 days before day 0 and identified the COVID-19 positive group had an average decrease of -0.3 mmHg/day; this was distinct compared to the average increase of 0.2 mmHg/day found in the COVID-19 negative group ($p<0.0001$) (**Table 6.2**).

Table 6.2 Average daily change in clinical and laboratory parameters 14 days prior to day 0.

Parameter	Mean daily change COVID-19 positive	Mean daily change COVID-19 negative	p-value
Pre-HD SBP (mmHg)	-0.2873	0.2293	<0.0001
Pre-HD pulse rate (BPM)	0.1728	0.0880	<0.0001
Pre-HD body temp (F)	0.0259	0.0140	<0.0001
Ferritin (ng/mL)	26.1162	3.6907	<0.0001
Lymphocytes (%)	-0.1772	-0.1208	0.0063
Neutrophils (%)	0.2445	0.1438	0.0002
Albumin (g/dL)	-0.0123	-0.0035	0.0004
IDWG (kg)	-0.0441	0.0068	<0.0001
Creatinine (mg/dL)	0.0127	0.0222	0.5561

Average daily changes are compared using linear mixed effects model in 14 days prior day 0. *P-value compares slopes in COVID-19 positive versus COVID-19 negative patients.

The COVID-19 positive group was also found to have subtle increases in the pre-HD pulse and body temperature in the week before day 0. These trends were consistent with the COVID-19 negative group, but less pronounced. The average daily change in pulse rate and body temperature was found to be significantly larger in the COVID-19 positive versus COVID-19 negative group ($p<0.0001$).

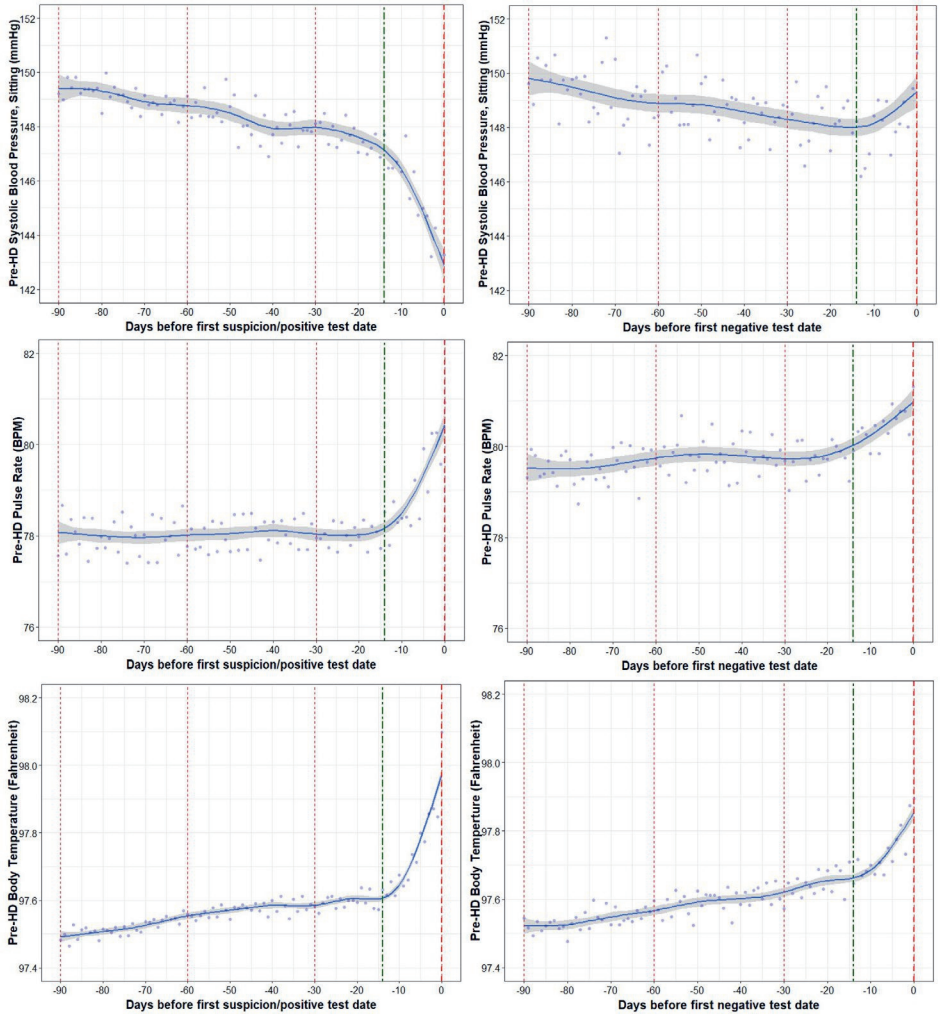


Figure 6.1 Trajectories of Pre-HD vital signs (SBP, pulse rate, body temperature) in COVID-19 positive and negative patients.

The average values for the pre-HD SBP, pulse, and body temperature on day 0 is shown in **Table 6.3**. There were significant differences between groups for pre-HD SBP and body temperature ($p<0.0001$).

Table 6.3 Average value of clinical and laboratory parameters on day 0.

Parameter	COVID-19 positive	COVID-19 negative	p value
Pre-HD SBP (mmHg)	143.28	150.74	<0.0001
Pre-HD pulse rate (BPM)	80.93	81.32	0.3266
Pre-HD body temp (F)	98.10	97.89	<0.0001
Ferritin (ng/mL)	1500.78	1038.42	0.0004
Lymphocytes (%)	17.42	18.01	0.1794
Neutrophils (%)	70.19	68.67	0.0064
Albumin (g/dL)	3.52	3.63	0.0074
IDWG (kg)	1.53	2.30	<0.0001
Creatinine (mg/dL)	9.04	8.76	0.3725

Trajectories of inflammatory markers before COVID-19 testing

Serum ferritin levels were found to have increased by around 400ng/ml in the COVID-19 positive group in the 14 days prior to day 0 (**Figure 6.2).** There was a daily change in ferritin of 26.12 ng/mL/day during the 14 days before day 0 in the COVID-19 positive group, which was distinct compared to the COVID-19 negative group that exhibited no remarkable changes ($p<0.0001$).

The percentage of lymphocytes were found to decrease below 20% in both COVID-19 positive and negative groups with a significant difference in trends of daily change ($p=0.0063$). Neutrophils showed an increasing trend in the COVID-19 positive and negative group 14 days prior to day 0 and the difference between the groups was significant ($p=0.0002$).

On day 0, ferritin and neutrophils were significantly higher in the COVID-19 positive group compared to the negative group ($p<0.05$).

Trajectories of nutritional markers before COVID-19 testing

Among both groups, there was a decline in serum albumin in the 14 days prior to day 0, yet the decline was more pronounced in the COVID-19 positive group compared to COVID-19 negative group. The decline in albumin was an average of 0.012g/dL per day in the COVID-19 positive group compared to the decline of an average of 0.004 g/dL in the COVID-19 negative group ($p=0.0004$) **(Figure 6.3)**. Notably, IDWG decreased by of 0.6kg per day in the COVID-19 positive group in the 14 days prior to day 0 compared to almost no change in COVID-19 negative group ($p<0.0001$). There were no differences in the trends of creatinine between the COVID-19 positive and negative group.

Albumin and IDWG were significantly lower in the COVID-19 positive group on day 0 compared to COVID-19 negative group on the negative test date.

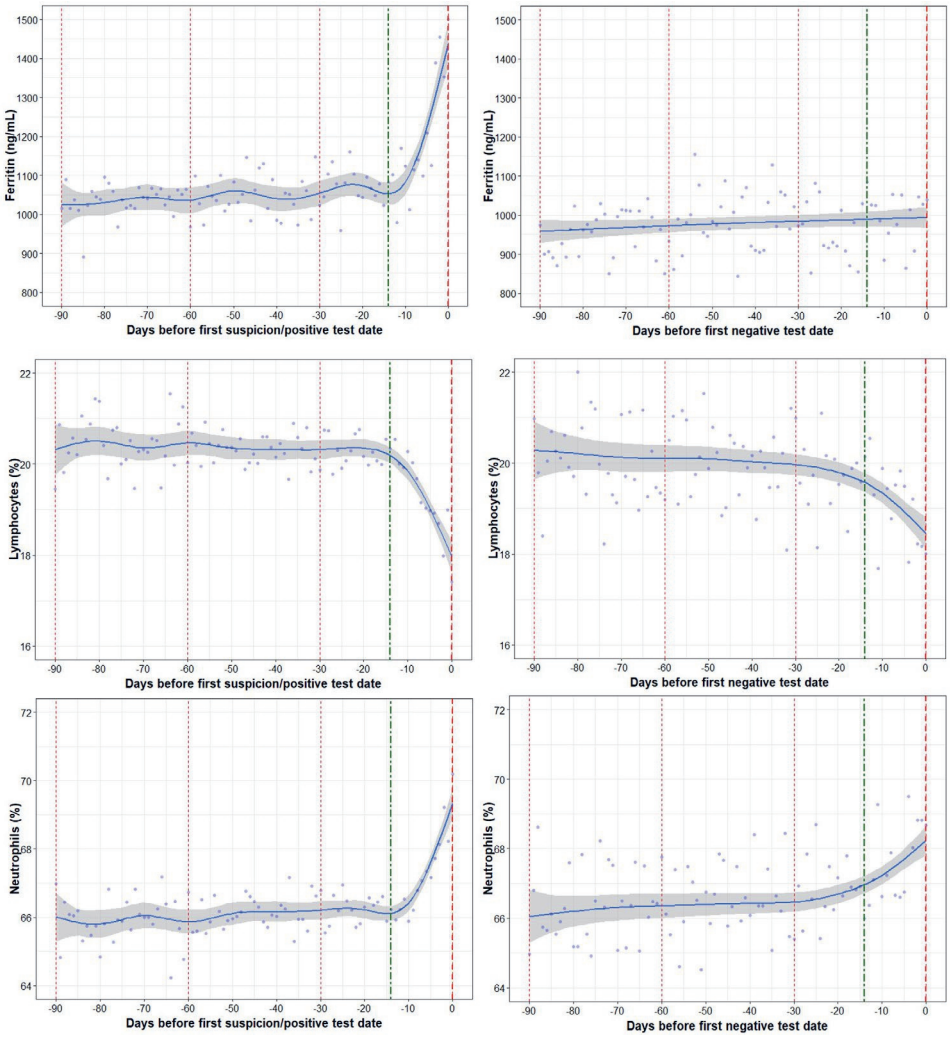


Figure 6.2 Trajectories in inflammatory markers (ferritin, lymphocytes, neutrophils) in COVID-19 positive and negative patients.

Characteristics of HD patients with COVID-19 by survival

Table 6.4 shows the demographics and comorbidities for patients diagnosed with COVID-19 who died within 30 days of COVID-19 positive date (non-survivors) versus those who survived (survivors). There were 7,897 survivors out of the 8,895 COVID-19 positive patients. The 998 non-survivors were more often older, male, white race, and had a higher comorbidity burden and longer dialysis vintage.

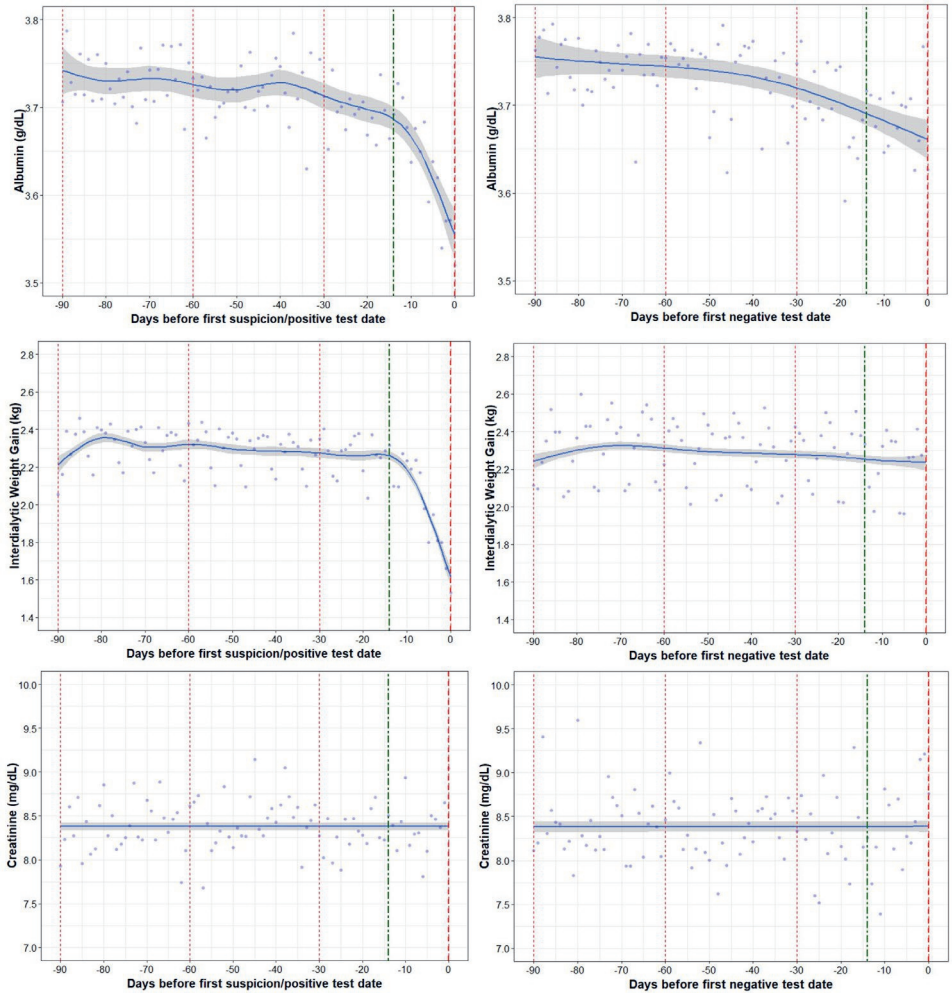


Figure 6.3 Trajectories in nutritional markers (albumin, IDWG, creatinine) in COVID-19 positive and negative patients.

Trajectories of vital signs before and after COVID-19 by survival

We observed a significant increase in the pre-HD pulse rate in the non-survivors versus the survivors with COVID-19. The linear mixed effects model showed an average increase of 0.29 BPM per day in the 14 days prior day 0 in the non-survivors compared to an average increase of 0.16BPM per day in survivors ($p<0.0001$) (**Table 6.5**). There was a difference in the pre-HD pulse rate between the survivors and non-survivors on day 0 ($p=0.0528$) (**Table 6.6**). Among survivors, it took 60 or more days after day 0 for pre-HD pulse to return to levels observed in the months before COVID-19.

Table 6.4 Demographics and comorbidities of COVID-19 survivors and non-survivors.

Parameter	Survivors	Non-survivors
Total number of patients	7,897	998
Male	53%	60%
White	37%	42%
Mean age as of first symptom date (std dev)	60.8 (14.1)	69.1 (12.5)
Mean vintage as of first symptom date (std dev)	3.9 (4.1)	4.7 (4.0)
Diabetes	68%	80%
Congestive heart failure	21%	28%
Ischemic heart disease	23%	29%

Table 6.5 Average daily changes in clinical and laboratory parameters 14 days prior to day 0.

Parameter	Survivors slope	Non-survivors slope	p-value
Pre-HD SBP (mmHg)	-0.2786	-0.3575	0.2222
Pre-HD pulse rate (BPM)	0.1582	0.2862	<0.0001
Pre-HD body temp (F)	0.0261	0.0245	0.5362
Ferritin (ng/mL)	24.4168	41.1784	0.1998
Lymphocytes (%)	-0.1569	-0.3420	<0.0001
Neutrophils (%)	0.2207	0.4345	<0.0001
Albumin (g/dL)	-0.0125	-0.0126	0.9854
IDWG (kg)	-0.0443	-0.0438	0.9365
Creatinine (mg/dL)	0.0129	0.0067	0.8301

Average daily changes are compared using linear mixed effects model in 14 days prior to day 0 in COVID-19 positive patients. *P-value compares slopes in survivors versus non-survivors.

Table 6.6 Average value of clinical and laboratory parameters on day 0.

Parameter	Survivors	Non-survivors	p value
Pre-HD SBP (mmHg)	143.98	136.5	0.0002
Pre-HD Pulse Rate (BPM)	80.73	82.77	0.0528
Pre-HD Body Temp (F)	98.1	98.07	0.7363
Ferritin (ng/mL)	1497.37	1567.57	0.7386
Lymphocytes (%)	17.75	14.45	0.0038
Neutrophils (%)	69.67	74.77	0.0003
Albumin (g/dL)	3.54	3.21	0.0204
IDWG (kg)	1.56	1.26	0.0432
Creatinine (mg/dL)	9.06	8.68	0.7392

There difference in the daily change of pre-HD DBP and body temperature in survivors versus non- survivors during the 14 days prior to day 0 was not significant. On day 0 there were distinctions in pre- HD SBP between the survivors and non-survivors (p=0.0002). Among survivors, pre-HD SBP began to increase back up again around 10 days following day 0 and it took 30 days or more to return to levels observed before the infection; trends in non-survivors showed further and more pronounced decreases in pre-HD SBP during the 30 days after day 0 (**Figure 6.4**). There were no differences in pre-HD body temperature between the survivors and non-survivors on day 0

(p=0.7363). Pre-HD body temperature returned to levels seen in the months before COVID-19 around 60 days after day 0.

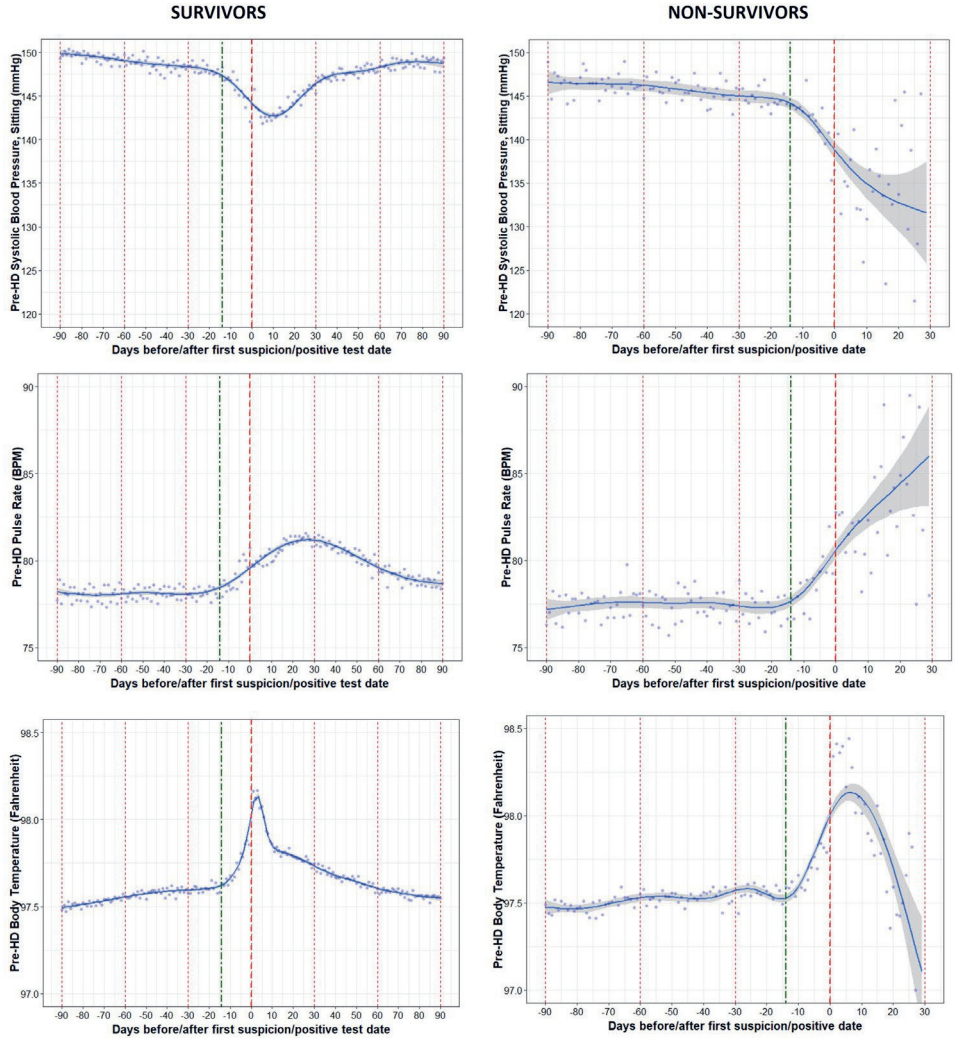


Figure 6.4 Trajectories in pre-HD vital signs (SBP, pulse rate, body temperature) in COVID-19 positive survivors and non-survivors

Trajectories of inflammatory markers before and after COVID-19 by survival

There were significant differences in inflammatory markers like lymphocytes and neutrophils in both COVID-19 positive survivors and non-survivors. The average decrease in the percentage of lymphocyte was about 0.34% in the 14 days prior to day 0 in non-survivors compared to 0.16% in survivors ($p<0.0001$). Similarly, there was a difference in the 14-day trend for neutrophils between the survivors and non-survivors ($p<0.0001$). There were also significant differences in the percentage of lymphocytes and neutrophils between the survivors and non-survivors on day 0.

As shown in **Figure 6.5**, the disturbances in inflammatory markers returned to the levels seen in the months before the infection within about 30 days after day 0 in patients who survived. Patients who died experienced more robust decreasing trends in lymphocyte levels as well as more robust increasing trends in neutrophil levels during the 30 days following day 0, as compared to those who survived.

Difference in Ferritin between survivors and non-survivors was not significant during 14 days prior to day 0 and on day 0.

Trajectories of nutritional markers before and after COVID-19 by survival

Albumin levels among patients who died decreased steadily during the 30 days following day 0, whereas albumin levels began to rise after approximately 15 days following day 0 for patients who survived.

Patients who died also experienced a larger decline in IDWG than the group who survived. Survivors experienced a decline in creatinine shortly after day 0, which remained lower than in the months before COVID-19 throughout the following 90 days. The average daily change was significantly different between groups for all nutritional markers (albumin and IDWG) in the 14 days prior to day 0 test date between the survivors and non-survivors ($p<0.05$). The difference in the slope of creatinine between the survivors and non-survivors was not significant($p=0.7392$).

There were also significant differences in albumin and IDWG between the survivors and non-survivors on day 0. Among survivors, the nutritional parameters took around 60-90 days to return to the same levels as in the months before the infection (**Figure 6.6**).

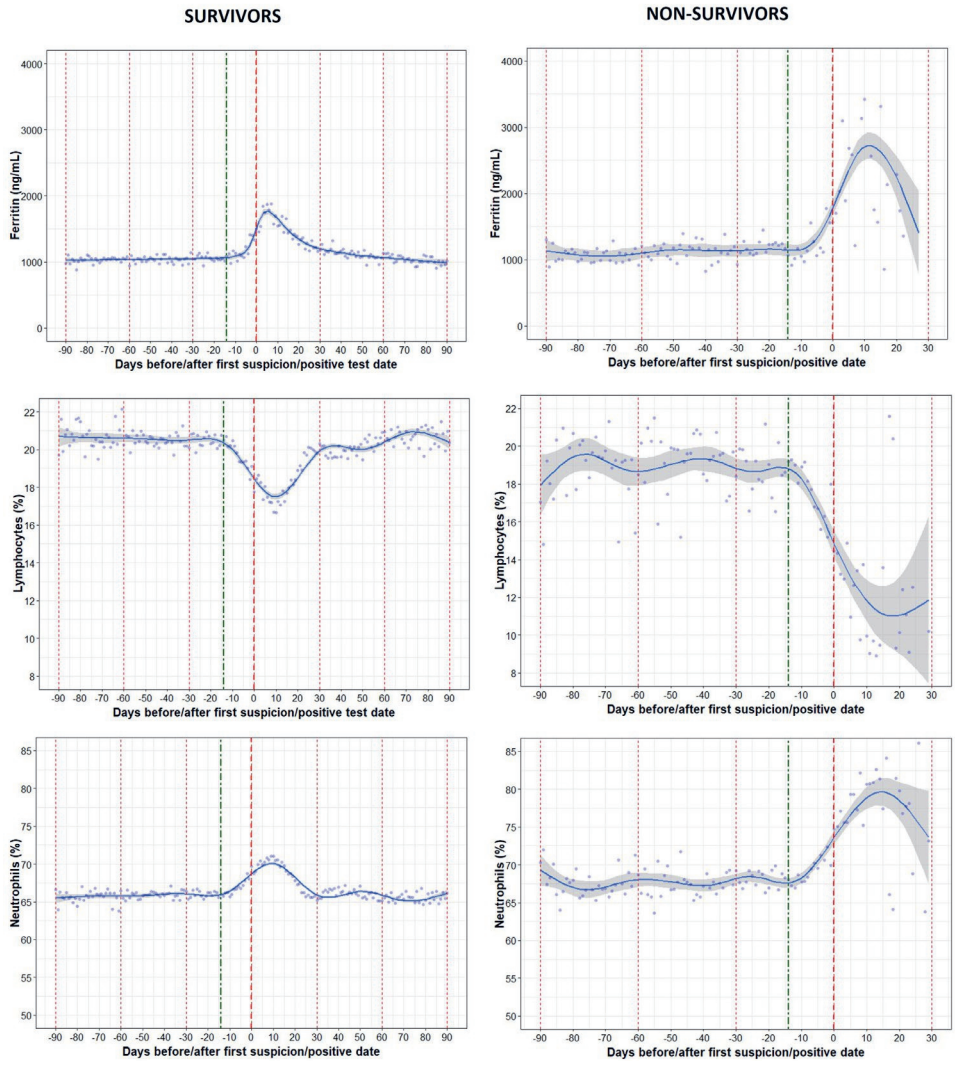


Figure 6.5 Trajectories in inflammatory markers (ferritin, lymphocytes, neutrophils) in COVID-19 positive survivors and non-survivors.

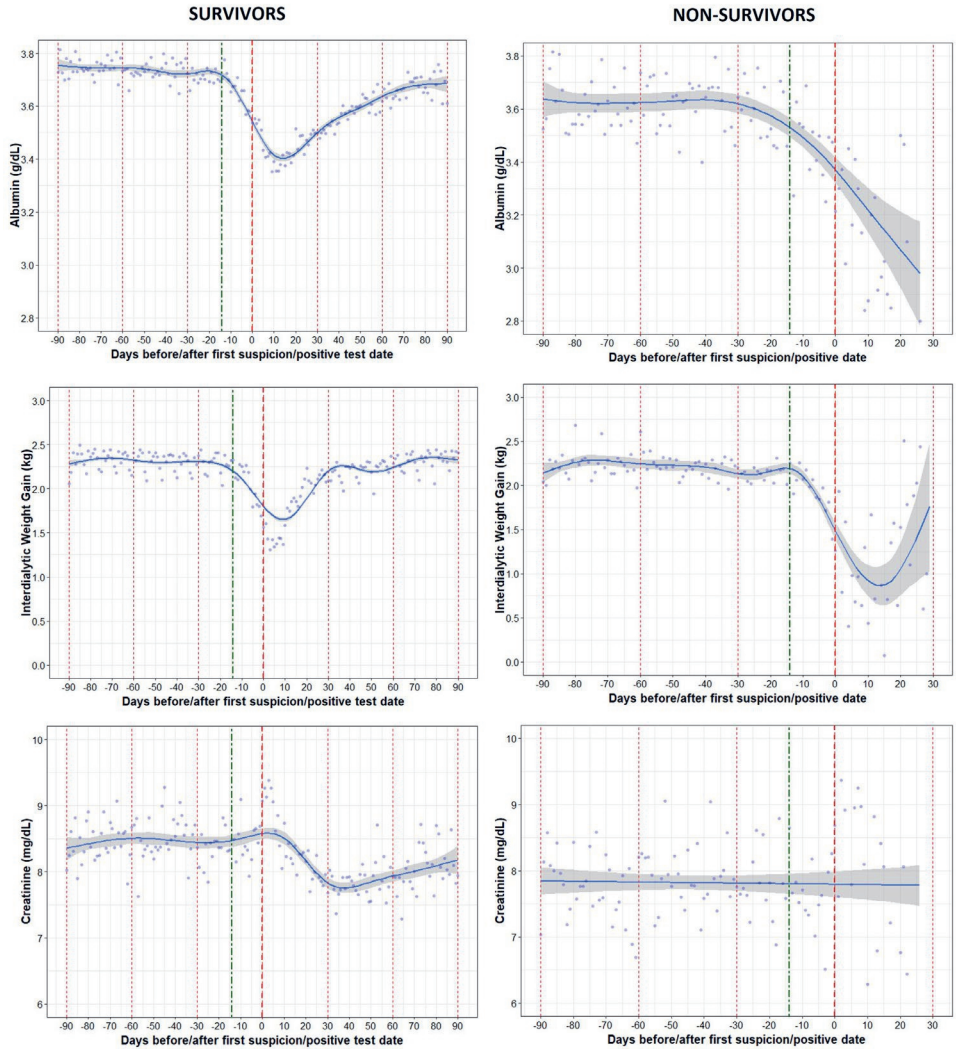


Figure 6.6 Trajectories in nutritional markers (albumin, IDWG, creatinine) in COVID-19 positive survivors and non-survivors.

Discussion

We observed unique temporal trends in various clinical and laboratory parameters among HD patients who tested positive versus negative for SARS-CoV-2 infection. Ultimately, these trends help to define the physiological disturbances that characterize the onset and course of COVID-19 in HD patients. The disturbances in various

parameters commonly started around 14 days before presentation and were more pronounced in patients who tested positive for COVID-19. We identified statistically distinct daily changes in vital signs, inflammatory, and nutritional markers across the weeks before presentation among patients who tested COVID-19 positive versus negative, as well as unique differences between groups on the date of suspicion/testing. Albeit statistically significant, the differences were often clinically subtle for many parameters. Nonetheless, these findings are anticipated to be of use in development of diagnostic support risk scores and prediction models. Among patients in the COVID-19 negative group, there were observed in various parameters which were anticipated since these patients likely underwent testing due to presentation with symptoms of a flu-like illness; however, we cannot exclude that some were tested secondary to exposure to someone with known COVID-19. Among the COVID-19 positive group, we identified clinically remarkable disturbances in vital signs, inflammatory, and nutritional markers after being diagnosed with COVID-19 that commonly took more than one month to return to normal among those who survived. The trajectories in parameters were statistically distinct between survivors and non-survivors brining insights that will be of importance to consider in the development of prognostic support risk scores and prediction models.

Published studies showing clinical trends before and at onset in COVID-19 HD patients are scarce. We recently showed how a machine learning predictive model that uses changes in clinical parameters before diagnosis had reasonable performance in classification of patients with a SARS-CoV-2 infection three days before symptoms onset (area under the curve is the testing dataset was 0.68). The model developed showed changes in IDWG had the highest variable feature importance (reflecting the impact of the value to the prediction).[11] Our present analysis adds to previous findings by showing quantitative trends in key parameters during the period before and after presentation with COVID-19. One of the novel findings we identified was the inverse trends in pre-HD SBP before presentation between COVID-19 positive versus negative patients. The decreases in SBP specifically associated with the onset of COVID-19 may be representative of some potential direct or indirect influences of the disease on the heart that caused a decompensation in cardiovascular system. The specificity of trends in physiological parameters for COVID-19 diagnosis should be assessed in future studies.

When comparing survivors and non-survivors, we found that mean age, dialysis vintage, and the proportion of males was higher in non-survivors, as was the presence of comorbidities such as diabetes and congestive heart failure and this consistent with previous findings.[3,4,12] Previous studies also have showed that higher body temperature, lower lymphocyte counts, higher levels of C-reactive protein and white blood cell counts were related to mortality in HD patients.[3,4,12-15] However, we observed that

average change in temperature 14 days prior to day 0 was not significantly different in survivors compared to non-survivors. The body temperature on day 0 for non-survivors was also slightly higher than survivors, although the difference was not significant.

In patients testing positive for COVID-19, the laboratory parameters anticipated to be directly related to the infection, such as neutrophils, lymphocytes, and ferritin, returned to their baseline values within 30 days after suspicion/testing. The relatively long time for infection related parameters to persist may be because SARS-CoV-2 infection may persist for longer period in HD patients. In a study of 19 HD patients with COVID-19 with repeat RT-PCR testing, SARS-CoV-2 tests remained positive in 68% patients after 20 days and in 32% after 40 days.[16]

We found a modest decline in IDWG in the weeks before presentation that was specific to the COVID-19 positive group. Among patients who contracted COVID-19, the decline in IDWG was more pronounced in non-survivors versus survivors, which may be a prognostic signal for malnutrition when accompanied by a deterioration in other nutritional parameters.[17] Parameters related to nutrition appear to take longer to return to the baseline in patients who survived suggesting the systemic effects of COVID-19 may be prolonged. Serum creatinine, which can also be considered a marker of lean tissue mass[18], even had not return to baseline 90 days after diagnosis. This suggests profound catabolic effects of COVID-19, which are due to multiple mechanisms such as anorexia, hypoxia, immobilization and increased levels of pro-inflammatory cytokines.[19] In our analysis, it also took 2-3 months for serum albumin to return to the baseline values seen in the months before infection.

This analysis assessed a large group of COVID-19 positive HD patients and provides novel information. However, it is important to note that patients could have been tested for presentation with symptoms or exposure to someone with known COVID-19, which cannot be fully deduced in cases where there is missing/unavailable data on the date of suspicion. Although most patients were likely tested for clinical reasons, which is known for most of the COVID-19 positive patients, this is a limitation of the analysis. In general, absolute differences in trends, although significant, were small and we therefore suggest using multiple markers in combination for risk prediction, as shown in our previous paper regarding a machine learning prediction model developed for early detection of patients with COVID-19.[11] Moreover, it is important to realize that the trends do not show the mean of individual trajectories, but an aggregate of the cohort groups. We cannot rule out that there might be some minimal temporal bias secondary to the definition of the reference date/day 0 for suspicion/testing that may have an impact on the trajectories. Also, there is a possibility that the incubation period of the virus was prolonged due to low immunity in dialysis patients. Nonetheless, the comparisons in trends in daily changes during the weeks prior suspicion/testing are

anticipated to have reasonably captured signals. Further investigations should consider inclusion of these insights to improve the precision of COVID-19 risk scores and prediction models being developed and used in care paradigms.

Conclusion

We found the trajectories of several clinical/laboratory parameters distinctly changed in patients during the weeks before presentation with COVID-19, as compared to patients who were tested for a SARS- CoV-2 infection and found negative. Many parameters changed in similar directions for both groups, yet significantly more for patients found to be COVID-19 positive versus negative. These included increasing pulse, body temperature, and neutrophils, along with decreasing lymphocytes and albumin. However, several factors uniquely changed in patients who were found to be COVID-19 positive, which included decreasing SBP and IDWG, as well as increasing ferritin. Among patients who were diagnosed with COVID-19, the survivors appeared to have statistically distinct trajectories showing less robust increases in pulse and neutrophils, and decreases in lymphocytes, compared to patients who died within 30 days after COVID-19 suspicion/testing. In patients with COVID-19 who survived, inflammatory and cardiovascular parameters returned to baseline levels in 1 month, however, the effect of COVID-19 on nutritional status appeared to be more prolonged and disturbances were seen up to 90 days. These findings appear to reveal some of the pathophysiologic trends defining the onset and course of the disease in the HD population, however, many changes were small. These insights are anticipated to be of high importance for development of prediction models for early identification and prognosis of COVID-19.

References

1. Corbett RW, Blakey S, Nitsch D, et al. Epidemiology of COVID-19 in an Urban Dialysis Center. J Am Soc Nephrol 2020;31:1815-1823.
2. Jager KJ, Kramer A, Chesnaye NC, et al. Results from the ERA-EDTA Registry indicate a high mortality due to COVID-19 in dialysis patients and kidney transplant recipients across Europe. Kidney Int 2020;98(6):1540–1548.
3. Ng JH, Hirsch JS, Wanchoo R, et al. Outcomes of patients with end-stage kidney disease hospitalized with COVID-19. Kidney Int 2020;98(6):1530–1539.
4. Hilbrands LB, Duivenvoorden R, Vart P, et al. COVID-19-related mortality in kidney transplant and dialysis patients: results of the ERACODA collaboration. Nephrol Dial Transplant 2020;35:1973-1983.
5. Neumann ME. Latest data show 305 dialysis patient deaths due to COVID-19 in the US. Nephrology News & Issues 2020; (Accessed 22 Apr 2020) https://www.healio.com/nephrology/infection-control/news/online/%7B3a263aa9-ad59- 4c3f-aab7-07b8395508e5%7D/latest-data-show-305-dialysis-patient-deaths-due-to-covid-19- in-the-us.
6. Medicine JHUSo. Mortality Analyses. Johns Hopkins Coronavirus Resource Center 2020; (Accessed December 2020) https://coronavirus.jhu.edu/data/mortality.
7. USRDS. COVID-19 Supplement. 2020; (Accessed November 2020) https://adr.usrds.org/2020/covid-19-supplement/1-covid-19-supplement.
8. Seshadri DR, Davies EV, Harlow ER, et al. Wearable Sensors for COVID-19: A Call to Action to Harness Our Digital Infrastructure for Remote Patient Monitoring and Virtual Assessments. Front Digit Health 2020;2:8.
9. Maddux DW, Usvyat LA, Ketchersid T, et al. Clinical parameters before and after the transition to dialysis. Hemodial Int 2018;22:235-244.
10. Wang Y. Smoothing Splines: Methods and Applications, Champman and Hall, New York. Chapman and Hall/CRC, 2011.
11. Monaghan CK, Larkin JW, Chaudhuri S, et al. Machine Learning for Prediction of Hemodialysis Patients with an Undetected SARS-CoV-2 Infection. Kidney360 2021:10.34067/KID.0003802020.
12. Alberici F, Delbarba E, Manenti C, et al. A report from the Brescia Renal COVID Task Force on the clinical characteristics and short-term outcome of hemodialysis patients with SARS-CoV-2 infection. Kidney Int 2020;98:20-26.
13. Goicoechea M, Sanchez Camara LA, Macias N, et al. COVID-19: clinical course and outcomes of 36 hemodialysis patients in Spain. Kidney Int 2020;98:27-34.
14. Keller N, Chantrel F, Krummel T, et al. Impact of first-wave COronaVIrus disease 2019 infection in patients on haemoDIALysis in Alsace: the observational COVIDIAL study. Nephrol Dial Transplant 2020; 35:1338-1411.
15. Valeri AM, Robbins-Juarez SY, Stevens JS, et al. Presentation and Outcomes of Patients with ESKD and COVID-19. J Am Soc Nephrol 2020;31:1409-1415.
16. Shaikh A, Zeldis E, Campbell KN, et al. Prolonged SARS-CoV-2 Viral RNA Shedding and IgG Antibody Response to SARS-CoV-2 in Patients on Hemodialysis. Clin J Am Soc Nephrol 2021;16:290–292.
17. Ye X, Dekker MJE, Maddux FW, et al. Dynamics of Nutritional Competence in the Last Year Before Death in a Large Cohort of US Hemodialysis Patients. J Ren Nutr 2017;27:412-420.
18. Canaud B, Ye X, Usvyat L, et al. Clinical and predictive value of simplified creatinine index used as muscle mass surrogate in end-stage kidney disease haemodialysis patients-results from the international MONitoring Dialysis Outcome initiative. Nephrol Dial Transplant 2020;35(12):2161–2171.
19. Morley JE, Kalantar-Zadeh K, Anker SD. COVID-19: a major cause of cachexia and sarcopenia? J Cachexia Sarcopenia Muscle 2020;11:863-865.

Chapter 7

Machine Learning for Prediction of Hemodialysis Patients with an Undetected SARS-CoV-2 Infection

Caitlin K Monaghan, John W Larkin, Sheetal Chaudhuri, Hao Han, Yue Jiao, Kristine M Bermudez, Eric D Weinhandl, Ines A Dahne-Steuber, Kathleen Belmonte, Luca Neri, Peter Kotanko, Jeroen P Kooman, Jeffrey L Hymes, Robert J Kossmann, Len A Usvyat , Franklin W Maddux

Kidney360 2021;2(3):456-468

Abstract

Background

We developed a machine learning (ML) model that predicts the risk of a hemodialysis (HD) patient having an undetected SARS-CoV-2 infection that is identified after the following 3 or more days.

Methods

As part of a healthcare operations effort, we used patient data from a national network of dialysis clinics (February-September 2020) to develop a ML model (XGBoost) that uses 81 variables to predict the likelihood of an adult HD patient having an undetected SARS-CoV-2 infection that is identified in the subsequent ≥3 days. We used a 60:20:20% randomized split of COVID-19 positive samples for the training, validation, and testing datasets.

Results

We used a select cohort of 40,490 HD patients to build the ML model (11,166 COVID-19 positive cases and 29,324 unaffected (control) patients). The prevalence of COVID-19 in the cohort (28% COVID-19 positive) was by design higher than the HD population. The prevalence of COVID-19 was set to 10% in the testing dataset to estimate the prevalence observed in the national HD population. The threshold for classifying observations as positive or negative was set at 0.80 to minimize false positives. Precision for the model was 0.52, the recall was 0.07, and the lift was 5.3 in the testing dataset. Area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC) for the model was 0.68 and 0.24 in the testing dataset, respectively. Top predictors of an HD patient having a SARS-CoV-2 infection were the change in interdialytic weight gain from the previous month, mean pre- HD body temperature in the prior week, and the change in post-HD heart rate from the previous month.

Conclusions

The developed ML model appears suitable for predicting HD patients at risk of having COVID-19 at least three days before there would be a clinical suspicion of the disease.

Introduction

The 2019 coronavirus disease (COVID-19) pandemic is challenging the world's healthcare systems, including bringing complexities to the maintenance of dialysis in people with end stage kidney disease (ESKD).[1-5] In the United States, most ESKD patients are treated by outpatient hemodialysis (HD) where social distancing can be difficult and heightened infection control measures are required (e.g. temperature screenings, universal masking, isolation treatments/shifts/clinics).[1-5] ESKD patients are typically older and have multiple comorbidities, placing the population at higher risk for requiring intensive care and dying if affected by COVID-19.[6-12]

Early reports from the United States show an 11% COVID-19 mortality in ESKD[13], which is higher than the 3% COVID-19 mortality shown in the national population.[14,15] This is not unexpected with reports from Asia and Europe suggesting a 16% to 23% COVID-19 mortality in ESKD.[16-19] Albeit the high mortality rate, an impaired immune response may render dialysis patients more frequently asymptomatic when infected by SARS-CoV-2.[16,17] In both the general and ESKD populations, the most prevalent symptoms of COVID-19 at presentation are fever (11%-66% in dialysis; 82% in general population) and cough (37%-57% in dialysis; 62% in general population).[16,20-22] The less frequent occurrence of signs and symptoms indicative of COVID-19 in dialysis patients could be making the outbreak even more challenging to manage.

Dialysis providers routinely capture patient/clinical data during care. The robust data collected during HD treatments (generally thrice weekly) provide unique opportunities to leverage artificial intelligence (AI) in predicting COVID-19 outcomes. AI modeling helped identify onset of the outbreak in China[23,24] and is currently being used to help with early detection of areas and individuals in the general population at risk for COVID-19.[25-27]

As part of an healthcare operations effort in response to the COVID-19 outbreak, an integrated kidney disease healthcare company aimed to develop a machine learning (ML) prediction model that identifies the risk of an HD patient having an undetected severe acute respiratory syndrome coronavirus-2 (SARS- CoV-2) infection. We analyzed the model performance to determine the possible utility for testing in the HD population.

Materials and methods

General

An integrated kidney disease healthcare company (Fresenius Medical Care, Waltham, MA, United States) used retrospective real-world data from its national network of dialysis clinics to develop a ML model that predicts the risk of an adult HD patient having an undetected SARS-CoV-2 infection that is identified after the following ≥3 days.

This analysis was performed in adherence with the Declaration of Helsinki under an initial and revised protocol reviewed by New England Independent Review Board (NEIRB). This retrospective analysis was determined to be exempt and did not require patient consent (Needham Heights, MA, United States; Protocol version 1.0 NEIRB#1-17-1302368-1; Protocol revision version 1.1 NEIRB#17-1348994-1).

COVID-19 mitigation and testing practices

The national network of dialysis clinics (Fresenius Kidney Care, Waltham, MA, United States) started implementing modified infection control measures in late Feb 2020 in response to the COVID-19 outbreak in the general population. Universal mitigation efforts at the provider included screening patients/staff before entry into the dialysis facility for high body temperature, signs or symptoms of flu- like illness, exposure to others with COVID-19, or a known infection diagnosed elsewhere.[28] Patients and staff were required to thoroughly wash their hands upon entering and leaving the facility. Patients were provided surgical masks and were required to wear them while in any area of the facility. Staff were required to wear enhanced personal protective equipment including use of masks, face shields, gowns, and gloves while being in the proximity of patients in any area. The first dialysis patients (n=2) at the provider were identified as COVID-19 positive on 03 Mar 2020.

All patients and staff with an elevated body temperature or symptoms of a flu-like illness were considered under investigation and had reverse transcription polymerase chain reaction (RT-PCR) laboratory testing for SARS-CoV-2 performed at a laboratory contracted by the dialysis provider. Patients under laboratory investigation for a SARS-CoV-2 infection were treated in dedicated isolation areas (rooms, shifts, or clinics) for suspected patients until confirmed negative by two RT-PCR tests that were more than 24 hours apart. Patients who had been exposed to others with COVID-19 were moved to unique isolation areas for exposed patients under investigation for 14 days and received RT-PCR testing if they presented with signs or symptoms of a flu-like illness. Patients with RT-PCR confirmed COVID-19 were treated in dedicated isolation areas for

infected patients until two negative RT-PCR tests more than 24 hours apart were documented.

Population and outcome

We considered data from adult (age ≥18 years) HD patients treated throughout the national network for development of a model to predict individuals with an undetected SARS-CoV-2 infection. The observation period started on 27 Feb 2020. Positive arm included data from patients who had ≥1 confirmed positive RT-PCR COVID-19 test as of the end of the observation period (08 Sep 2020, n=11,166). Negative arm included data from patients who: 1) were found COVID-19 negative (n=7,959), or 2) were randomly sampled from all active patients at the dialysis provider without a reported suspicion of COVID-19 as of the end of the observation period (n=21,365). The random sampling was performed using the 'sample' function from the 'pandas' Python package.

We defined the index date of a HD patient having a SARS-CoV-2 infection as the date of the COVID-19+ test. In control patients with a negative COVID-19 test result, the test date was used as the index date. In controls without a test, the index date was randomly sampled from the positive cases' index dates occurring before 25 Aug 2020, two weeks before the end of the observation period. This cutoff was chosen to minimize the possibility that control patients were infected but had not displayed signs or symptoms leading to testing before the end of the observation period. We included data from patients with 1) ≥1 hemoglobin sample collected both 1-14 days and 31-60 days before the individual's prediction date (3 days prior to index date, further defined below), and 2) ≥1 HD treatment both 1-7 days and 31-60 days preceding the prediction date. This was done to ensure we included only active patients as hemoglobin draws are conducted weekly for in-center HD (typically thrice weekly treatments). We excluded data from patients suspected to have COVID-19 who were pending laboratory testing or were classified as person under investigation (PUI) where no laboratory testing was performed or documented.

AI model development

Software and ML model logic

We used Python version 3.7.7 (Python Software Foundation, Delaware, United States) to build the ML model utilizing the XGBoost package (29). The XGBoost Python package used input variables from the training dataset to construct multiple decision trees, giving each a random sample, and established a series of thresholds that split variables to maximize the information gain. Decision trees were constructed iteratively, and new decision trees were added to predict prior errors. The decision trees made by the XGBoost ML model are inherently able to handle missing values without imputation by

including their presence when determining the splits (e.g. splitting observations with temperatures ≥98.0°F (≥36.7°C) from temperatures <98.0°F (<36.7°C) or missing temperatures). After no further improvements in performance were achieved using the validation dataset (also used for hyperparameter tuning), the ensemble of decision trees produced the final ML model that was assessed with the testing dataset.

Undetected SARS-CoV-2 prediction model

We used 81 a priori selected treatment/laboratory variables up to the individually defined prediction date (3 days prior to the index date defined above) to predict the risk of a SARS-CoV-2 infection being identified in the following ≥3 days (**Figure 7.1**). This is intended to yield individual predictions at least 3 days in advance of symptoms that warranted testing. We used a 60:20:20% randomized split of COVID-19+ samples for the training, validation, and testing datasets, and added the same number of COVID-19 negative patients to only the training and validation datasets. The testing dataset used to evaluate final model performance had a higher number of COVID-19 negative samples added to more closely match the prevalence observed in the overall national HD population.[30,31]

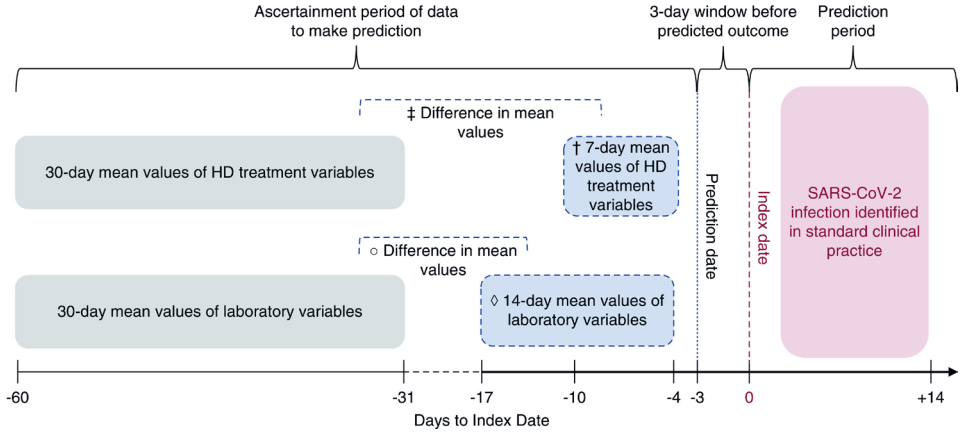


Figure 7.1 Prediction timeline for data ascertainment and prediction of HD patients with and without SARS-CoV-2 infection identified in the subsequent ≥3 days. ML model used HD treatment variables († mean values 1-7 days before the prediction date; ‡ difference in mean values 31-60 days to 1-7 days before the prediction date) and laboratory variables (◊ mean values 1-14 days before the prediction date; ○ difference in mean values 31-60 days to 1-14 days before the prediction date) for prediction of SARS-CoV-2 infection.

Statistical methods

Descriptive statistics

Descriptive statistics for HD patients were tabulated for demographics and variables at the time of the prediction for an undetected SARS-CoV-2 infection. Data are stratified by HD patients who did, or did not, have laboratory confirmation of COVID-19 after the date of prediction.

Analysis of ML model feature importance

Shapley values[32,33] were calculated using the SHAP python package to determine the influence of each variable on the predictions.[34,35] SHAP values are calculated for each variable and each observation, representing a measure of impact (positive or negative value) of the observed value on each individual prediction. SHAP methods withhold and include individual inputs in all possible combinations, and compare differences between withheld and included data, to compute the mean value of all possible differences for attributing the feature importance. SHAP values are output as log odds (i.e. the logarithm of the odds ratio), meaning they are additive explanations of feature importance. SHAP values for each variable are summed for each set of observations (in this case, for each patient) and converted from log odds to probability, which is then output by the model as the prediction. Thus, the more positive SHAP values increase the predicted probability, while more negative SHAP values decrease it. Overall feature importance for individual variables in the model were calculated from the SHAP values using the mean absolute values for each variable across all observations.

Analysis of ML model performance

Performance of ML model was measured by the area under the receiver operating characteristic curve (AUROC) in the training, validation, and testing datasets, as well as the recall, precision, and lift in the testing datasets. Additionally, we evaluated the area under the precision-recall curve (AUPRC) in the testing dataset.

AUROC measures the rate of true and false positives classified by the prediction model across probability thresholds. The definition of true/false positives and negatives is shown in **Table 7.1**.

Table 7.1 Definition of true/false positive and negative predictions classified by the model in the assessment of performance in the testing dataset.

True positives	Patients classified as COVID-19 positive by the model who were in the COVID-19 positive group
False positives	Patients classified as COVID-19 positive by the model who were in the COVID-19 negative group
True negatives	Patients classified as COVID-19 negative by the model who were in the COVID-19 negative group
False negatives	Patients classified as COVID-19 negative by the model who were in the COVID-19 positive group

Recall (sensitivity) measures the rate of true positives classified by the model at a specified threshold and is calculated as follows:

Recall = number of true positives classified by model / (number of true positives classified by model + number of false negatives classified by model)

Precision measures the positive predictive value for the model at a specified threshold and is calculated as follows:

Precision = number of true positives classified by model / (number of true positives classified by model + number of false positives classified by model

Lift measures the effectiveness of the model compared to random sampling and is calculated as follows:

Lift = model precision / proportion of positives in dataset

AUPRC measures the ratio of precision for corresponding recall values across probability thresholds.[36]

AUROC, AUPRC, recall, and precision metrics yield scores on a scale of 0 (lowest) to 1 (highest). A model performing at chance would yield an AUROC of 0.5, an AUPRC equal to the proportion of positives in the dataset, and a lift value of 1. The cutoff threshold for classifying predictions were selected to optimize recall, precision, and lift according to the use case.

Results

Patient characteristics

We identified data from a select cohort of 40,490 HD patients meeting eligibility criteria (11,166 COVID- 19+ cases and 29,324 unaffected (control) patients). The prevalence of COVID-19 in the cohort (28% COVID-19 positive) was by design higher than the HD population. The prevalence of COVID-19+ cases (about 28% COVID-19+) in the training and validation datasets was consistent within the cohort. For the testing dataset used to evaluate final model performance, there was a 10% prevalence of COVID-19+ cases based on the designed data split that was made to estimate the prevalence observed in the national HD population.[30,31]

In the cohort, there was a higher proportion of HD patients with a SARS-CoV-2 infection of black race, Hispanic ethnicity, and with diabetes (**Table 7.2**). Mean values for the 81 treatment and laboratory variables before a SARS-CoV-2 infection being identified in the subsequent ≥3 days (or concurrent index date in controls) are shown in **Tables 7.3 & 7.4**.

Table 7.2 Demographics and comorbidities of HD patients with and without an undetected SARS- CoV-infection dentified in the ubsequent ≥3 Days

Variable	Unaffected patients N (%) or Mean±SD	COVID-19+ patients N (%) or Mean±SD
Number of HD patients	29,324	11,166
Age (yrs)	62.66±14.25	62.62±13.92
Male	16,614 (56.66%)	6,149 (55.07%)
White race	12,021 (40.99%)	4,338 (38.85%)
Black race	7,838 (26.73%)	3,354 (30.04%)
Other race	1,223 (4.17%)	372 (3.33%)
Unknown race	8,242 (28.11%)	3,102 (27.78%)
Hispanic ethnicity	2,849 (14.03%)	1,831 (23.34%)
BMI (kg/m^2)	29.26±7.71	29.45±7.83
Dialysis vintage (yrs)	3.75±4.11	3.96±4.09
Diabetes	19,186 (65.58%)	8,085 (73.11%)
CHF	6,710 (22.93%)	2,595 (23.47%)
Ischemic heart disease	7,647 (26.14%)	2,830 (25.59%)
Central venous catheter access	6,799 (23.19%)	2,738 (24.52%)

Age, gender, and catheter access variables were included in the ML prediction model to classify the risk of an individual HD patient having a SARS-CoV-2 infection being identified in the following ≥3 days. HD: hemodialysis; CHF: congestive heart failure; BMI: body mass index; N: patient count; SD: standard deviation.

HD patients who contracted COVID-19 had only subtle, clinically unremarkable distinctions in treatment and laboratory characteristics before being suspected to have a SARS-CoV-2 infection compared to unaffected patients. Mean pre-/post-HD body temperatures (**Table 7.3**) and inflammatory markers (white blood cell (WBC) count and differential) (**Table 7.4**) before a SARS-CoV-2 infection being identified did not did not

show a clinically relevant difference differ between groups. HD patients who had a SARS-CoV-2 infection identified in the following 3 days did appear to have somewhat higher ferritin levels compared to unaffected patients.

Table 7.3 Clinical and treatment characteristics of HD patients with and without an undetected SARS-CoV-2 infection dentified in the ubsequent ≥3 Days.

Variable	Unaffected patients Mean±SD; N	COVID-19+ patients Mean±SD; N
Number of HD patients	29,324	11,166
Pre-HD sitting SBP (mmHg)†	148.31±22.83; 29,324	146.03±23.03; 11,166
Change in pre-HD sitting SBP (mmHg)‡	-0.40±15.78; 29,324	-1.95±16.72; 11,166
Pre-HD sitting DBP (mmHg)†	76.87±13.86; 29,322	75.44±13.58; 11,166
Change in pre-HD sitting DBP (mmHg)‡	-0.32±9.03; 29,322	-0.88±9.48; 11,166
Pre-HD weight (kg)†	85.71±24.51; 29,323	85.09±24.51; 11,165
Change in pre-HD weight (kg)‡	-0.17±2.24; 29,323	-0.66±2.73; 11,165
Pre-HD body temperature (°F)†	97.56±0.61; 29,324	97.76±0.66; 11,166
Change in pre-HD body temperature (°F)‡	0.07±0.56; 29,324	0.22±0.65; 11,166
Post-HD sitting SBP (mmHg)†	140.40±21.60; 29,321	144.44±21.62; 11,166
Change in post-HD sitting SBP (mmHg)‡	0.43±14.98; 29,320	1.55±15.74; 11,166
Post-HD sitting DBP (mmHg)†	73.91±12.58; 29,319	73.56±12.33; 11,166
Change in post-HD sitting DBP (mmHg)‡	0.15±8.49; 29,318	0.41±8.79; 11,166
Post-HD body temperature (°F)†	97.58±0.56; 29,318	97.70±0.62; 11,166
Change in post-HD body temperature (°F)‡	0.03±0.50; 29,317	0.14±0.57; 11,165
Pre-HD respirations per minute†	17.64±1.16; 29,324	17.72±1.15; 11,166
Change in pre-HD respirations per minute‡	-0.001±0.97; 29,324	0.01±1.02; 11,166
Pre-HD pulse (BPM)†	79.00±12.11; 29,324	79.02±11.90; 11,166
Change in pre-HD pulse (BPM)‡	0.11±7.26; 29,324	1.06±7.56, 11,166
Post-HD respirations per minute†	17.56±1.15; 29,320	17.65±1.13; 11,165
Change in post-HD respirations per minute‡	-0.007±0.95; 29,319	0.0004±0.99; 11,165
Post-HD pulse (BPM)†	75.80±11.23; 29,321	77.23±11.16; 11,166
Change in post-HD pulse (BPM)‡	-0.32±7.16; 29,320	1.30±7.87; 11,166
IDWG (kg)†	2.24±1.21; 29,083	1.95±1.29; 11,039
Change in IDWG (kg) ‡	0.01±0.90; 29,004	-0.26±1.09; 10,991
Post-HD weight loss (kg)†	-2.26±1.07; 29,317	-2.06±1.07; 11,160
Change in post-HD weight loss (kg)‡	-0.01±0.68; 29,316	0.18±0.77; 11,159
Post-HD body temperature change†	0.01±0.66; 29,318	-0.06±0.70; 11,165
Change in post-HD body temperature change‡	-0.04±0.66; 29,317	-0.07±0.71; 11,165
Post-HD respirations per minute change†	-0.08±0.97; 29,320	-0.07±0.97; 11,165
Change in post-HD respirations per minute change‡	-0.01±1.04; 29,319	-0.01±1.07; 11,165
Post-HD pulse change (BPM)†	-3.20±8.86; 29,321	-1.79±8.77; 11,166
Change in post-HD pulse change (BPM) ‡	-0.43±7.75; 29,320	0.24±8.06; 11,166
% HD treatments with nasal oxygen administered†	5.23±18.52; 29,324	5.67±19.16; 11,166
Change in % HD treatments with nasal oxygen administered‡	0.37±13.40; 29,324	0.72±14.12; 11,166

All variables were included in the ML prediction model to classify the risk of an individual HD patient having a SARS-CoV-2 infection being identified in the following ≥3 days. † Mean values of HD treatment variables 1-7 days before the prediction date (i.e. 3 days before suspicion of SARS-CoV-2 infection in standard clinical practice). ‡ Mean values of the difference in HD treatment variables 31-60 days to 1-7 days before the prediction date. HD: hemodialysis; SBP: systolic blood pressure; DBP: diastolic blood pressure; IDWG: interdialytic weight gain; Post-HD Weight Loss: post-HD minus pre-HD weight (kg); N: patient count; SD: standard deviation. (100°F − 32) × 5/9 = 37.8°C

Table 7.4 Laboratory characteristics of HD patients with and without an undetected SARS-CoV-2 infection identified in the subsequent ≥3 days.

Variable	Unaffected patients Mean±SD, N	COVID-19+ patients Mean±SD, N
Number of HD patients	29,324	11,166
Albumin (g/dL)†	3.79±0.40; 13,723	3.69±0.46; 5,252
Change in albumin (g/dL)‡	-0.002±0.25; 13,139	-0.03±0.27; 5,012
Creatinine (mg/dL)†	8.42±3.06; 13,323	8.41±3.14; 5,113
Change in creatinine (mg/dL)‡	0.08±1.40; 12,711	0.16±1.52; 4,860
Bicarbonate (mmol/L)†	24.24±3.05; 13,395	24.22±3.22; 5,137
Change in bicarbonate (mmol/L)‡	0.02±2.97; 12,772	-0.16±3.12; 4,864
BUN (mg/dL)†	56.21±18.53; 14,941	56.17±19.27; 5,631
Change in BUN (mg/dL)‡	-0.21±15.50; 14,400	-0.13±16.55; 5,416
URR†	74.92±6.52; 14,273	75.05±6.61; 5,348
Change in URR‡	0.09±5.89; 13,548	0.07±6.08; 5,054
Sodium (mmol/L)†	137.50±3.37; 13,139	137.08±3.52; 5,046
Change in sodium (mmol/L)‡	-0.10±2.83; 29,324	-0.25±3.10; 4,772
Potassium (mmol/L)†	4.80±0.68; 16,051	4.78±0.70; 6,217
Change in potassium (mmol/L)‡	0.01±0.60; 15,499	-0.01±0.63; 6,003
Phosphate (mg/dL)†	5.55±1.74; 15,489	5.37±1.71; 5,913
Change in phosphate (mg/dL)‡	0.01±1.48; 14,918	-0.03±1.46; 5,692
Chloride (meq/L) †	98.66±4.14; 12,602	98.33±4.13, 4,702
Change in chloride (meq/L) ‡	-0.19±3.35; 11,708	-0.24±3.50; 4,450
Calcium (mg/dL)†	8.89±0.69; 15,420	8.78±0.73; 5,878
Change in calcium (mg/dL)‡	0.02±0.58; 14,882	-0.07±0.60; 5,659
Corrected calcium (mg/dL)†	9.06±0.66; 12,865	9.04±0.71; 4,903
Change in corrected calcium (mg/dL)‡	0.01±0.54; 12,148	-0.03±0.59; 4,608
iPTH (pg/mL)†	489.46±454.13; 10,090	497.22±490.12; 3,801
Change in iPTH (pg/mL)‡	-21.39±280.41; 7,245	-21.84±296.17; 2,734
Ferritin (ng/mL)†	1029.94±576.07; 8,229	1197.32±900.22; 3,138
Change in cerritin (ng/mL)‡	52.90±505.99; 4,400	142.00±739.89; 1,589
TSAT (%)†	33.07±14.10; 13,051	31.29±14.42; 5,008
Change in TSAT (%)‡	0.17±15.33; 12,310	-1.59±16.54; 4,689
Hgb (g/dL)†	10.76±1.24; 29,324	10.61±1.26; 11,166
Change in Hgb (g/dL)‡	0.05±1.07; 29,324	0.01±1.13; 11,166
Platelet count (x 10^9/L)†	195.49±72.47; 11,378	192.35±77.10; 4,293
Change in platelet count (x 10^9/L)‡	-1.93±49.23; 10,595	-7.82±55.06; 3,963
WBC count (x 10^9/L)†	6.93±2.36; 13,043	6.55±2.39; 5,027
Change in WBC count (x 10^9/L)‡	0.03±1.76; 12,344	-0.36±1.93; 4,733
% of neutrophils†	66.11±9.50; 17,215	66.59±9.53; 6,941
Change in % of neutrophils‡	0.06±6.72; 14,931	0.47±7.37; 5,997
% of lymphocytes†	20.22±7.98; 17,215	19.76±7.96; 6,941
Change in % of lymphocytes‡	-0.04±4.99; 14,931	-0.53±5.57; 5997
% of monocytes†	6.38±1.90; 17,215	6.69±2.14; 6,941
Change in % of monocytes‡	0.02±1.48; 14,931	0.37±1.82; 5,997
% of eosinophils†	4.29±2.88; 17,212	3.95±2.84; 6,939
Change in % of eosinophils‡	-0.10±2.03; 14,927	-0.40±2.28; 5,995
% of basophils†	0.75±0.47; 17,206	0.73±0.45; 6,934
Change in % of basophils‡	0.05±0.54; 14,917	0.03±0.52; 5,988

All variables were included in the ML prediction model to classify the risk of an individual HD patient having a SARS-CoV-2 infection being identified in the following ≥3 days. † Mean values of laboratory variables 1-14 days before the prediction date (i.e. 3 days before suspicion of SARS- CoV-2 infection in standard clinical practice). ‡ Mean values of the difference in laboratory variables 31-60 days to 1-14 days before the prediction date. HD: hemodialysis; Hgb: hemoglobin; WBC: white blood cell; TSAT: transferrin saturation; URR: urea reduction ratio; iPTH: intact parathyroid hormone; BUN: blood urea nitrogen; N: patient count; SD: standard deviation.

Prediction model feature importance

Calculation of variable feature importance with SHAP values found the top three predictors of HD patients having a SARS-CoV-2 infection were the change in interdialytic weight gain (IDWG) from the previous month, mean pre-HD body temperature in the prior week, and the change in post-HD pulse from the previous month (**Figure 7.2A**).

The SHAP value plot in **Figure 7.2B** further shows the degree of positive or negative impact of each individual measurement for each individual prediction. Each dot corresponds to an individual patient, where the dot's position on the x-axis represents that feature's impact on the model prediction while the color indicates how high or low that feature's value was. Features with missing values are indicated in gray.

For the top predictor of the change in interdialytic weight gain in the week before compared to the month before a SARS-CoV-2 infection, smaller (negative) values (cooler colors) were associated with a positive SHAP value, while larger values (warmer colors) were associated with a negative SHAP value. These results showed for each individual prediction, the model generally considered decreases in interdialytic weight gain from the previous month to be associated with a greater probability of an undetected SARS-CoV-2 infection and an increase in interdialytic weight gain to be associated with a lower likelihood of an undetected SARS-CoV-2 infection. In other words, patients who do not gain as much weight as usual in between dialysis treatments are deemed more likely to have an undetected SARS-CoV-2 infection by the model.

Along with highlighting directional effects as previously stated, **Figure 7.2B** also highlights different distributions of effects that might not be apparent when viewing the mean absolute values as in **Figure 7.2A**. For example, the eighth most important variable, change in monocytes from the previous month, produces the largest (most positive) SHAP values out of all the variables shown. This long, rightward tail along the x-axis indicates that despite having a lower mean absolute value in comparison to other variables, for some individuals this is very important. Specifically, the model assessed that patients with increased monocyte levels from the previous month are deemed more likely to have a SARS-CoV-2 infection, whereas the SHAP values for those with similar or lower levels of monocytes do not significantly decrease the prediction.

Prediction model performance

The ML model had adequate performance in prediction of the 3-day risk for having an undetected SARS- CoV-2 infection. The ML model had an AUROC of 0.77, 0.67, and 0.68 in the training, validation, and testing datasets respectively (**Figure 7.3**). The ML model had an AUPRC of 0.24 in the testing dataset (**Figure 7.4**).

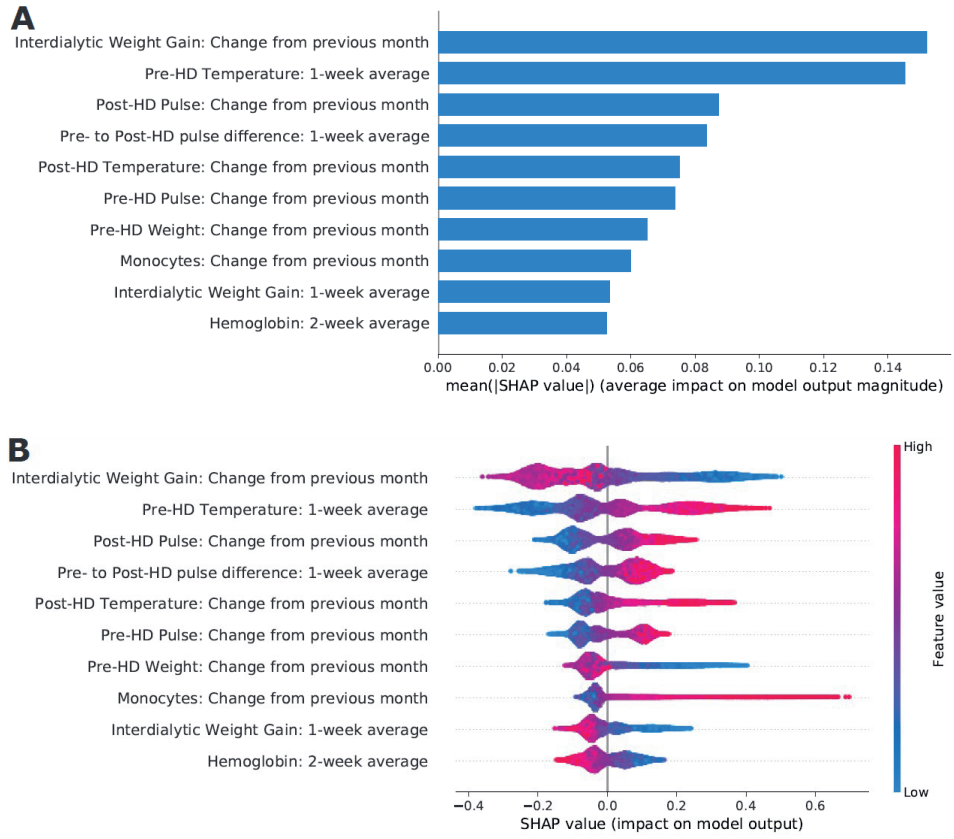


Figure 7.2 SHAP value plots for the ML model showing the extent each predictor contributes (positively or negatively) to each individual prediction. **A**) Bar plot of the mean absolute SHAP values for the top 10 predictors in descending order. **B**) SHAP value plot for the degree of the positive or negative impact of each individual measurement on the prediction (x-axis), with warmer colors representing higher observed values for that measurement, cooler colors indicating lower values for that measurement, and gray color representing a missing value for that measurement.

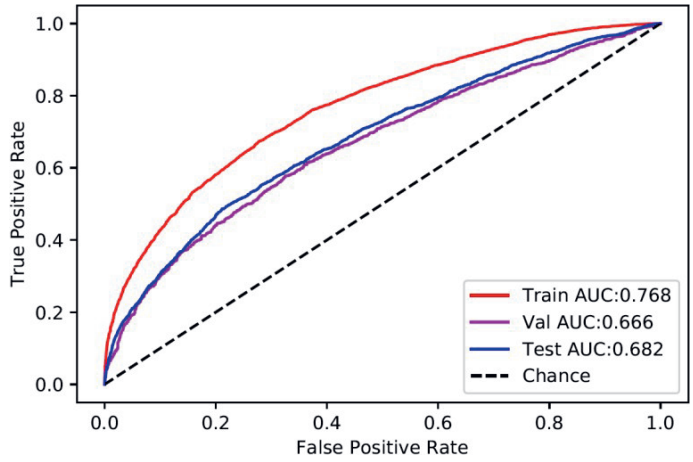


Figure 7.3 Area under the receiver operating characteristic curve (AUROC) plot for the ML model showing the rate of true and false positives classified by the prediction model across probability thresholds.

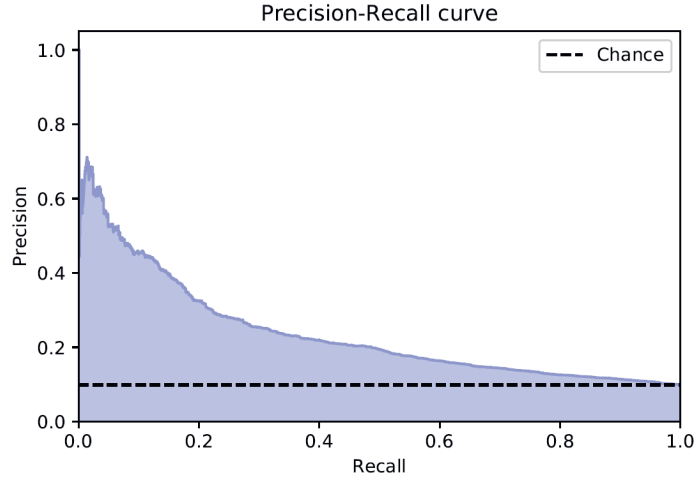


Figure 7.4 Area under the precision-recall curve (AUPRC) plot for the ML model showing the ratio of precision for corresponding recall values across probability thresholds.

Setting the threshold for classifying observations as positive or negative at 0.80 to minimize false positives, the precision for the ML model in the testing dataset was 0.52 showing 52% of patients predicted to have a SARS-CoV-2 infection actually had symptoms in the subsequent ≥3 days and were confirmed to have COVID-19. Given the high threshold, recall was 0.07 showing the model correctly predicted true positives for a SARS-CoV-2 infection in 7% of positive HD patients. The lift was 5.3, suggesting model

use is 5.3 times more effective in predicting a HD patient who contracts COVID-19, as compared to not having a model (**Figure 7.5**).

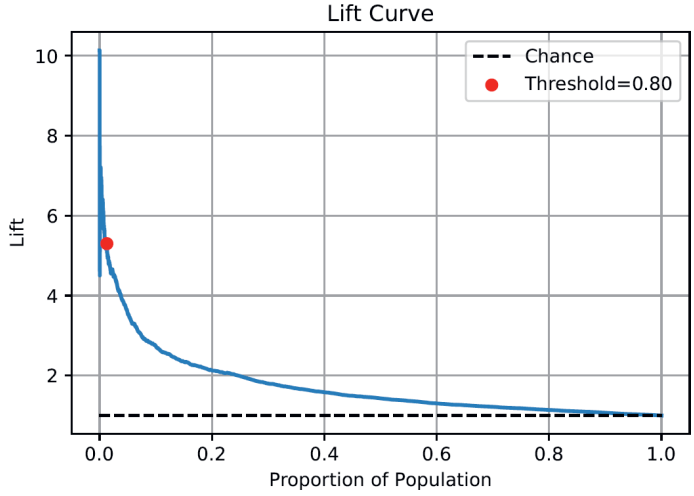


Figure 7.5 Lift curve for the ML model showing the lift value (y-axis) by the proportion of the population predicted to have an undetected SARS-CoV-2 infection (x-axis).

Discussion

We successfully developed a ML prediction model using retrospective data that appears to have suitable performance in identifying HD patients at risk of having an undetected SARS-CoV-2 infection that is identified in the following ≥3 days. The top predictors of a patient having a SARS-CoV-2 infection were the change in interdialytic weight gain from the previous month, mean pre-HD body temperature in the prior week, and the change in post-HD pulse from the previous month.

Albeit some top predictors are not surprising, the observed distinctions were subtle. Without insights from the model considering an array of variables, it would not be clear where one should classify a higher or lower risk for an individual patient that is meaningful. For instance, assessing for a decrease in weekly IDWG of about 0.3 kg alone may not be considered actionable, and the same is true for assessing for an increase of about 0.2°F (0.1°C) in weekly pre-HD body temperature, or an increase in pulse of about 1 beat per min (BPM). Notably, the average pre-HD body temperature was 97.6°F (36.4°C) (primarily oral measurements) in our analysis and has been previously reported as 98.2°F (36.7°C).[37] Given 98.6°F (37°C) is the expected average in healthy populations, the lower body temperature of HD patients is of importance with the

rather low incidence of fever presenting in dialysis patients with COVID-19 (11%-to-66% with fever[16,20,22]). Overall, the small changes observed for each individual variable suggest any one parameter alone has minimal value for detecting a patient's risk of having COVID-19, especially since every affected patient will not have every symptom of COVID-19 consistently. However, the combinations of minor changes appear to be meaningful in the individualized ML model we developed, with each small change being one piece of the puzzle for each patient's unique prediction.

Individual predictions can be further used to identify the risk level for dialysis clinics through the proportion of patients classified with an undetected SARS-CoV-2 infection. We anticipate using a combination of individual predictions along with reporting of the percent of patients at risk in each clinic may yield the greatest early insights on: 1) what otherwise asymptomatic HD patients might be most appropriate for enhanced screening, COVID-19 testing, and triage to an isolation area, and 2) where providers can focus additional resource allocations to combat COVID-19. Furthermore, flagging patients as potentially infectious may cut through some of the 'COVID fatigue' occurring during this prolonged pandemic. By adding this additional novelty and warning, the hope is additional care may be given in identifying of potential symptoms during screening. Prospective evaluation of ML model directed mitigation is currently being piloted at the national network of dialysis clinics.

The authors propose a conceptual workflow for the application of the ML model predictions to assist with directing care to individual patients and assisting with directing resource allocations to clinics (**Figure 7.6**). The model was trained using a target date of three days prior to patients presenting with COVID-19 symptoms to alert clinicians at least one dialysis treatment earlier. Given this timeline, we believe it is prudent to run the prediction model on a per treatment basis. The delivery of reports on individual patient predictions to clinic staff would optimally be delivered on interdialytic days to provide the care team time to prepare for a more comprehensive screening by an advanced clinician at the next encounter and potential isolation of subsequent HD treatments. The delivery of reports on the percent of patients in each clinic at risk can be performed on a weekly basis to allow leadership and regional managers to meet with clinical managers and prepare for allocation of resources including additional staff, protective equipment, and isolation areas. We propose categorizing clinic-level reports to detail facilities with more than 5% of patients at risk for undetected SARS-CoV-2 infection.

Mitigation efforts at the national dialysis network include universal RT-PCR testing of patients with symptoms of a flu-like illness along with distinct isolation areas (rooms, shifts, clinics) for suspected patients under investigation and COVID-19 positive patients. We propose patients predicted to be at risk receive a comprehensive

screening for signs and symptoms of a flu-like illness by an advanced practitioner (e.g. physician, physician assistant, nurse practitioner, experienced dialysis nurse) since there is a possibility of false positives. However, the comprehensive assessments should consider any minor sign or symptoms of a flu-like illness that may otherwise be considered normal based on the patient's uremia and medical history[38,39] to be a reason for suspicion of COVID-19. Given the predictions are derived for each individual patient, the reasons underlying the risk predictions can be provided for each patient, as well as the global importance of features for the model (**Figure 7.2**). This may help to provide additional insight into what the what a the more comprehensive screening assessment should focus on for each individual patient. For example, if a patient is classified by the model at risk with the top reason being related to a decrease in intradialytic weight gain, the next screening before entry to the clinic could include assessment of any change in appetite or fluid intake. High risk patients suspected with any mild sign of a flu-like illness could be triaged to unique isolation areas for patients under investigation and receive RT-PCR testing. HD would be continued in a distinct isolation area until diagnosis of COVID-19 or not (determined by two negative RT-PCR tests >24 hours apart), whereby laboratory positive patients would be triaged to unique isolation areas for COVID-19, and negative patients would return to be treated with the general HD population (**Figure 7.6**), which is consistent with the providers practices without the model. Patients diagnosed with COVID-19 at the provider are treated in distinct isolation areas until they have two negative RT-PCR tests >24 hours apart, after which recovered patients are transferred back to receive HD with the unaffected HD population.

The developed model has potential to provide a data-driven way for providers to identify individuals with undetected SARS-CoV-2 infections. The conceptual workflow provides a hypothetical strategy that can be adapted within the practice patterns of other providers, which may not include universal testing and require periods of isolation. Different strategies could utilize different thresholds for flagging patients depending on the intervention and implications of false positives and false negatives.

Considering the possibility of prolonged viral shedding observed in the general and dialysis populations[40-42], the optimal period for isolation of dialysis patients affected by COVID-19 appears to be longer than 14 days.[42] In countries or areas with testing limitations, especially those with a high positive-to- negative testing ratio (e.g. >25% positive test rate), it may be reasonable to consider having separate isolation areas for patients predicted at risk in addition to isolation areas for patients with symptoms of a flu-like illness. In this scenario, the 14-day timeframe for isolation of patients predicted to be at risk is anticipated to be appropriate if no signs or symptoms of a flu-like illness arise.

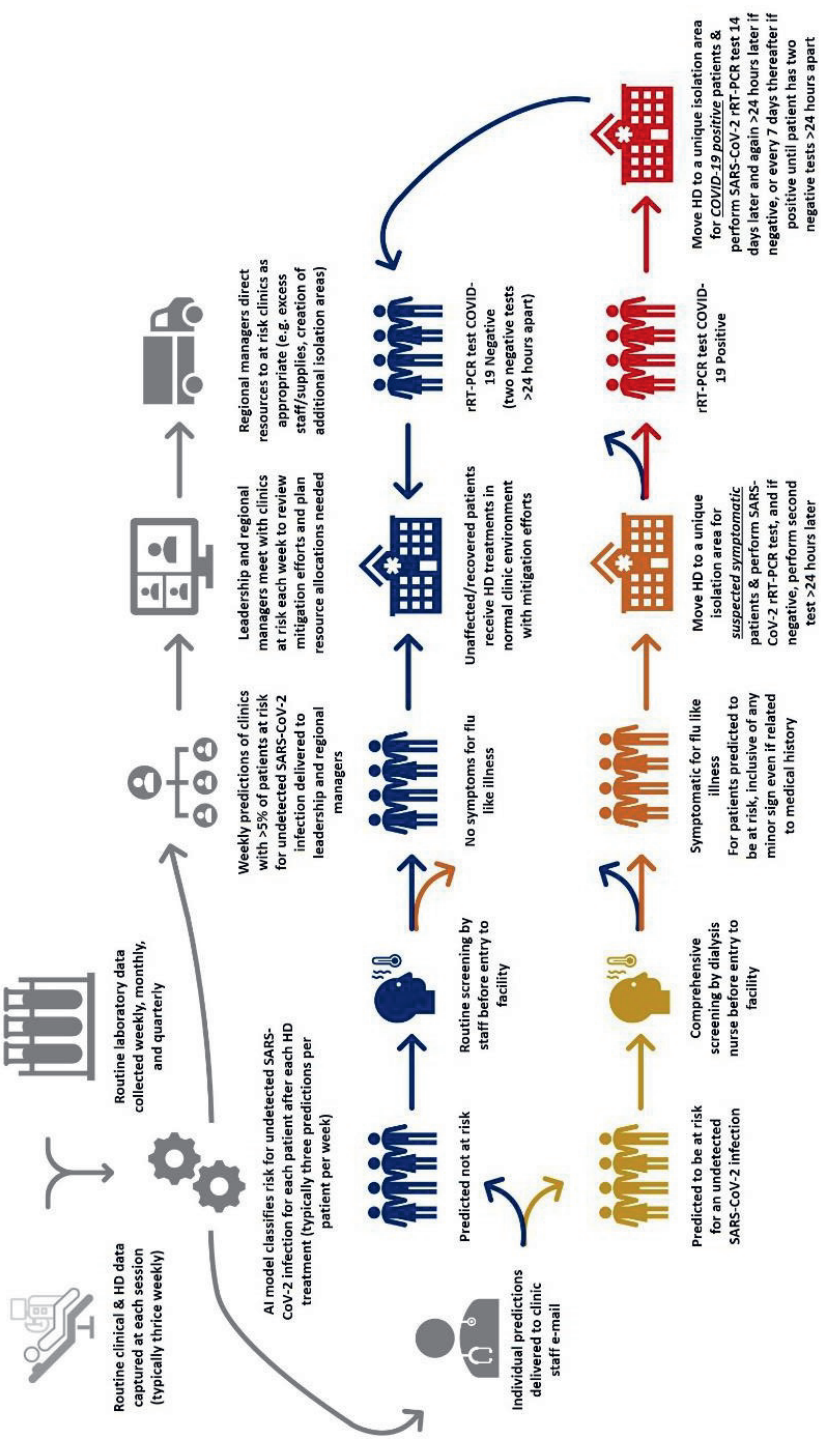


Figure 7.6 Conceptual workflow for application of ML model predictions within current mitigation and testing practices at the provider.

As more data is captured in the COVID-19 outbreak, further prediction models that can classify the risk of morbid/mortal outcomes in dialysis patients affected by COVID-19 need to be developed. The potential applications of AI for COVID-19 have been previously detailed[43]; the first priority was suggested as "early detection and diagnosis of the infection". The robustness of data and an *a priori* selection of variables to be included in our ML model bring value through assessment of feature importance; this allows for interpretation of meaningfulness of predictors, albeit it does not determine causality. The selection of input variables was focused on biological changes reflected in clinical presentations and biomarkers allowing the model to be generalizable to all individual HD patients in the overall population, and not specific to the characteristics of outbreaks or the local population where patients reside. Although this approach yields more generalizability for the model to be used in the HD populations worldwide, external factors such as local incidence rates or social determinants of health are anticipated to impact the likelihood of a patient contracting COVID-19 and can be considered as appropriate. Ultimately, this strategy has the potential to allow for COVID-19 to be detected sooner than HD patients show symptoms, and for a localized HD population, earlier than it would be reported by national authorities.

A systematic review identified several models developed using data from China for early detection of COVID-19 in suspected individuals in the general population.[27] One is an externally validated ML model that predicts COVID-19 in suspected asymptomatic patients (AUROC validation=0.872).[44]

Another effort used a prediction model (AUROC validation=0.966) to develop logic for an 8 variable COVID-19 risk chart.[45] A further model with an AUROC of 0.938 was created to detect COVID-19 pneumonia in patients admitting to a fever clinic.[46] Other models used genomic/computed tomography data to diagnose COVID-19.[27] An effort using data from China not included in prior reviews developed various ML models to predict (AUROC testing=0.87 to 0.95) and identify features indicative of COVID-19 status across age categories among people in the general population presenting to a clinic/hospital.[47] This model found the most important features for prediction of COVID-19 at presentation were lung infection, cough, pneumonia. Consistent variables used across models for predictions included age, body temperature, and flu-like illness symptoms.[27,47] Another distinct effort reported in the literature included the development of ML and traditional models using only full blood count data to predict the likelihood of a COVID-19 among people in the general population presenting to the emergency department (AUROC training=0.80 to 0.86) of, or patients admitted at (AUROC training=0.94 to 0.95), a large hospital in Brazil.[48] Although these models were all reported to have suitable performance, all were subject to bias due to non-generalizable sampling of controls without COVID-19 and possible overfitting. We cannot rule out that our ML model may have similar bias, although it included a large

sample and the testing dataset had relatively generalizable sampling for the dialysis population with respect to positives and negatives.[30,31] Also, since we randomly selected a subset of patients for the negative arm who never had symptoms of COVID-19 and did not receive PCR testing, it is possible that we might have unintentionally included a small number of patients who were asymptomatic. However, this would have required patients to have had an asymptomatic SARS-CoV-2 infection that aligned with the randomly sampled time window. Given the balanced class design of the training and validation data splits, it is unlikely this introduced a remarkable bias in the model during training and validation. Yet, there is a possibly this could have introduced a minimal bias in evaluation of performance in the testing data since there were fewer positive cases to identify to offset any impact of an incorrectly labeled negative patient as positive. Additionally, the reported model performance may be on the conservative side when considering the constraints of the "ground truth" labels as they relate to how positive patients are identified by conventional screening. The extent of this depends on how well the model identifies individuals not included in the training sample but might show similar patterns and on the intervention. In any case, our model is unique in its ability to identify the risk of SARS-CoV-2 infection in patients without any suspicion of being affected with the disease.

The developed model holds promise to help providers through the COVID-19 pandemic and subsequent wave(s) of outbreak.[49,50] We recommend model use as augmentation and not replacement of symptom screening, as AI modeling is never 100% accurate and model risk classifications need to be interpreted within the extent of the model's performance.

Conclusions

The developed AI model showed a clinically meaningful performance in prediction of individual HD patients at risk of having an undetected SARS-CoV-2 infection at least three days before there would be any suspicion of the disease. Prospective testing is needed and underway at the national network of dialysis clinics. We proposed a conceptual workflow for application of ML model directed mitigation and testing. These efforts should provide key insights for consideration by healthcare providers.

References

1. Kliger AS, Silberzweig J. Mitigating Risk of COVID-19 in Dialysis Facilities. Clin J Am Soc Nephrol 2020;15(5):707-709.
2. Ikizler TA. COVID-19 and Dialysis Units: What Do We Know Now and What Should We Do? Am J Kidney Dis 2020;76(1):1-3.
3. Basile C, Combe C, Pizzarelli F, Covic A, Davenport A, Kanbay M, Kirmizis D, Schneditz D, van der Sande F, Mitra S. Recommendations for the prevention, mitigation and containment of the emerging SARS-CoV-2 (COVID-19) pandemic in haemodialysis centres. Nephrol Dial Transplant 2020;35(5):737-741.
4. Mokrzycki MH, Coco M. Management of hemodialysis patients with suspected or confirmed COVID-19 infection: perspective of two nephrologists in the United States. Kidney360 2020;1(4):273-278
5. Gallieni M, Sabiu G, Scorza D. Delivering safe and effective hemodialysis in patients with suspected or confirmed COVID-19 infection: a single-center perspective from Italy. Kidney360 2020;1(5):403-409..
6. United States Renal Data System. 2019 USRDS annual data report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2019
7. Roncon L, Zuin M, Rigatelli G, Zuliani G: Diabetic patients with COVID-19 infection are at higher risk of ICU admission and poor short-term outcome. J Clin Virol 2020;127:104354.
8. Guo T, Fan Y, Chen M, Wu X, Zhang L, He T, Wang H, Wan J, Wang X, Lu Z. Cardiovascular Implications of Fatal Outcomes of Patients With Coronavirus Disease 2019 (COVID-19). JAMA Cardiol 2020;5(7): 811-818.
9. Li X, Xu S, Yu M, Wang K, Tao Y, Zhou Y, Shi J, Zhou M, Wu B, Yang Z, Zhang C, Yue J, Zhang Z, Renz H, Liu X, Xie J, Xie M, Zhao J. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. J Allergy Clin Immunol 2020;146(1):110-118.
10. Cheng Y, Luo R, Wang K, Zhang M, Wang Z, Dong L, Li J, Yao Y, Ge S, Xu G. Kidney disease is associated with in-hospital death of patients with COVID-19. Kidney Int 2020;97(5):829-838.
11. Du RH, Liang LR, Yang CQ, Wang W, Cao TZ, Li M, Guo GY, Du J, Zheng CL, Zhu Q, Hu M, Li XY, Peng P, Shi HZ. Predictors of Mortality for Patients with COVID-19 Pneumonia Caused by SARS-CoV-2: A Prospective Cohort Study. Eur Respir J 2020;55(5):2000524.
12. Adams ML, Katz DL, Grandpre J. Population-Based Estimates of Chronic Conditions Affecting Risk for Complications from Coronavirus Disease, United States. Emerg Infect Dis 2020;26(8):1831-1833.
13. Neumann ME: Latest data show 305 dialysis patient deaths due to COVID-19 in the US. Nephrology News & Issues, (Accessed 22 Apr 2020) https://www.healio.com/nephrology/infection-control/news/online/%7B3a263aa9-ad59-4c3f-aab7-07b8395508e5%7D/latest-data-show-305- dialysis-patient-deaths-due-to-covid-19-in-the-us, 2020
14. Team CC-R: Geographic Differences in COVID-19 Cases, Deaths, and Incidence - United States, February 12-April 7, 2020. MMWR Morb Mortal Wkly Rep 2020;69:465-471.
15. Johns Hopkins Coronavirus Resource Center. Johns Hopkins University School of Medicine, (Accessed 15 Oct 2020) https://coronavirus.jhu.edu/data/mortality, 2020
16. ERACODA - The ERA-EDTA COVID-19 Database for Patients on Kidney Replacement Therapy. ERA- EDTA, (Accessed 15 Oct 2020) https://www.era-edta.org/en/wp-content/uploads/2020/04/ERACODA-Study-Report-2020-04-29.pdf, 2020
17. Wang H. Maintenance Hemodialysis and Coronavirus Disease 2019 (COVID-19): Saving Lives With Caution, Care, and Courage. Kidney Med 2020;2(3):365-366.
18. Jager KJ, Kramer A, Chesnaye NC, Couchoud C, Sanchez-Alvarez JE, Garneata L, Collart F, Hemmelder MH, Ambuhl P, Kerschbaum J, Legeai C, Dolores Del Pino YPM, Mircescu G, Mazzoleni L, Hoekstra T, Winzeler R, Mayer G, Stel VS, Wanner C, Zoccali C, Massy ZA. Results from the ERA- EDTA Registry indicate a high mortality due to COVID-19 in dialysis patients and kidney transplant recipients across Europe. Kidney Int 2020;98(6):1540-1548.
19. Physicians C-TFCotJAoD, Japanese Society for Dialysis T, Japanese Society of N, Kikuchi K, Nangaku M, Ryuzaki M, Yamakawa T, Hanafusa N, Sakai K, Kanno Y, Ando R, Shinoda T, Nakamoto H, Akizawa T: COVID-19 of dialysis patients in Japan: Current status and guidance on preventive measures. Ther Apher Dial 2020;24:361-365.

20. Ma Y, Diao B, Lv X, Zhu J, Liang W, Liu L, Bu W, Cheng H, Zhang S, Yang L, Shi M, Ding G, Shen B, Wang H. 2019 novel coronavirus disease in hemodialysis (HD) patients: Report from one HD center in Wuhan, China. medRxiv: 2020.2002.2024.20027201, 2020
21. Siordia JA, Jr. Epidemiology and clinical features of COVID-19: A review of current literature. J Clin Virol 2020;127:104357.
22. Xiong F, Tang H, Liu L, Tu C, Tian JB, Lei CT, Liu J, Dong JW, Chen WL, Wang XH, Luo D, Shi M, Miao XP, Zhang C. Clinical Characteristics of and Medical Interventions for COVID-19 in Hemodialysis Patients in Wuhan, China. J Am Soc Nephrol 2020;31:1387-1397.
23. Niiler E. An AI Epidemiologist Sent the First Warnings of the Wuhan Virus. Wired, (Accessed 22 Apr 2020) https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/, 2020
24. Bogoch, II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. J Travel Med 2020;27(2):taaa008.
25. McCall B. COVID-19 and artificial intelligence: protecting health-care workers and curbing the spread. Lancet Digit Health 2020;2:e166-e167.
26. Alimadadi A, Aryal S, Manandhar I, Munroe PB, Joe B, Cheng X. Artificial intelligence and machine learning to fight COVID-19. Physiol Genomics 2020;52:200-202.
27. Wynants L, Van Calster B, Bonten MMJ, Collins GS, Debray TPA, De Vos M, Haller MC, Heinze G, Moons KGM, Riley RD, Schuit E, Smits LJM, Snell KIE, Steyerberg EW, Wallisch C, van Smeden M. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ 2020;369:m1328.
28. COVID-19 Resource and Education Center. Fresenius Medical Care North America, (Accessed 31 Aug 2020) https://fmcna.com/company/covid-19-resource-center/,
29. Chen T, Guestrin C: XGBoost. A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA, Association for Computing Machinery. 2016:785–794
30. Preliminary Medicare COVID-19 Data Snapshot. Centers for Medicare & Medicaid Services, (Accessed 14 Oct 2020) https://www.cms.gov/research-statistics-data-systems/preliminary-medicare-covid-19-data-snapshot, 2020
31. Anand S, Montez-Rath M, Han J, Bozeman J, Kerschmann R, Beyer P, Parsonnet J, Chertow GM. Prevalence of SARS-CoV-2 antibodies in a large nationwide sample of patients on dialysis in the USA: a cross-sectional study. Lancet 2020;396(10259):1335-1344.
32. Shapley LS. “A Value for n-Person Games,” In: H. W. Kuhn and A. W. Tucker, Eds. Contributions to the Theory of Games II. Annals of Mathematics Studies, Princeton University Press, Princeton 1953;28: 307-317.
33. Štrumbelj E, Kononenko I. Explaining prediction models and individual predictions with feature contributions. J Knowledge and Information Systems 2013;41:647-665.
34. Lundberg S, Lee SI. "A Unified Approach to Interpreting Model Predictions." In: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, Eds., Advances in Neural Information Processing Systems 30. Curran Associates, Inc: 2017:4765-4774
35. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, Lee S-I. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence 2020;2:56-67.
36. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 2015;10:e0118432.
37. Usvyat LA, Kotanko P, van der Sande FM, Kooman JP, Carter M, Leunissen KM, Levin NW. Circadian variations in body temperature during dialysis. Nephrol Dial Transplant 2012;27:1139-1144.
38. Gedney N. Long-Term Hemodialysis during the COVID-19 Pandemic. Clin J Am Soc Nephrol 2020;15: 1073- 1074.
39. Gagliardi I, Patella G, Michael A, Serra R, Provenzano M, Andreucci M. COVID-19 and the Kidney: From Epidemiology to Clinical Practice. J Clin Med 2020;9(8):2506.

40. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Chen H, Cao B. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395:1054-1062.
41. Fontana F, Giaroni F, Frisina M, Alfano G, Mori G, Lucchi L, Magistroni R, Cappelli G. SARS-CoV-2 infection in dialysis patients in northern Italy: a single-centre experience. Clin Kidney J 2020;13:334-339.
42. Shaikh A, Zeldis E, Campbell KN, Chan L. Prolonged SARS-CoV-2 Viral RNA Shedding and IgG Antibody Response to SARS-CoV-2 in Patients on Hemodialysis. Clin J Am Soc Nephrol 2021;16(2):290-292.
43. Vaishya R, Javaid M, Khan IH, Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr 2020;14:337-339.
44. Meng Z, Wang M, Song H, Guo S, Zhou Y, Li W, Zhou Y, Li M, Song X, Zhou Y, Li Q, Lu X, Ying B. Development and utilization of an intelligent application for aiding COVID-19 diagnosis. medRxiv: 2020.2003.2018.20035816, 2020
45. Song C-Y, Xu J, He J-Q, Lu Y-Q. COVID-19 early warning score: a multi-parameter screening tool to identify highly suspected patients. medRxiv: 2020.2003.2005.20031906, 2020
46. Feng C, Huang Z, Wang L, Chen X, Zhai Y, Zhu F, Chen H, Wang Y, Su X, Huang S, Tian L, Zhu W, Sun W, Zhang L, Han Q, Zhang J, Pan F, Chen L, Zhu Z, Xiao H, Liu Y, Liu G, Chen W, Li T. A Novel Triage Tool of Artificial Intelligence Assisted Diagnosis Aid System for Suspected COVID-19 pneumonia In Fever Clinics. medRxiv: 2020.2003.2019.20039099, 2020
47. Ahamad MM, Aktar S, Rashed-Al-Mahfuz M, Uddin S, Lio P, Xu H, Summers MA, Quinn JMW, Moni MA: A machine learning model to identify early stage symptoms of SARS-Cov-2 infected patients. Expert Syst Appl 2020;160:113661.
48. Banerjee A, Ray S, Vorselaars B, Kitson J, Mamalakis M, Weeks S, Baker M, Mackenzie LS. Use of Machine Learning and Artificial Intelligence to predict SARS-CoV-2 infection from Full Blood Counts in a population. Int Immunopharmacol 2020;86:106705.
49. Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet 2020;395(10233):1382-1393.
50. Xu S, Li Y. Beware of the second wave of COVID-19. Lancet 2020;395(10233):1321-1322.

Chapter 8

Predicting Mortality Risk in Dialysis: Assessment of Risk Factors using Traditional and Advanced Modelling Techniques within the MONDO Initiative

Sheetal Chaudhuri, John Larkin, Murilo Guedes, Yue Jiao, Peter Kotanko, Yuedong Wang, Len Usvyat, Jeroen P. Kooman

Hemodialysis International. 2023;27(1):62-73

Abstract

Introduction
Several factors affect the survival of End Stage Kidney Disease (ESKD) patients on dialysis. Machine learning (ML) models may help tackle multivariable and complex, often non-linear predictors of adverse clinical events in ESKD patients. In this study, we used advanced ML method as well as a traditional statistical method to develop and compare the risk factors for mortality prediction model in hemodialysis (HD) patients.

Materials and methods
We included data HD patients who had data across a baseline period of at least one year and one day in the internationally representative Monitoring Dialysis Outcomes (MONDO) Initiative dataset. Twenty-three input parameters considered in the model were chosen in an a priori manner. The prediction model used 1 year baseline data to predict death in the following 3 years. The dataset was randomly split into 80% training data and 20% testing data for model development. Two different modeling techniques were used to build the mortality prediction model.

Findings
A total of 95,142 patients were included in the analysis sample. The area under the receiver operating curve (AUROC) of the model on the test data with XGBoost ML model was 0.84 on the training data and 0.80 on the test data. AUROC of the logistic regression model was 0.73 on training data and 0.75 on test data. Four out of the top five predictors were common to both modelling strategies.

Discussion
In the internationally representative MONDO data for HD patients, we describe the development of a ML model and a traditional statistical model that was suitable for classification of a prevalent HD patient's three-year risk of death. While both models had a reasonably high AUROC, the ML model was able to identify levels of hematocrit (HCT) as an important risk factor in mortality. If implemented in clinical practice, such proof-of-concept models could be used to provide pre- emptive care for HD patients.

Introduction

There are several factors affecting the survival of End Stage Kidney Disease (ESKD) patients on dialysis. In addition to demographic factors such as age and gender, fluid overload, inflammation, serum phosphate levels, indications of malnutrition and loss of lean tissue mass assessed by bioimpedance spectroscopy (BIS) have a strong association with clinical outcomes.[1-5] Adding to the complexities related to such multiple prognostic markers, many of them have a bimodal relation with clinical outcomes and are also dependent on the interaction with other parameters. For example, blood pressure, inter-dialytic weight gain (IDWG), and phosphate have non-linear, bimodal, associations with worse outcomes. Also illustrative, the association between systolic blood pressure and clinical outcomes are modified by the fluid status, whereas the associations between phosphate and clinical outcomes are influenced by nutritional state.[6,7] While we know inflammatory, nutritional, body composition, and fluid overload markers are associated to clinical outcomes in hemodialysis (HD) patients, they have not been studied together in a multi-modal analysis using advanced analytical models. Furthermore, current risk prediction models lack detailed assessments of fluid state and body composition.

There are several differences between traditional statistical methods vs advanced analytical techniques.[8] Traditional statistical methods are easily interpretable, while advanced machine learning (ML) techniques are powerful at analyzing complex data formats such as audio or images. Advanced ML models can deal with such non-linear relationships, efficiently handle missing data, and thus could be used to enhance risk prediction models in HD patients.

The goal of this study is to understand if it is feasible to develop a prediction model to predict mortality in HD patients using a large globally representative dialysis database with clinical parameters such as nutritional, inflammatory, hydration, anemia, and mineral metabolism related parameters collected consistently across different regions in various electronic medical records (EMR). In this effort, we also assessed the application of advanced analytical ML method as well as simple traditional statistical method and highlighted how the output of the two approaches are similar and/or different. We further attempted to explain the results of the advanced ML model while comparing it to the output of the traditional model which has been generally well understood in the clinical community.

Materials and design

Patient cohort

This was a retrospective observational cohort study that used the Monitoring Dialysis Outcomes (MONDO) Initiative dataset.[9,10] We included data from all unique HD patients who had data across a baseline period of at least one year and one day (i.e., ≥2 HD treatment records ≥1 year & 1 day apart). The follow-up period for the assessment of mortal outcomes was up to three years. The research activities conducted in the MONDO Initiative comply with all applicable national and international ethical standards. Western Institutional Review Board (Puyallup, WA, United States) approved a protocol on the MONDO initiative study projects and determined analyses are exempt due to use of de-identified data (Work Order 1-939512-1). Furthermore, the MONDO dataset has had a re-identification risk assessment performed by Privacy Analytics (Ottawa, ON, Canada) and has been fully anonymized using techniques that satisfy the concept of anonymization by the European Union General Data Protection Regulation (GDPR) and satisfy the concept of de-identification by the United States Department of Health and Human Services (HHS). This study was performed in adherence with the Declaration of Helsinki.

Model data and features

The outcome of all-cause mortality (dependent variable) was recorded during a 3-year follow-up period after baseline. The clinical data was collected per the standard practices in each country.

While most patients in regions receive HD treatment thrice a week, it could vary from region to region. Only those input parameters that were consistently captured across regions were chosen in an a priori manner and considered in the model (independent variables). Patient characteristics such as age and body mass index (BMI) were calculated as the maximum recorded value during the baseline period. BMI was calculated by dividing maximum of the post HD weight and the square of maximum height in meters. Laboratory and clinical parameters such as albumin, normalized protein catabolic rate (NPCR), bicarbonate, calcium, creatinine (CREAT), c reactive protein (CRP), ferritin, hematocrit (HCT), phosphate, potassium, pre-HD weight, post-HD weight, white blood cell (WBC) count, and residual renal function (RRF) were derived from the average value during the baseline period. RRF was captured as the glomerular filtration rate determined by urine collection. BIS measurements such as fat tissue mass (FTM), lean tissue mass (LTM), and total body water (TBW) were also averaged over the baseline period. IDWG was calculated as the difference between pre-HD weight at the index treatment and post-HD weight at the prior treatment, and the

IDWG values were averaged over the baseline period. Overhydration (OH) was calculated as the difference between pre-weight and normal hydration weight measured by the BIS, and were averaged over the baseline period. **Figure 8.1** shows categories of various input parameters used in the model.

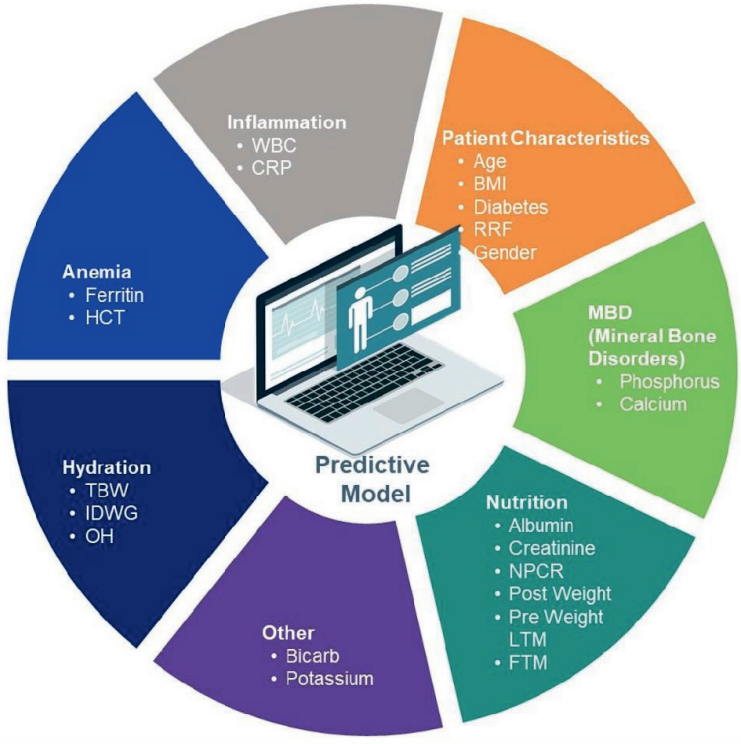


Figure 8.1 Groups of input parameters.

Predictive model

The prediction model used 1 year baseline data for in-center HD patients to predict death in the following 3 years. **Figure 8.2** shows the ascertainment period and prediction period of the advanced ML and logistic regression prediction model. The model only predicts mortality for patients who survived and have data for at least one year. The cohort dataset was randomly split into 80% training data and 20% testing data for model development.

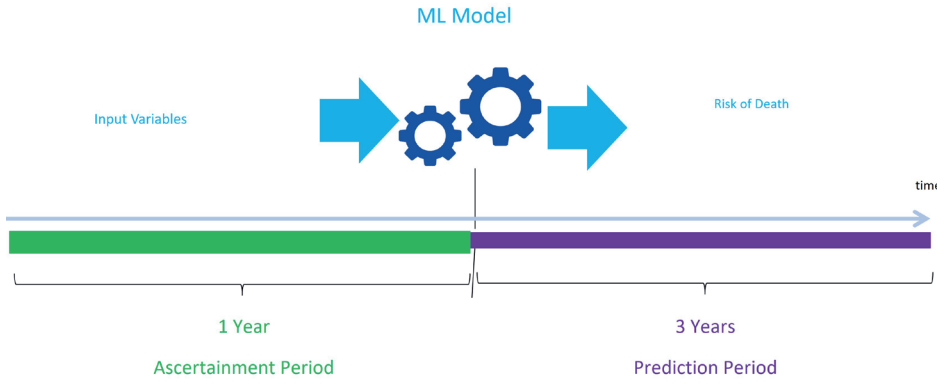


Figure 8.2 Ascertainment period and prediction period

Advanced analytical model development

Python version 3.7.7 (Python Software Foundation, Delaware) was used to develop the advanced ML model utilizing the XGBoost package.[11] The XGBoost Python package uses input parameters based on the training dataset to construct decision trees. Each decision tree provided a random sample and a series of thresholds that split parameters to maximize information gain. These decision trees are created iteratively and new decision trees are developed to minimize prior prediction errors.[12] The decision trees made by the XGBoost ML model can handle missing values without imputation by recognizing their presence when determining the splits. The ML model was constructed using the training data (80% of the cohort) and the final model was assessed using the unseen testing data (20% of the cohort).

Traditional analytical model development

SAS 9.4 was used to build the logistic regression model. Stepwise logistic regression model (slentry=0.3 and slstay=0.35)was developed using the same input parameters and the training data as the advanced ML model.[13]

Imputation in the form of mean or median was used to fill in data where it was incomplete. The stepwise logistic regression model was also tested on the 20% test data.

Analysis of ML and logistic regression models performance

Performance of the advanced ML model and the logistic regression model was evaluated by the area under the receiver operating curve (AUROC) in the training and testing datasets.[14]

Analysis of feature importance advanced ML model

Shapley values are calculated using the SHAP Python package to define the effect of each parameter on the predictions.[15,16] SHAP values are computed for each input parameter, representing a measure of effect (positive or negative value) of the input parameter on each individual prediction. SHAP methods withhold and include all combinations of individual input parameters and then compare differences between withheld and included data. Mean value of all possible differences is then used to calculate the feature importance. SHAP values are additive explanations of feature importance and are presented as log odds (i.e., the logarithm of the odds ratio). SHAP values for each set of observations are summed, and converted from log odds to probability, which is then output by the model as the prediction. Positive SHAP values increase the predicted probability, whereas negative SHAP values decrease the predicted probability.

Partial dependence plots (PDP) were created using the SHAP values to analyze the bimodal associations between two parameters and the impact on mortality. It shows the localized effects of two parameters on the predicted outcome. It also shows whether the associations are linear or more complex.

Analysis of feature important in traditional model

The summary statistics from the logistic regression model shows a list of the input parameters from the training dataset in the order of importance. P-value of <0.05 was considered statistically significant. Odds Ratio Estimates (OR) estimate from the stepwise logistic regression model shows the association between the risk factor and outcome. It represents the odds or the probability of the risk factor altering the predicted outcome.[17]

Results

Patient characteristics

Out of 150,496 unique patients in the MONDO Initiative dataset, 95,142 patients who had data recorded across a baseline period of one year and one day were included and assessed during a three year follow up period. **Figure 8.3** shows a flow diagram of the data used in the study. The overall follow up time was an average 2.8 years. Among patients who died, the average follow-up time was years. Among those who survived, the average follow- up time was 3 years. The majority was male (57.4%) with an average age of 61.7 years and 62% of them were diagnosed with diabetes mellitus.

Table 8.1 shows the regional spread of the 95,142 patients. **Table 8.2** shows the descriptive statistics of the numeric input parameters during the baseline period.

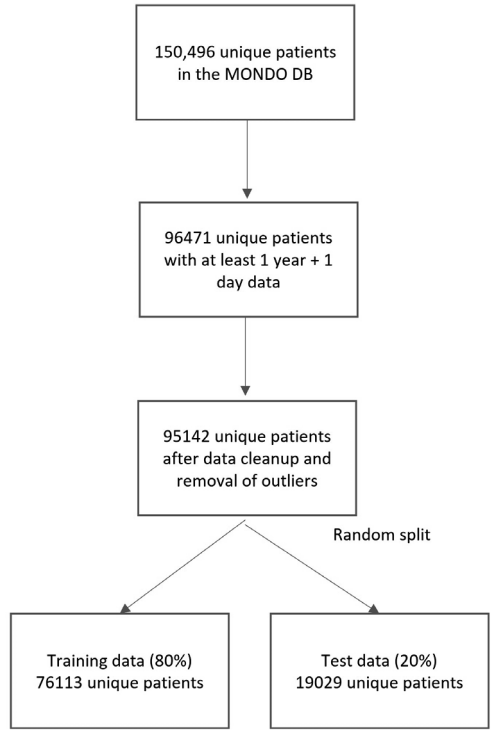


Figure 8.3 Flow diagram of study data.

ML model performance and feature importance

The resulting advanced analytical predictive model was tested on 20% of the patient's data, which was withheld and unseen during training. The AUROC of the model on the test data with XGBoost ML model was 0.84 on the training data and 0.80 on the test data. **Figure 8.4** shows the AUROC of the XGBoost ML model. Similarly, **Figure 8.5** shows the AUROC curve from the logistic regression model. It was 0.73 on the training and 0.75 on the test data.

Table 8.1 Distribution of patients by region.

Region	Count	Percentage
Southern Europe	20178	21.21%
Eastern Europe	19705	20.71%
South America	17530	18.43%
Eastern Asia	12402	13.04%
Western Asia	9465	9.95%
Northern Europe	7046	7.41%
Northern America	4419	4.64%
Western Europe	2825	2.97%
Southeastern asia	805	0.85%
Oceania	731	0.77%
Other	36	0.04%
Total Worldwide	95142	100%

Table 8.2 Descriptive statistics of numeric input parameters.

	Count	Mean	Std	Min	25%	50%	75%	Max
Albumin [g/L]	52099	3.78	0.42	1.36	3.54	3.81	4.05	6.00
NPCR [g/kg/day]	37324	1.00	0.24	0.01	0.92	1.01	1.08	22.64
Bicarb [mEq/L]	22361	22.41	2.86	2.00	20.67	22.43	24.20	45.00
Calcium [mg/dL]	43821	8.86	0.64	4.01	8.48	8.84	9.21	14.92
Creat [mg/dL]	44172	7.40	2.48	0.30	5.67	7.15	8.84	25.00
CRP [mg/dL]	39576	11.07	16.12	0.10	1.65	5.10	13.35	160.00
Ferritin [ng/mL]	52340	418.08	257.85	2.00	230.72	370.03	551.40	1650.00
HCT [%]	38380	34.19	3.62	15.00	32.14	34.39	36.50	53.00
Phosph [mg/dL]	54362	4.27	1.59	1.00	3.47	4.46	5.27	18.29
Postweight [kg]	46141	69.95	16.49	22.35	58.51	67.97	78.93	212.12
Preweight [kg]	46284	71.95	16.73	31.70	60.38	69.96	81.09	214.40
Potassium [mEq/L]	54151	4.87	0.62	2.15	4.45	4.84	5.26	8.70
IDWG [kg]	46128	1.99	0.83	0.00	1.43	1.95	2.48	14.38
WBC [1000/mc]	35527	6.91	2.72	0.00	5.63	6.81	8.13	100.00
RRF [mL/min]	3168	5.77	5.75	0.00	2.48	4.40	8.00	97.00
BSI FTM [%]	2297	22.72	10.08	5.06	15.11	21.43	29.04	54.39
BSI LTM [%]	2308	46.96	6.09	40.00	42.47	45.29	49.70	84.50
BSI TBW [%]	2345	41.83	5.15	30.11	38.20	41.21	44.84	71.62
Age [Yrs]	64313	61.73	15.08	18.00	52.00	64.00	73.00	90.00
BMI [kg/m^2]	18706	25.13	5.53	11.68	21.49	24.22	27.63	83.19
OH [kg]	2334	1.36	3.11	-14.74	-0.11	1.49	2.90	14.65

NPCR: normalized protein catabolic rate; Bicarb: bicarbonate; Creat: Creatinine; CRP: C reactive protein;HCT: heamtocrit; IDWB: inter-dialytic weight Gain; WBC: white blood cell; RRF: residual renal function; BIS: bioimpendance spectroscopy; FTM: fat tissue mass; LTM: lean tissue mass; TBW: total body water; BMI: body mass index; OH: overhydration = preweight – normal hydration weight (BIS measurement).

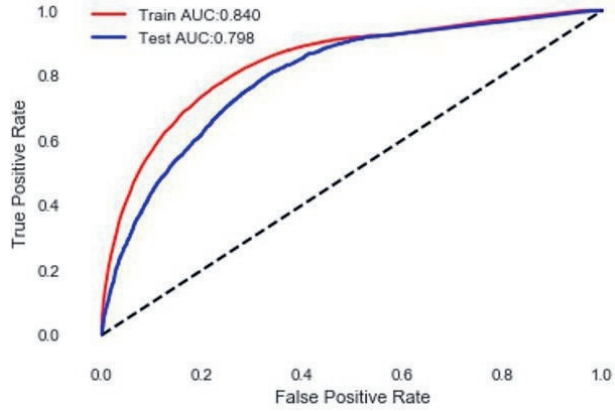


Figure 8.4 Area under the curve for XGBoost model.

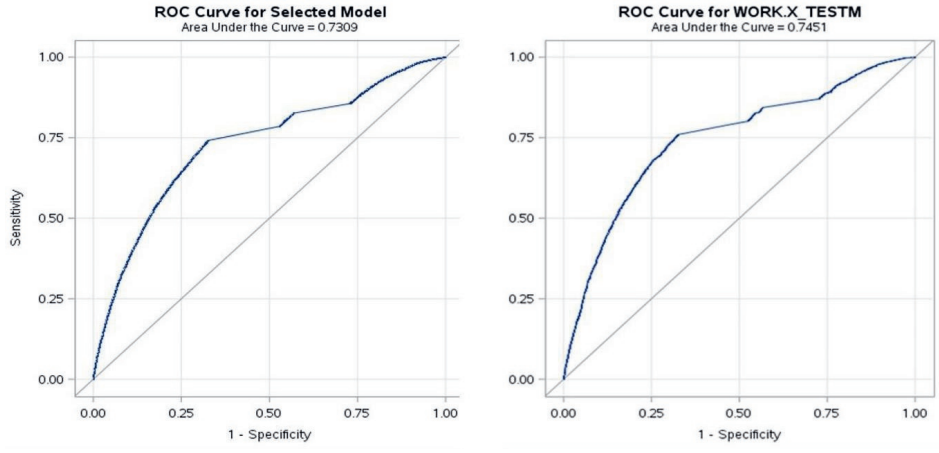


Figure 8.5 Area under the receiver operating curve(AUROC) for logistic regression model with training and test data.

Important features using SHAP values

Figure 8.6A shows the magnitude of the SHAP value for the top predictors identified by the XGBoost Model in a descending order of importance. Warmer colors on the figure shows a higher impact of the parameter in predicting mortality, while cooler colors show a the negative (i.e., protective) impact of the parameter on predicting mortality. Patients with higher age, lower HCT, lower albumin, lower CREAT, higher CRP, lower NPCR, higher IDWG, higher ferritin, higher phosphorous, higher WBC, presence of

diabetes, lower potassium, lower BMI, low RRF, lower pre and post weight have a higher chance of mortality in the following three years. Similarly, higher CREAT, higher HCT, higher NPCR, higher BICARB, lower overhydration, higher BIS measurements of FTM, lower LTM and slightly higher TBW have a lower chance of mortality in the following three years. **Figure 8.6B** shows the absolute value of the average SHAP value that is the value shows on average how much each predictor impacts the mortality prediction either in the positive or negative direction.

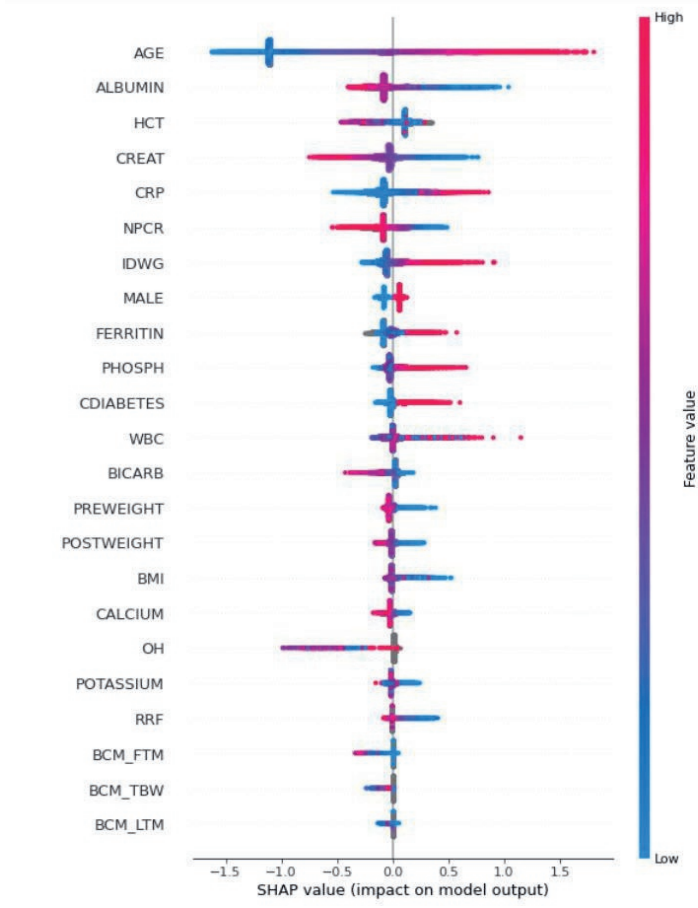


Figure 8.6A SHAP value plots show the size and direction (more positive=higher risk or more negative=lower risk) of each variable's influence on the outcome for each unique patient on the x- axis, with warmer colors representing higher observed values for that measurement, cooler colors indicating lower values for that measurement, and gray representing a missing value for that measurement. SHAP values are presented in the unit of log odds (i.e. logarithm of the odds ratio).

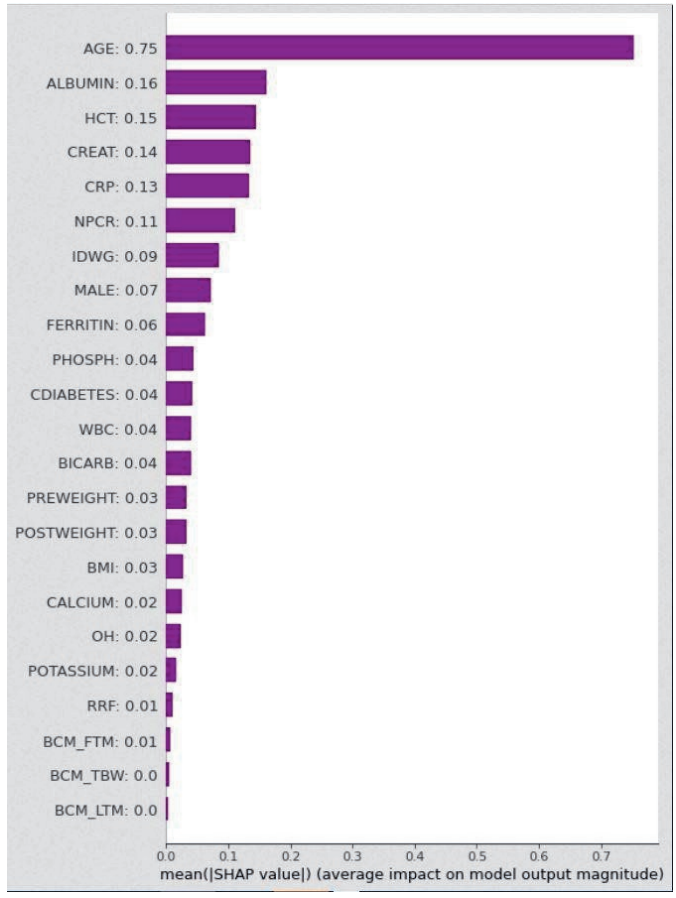


Figure 8.6B Absolute value of the average SHAP value shows on average how much each predictor impacts the mortality prediction

Figures 8.7A and 8.7B show a couple of PDP plots for individual parameters used in the model. These plots help understand the interaction of select top predictors in predicting mortality among different domains shown in **Figure 8.1**. Each dot corresponds to an individual person in the study. The dot's position on the x-axis shows the impact that parameter has on the model's prediction for that person. Multiple dots in the same position show density. Risk of mortality based on the SHAP value of the parameter on the X axis is shown on the Y axis.

Figure 8.7 Partial Dependence Plots: Each dot corresponds to an individual person in the study. The dot's position on the x-axis shows the impact that parameter has on the model's prediction for that person. Multiple dots in the same position show density. Risk of mortality based on the SHAP value of the parameter on the X axis is shown on the Y axis.

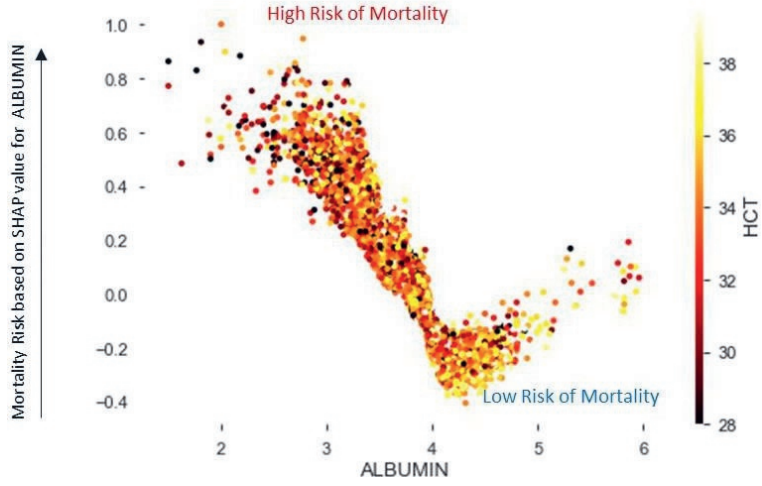


Figure 8.7A Partial dependence plots for nutrition related parameters.

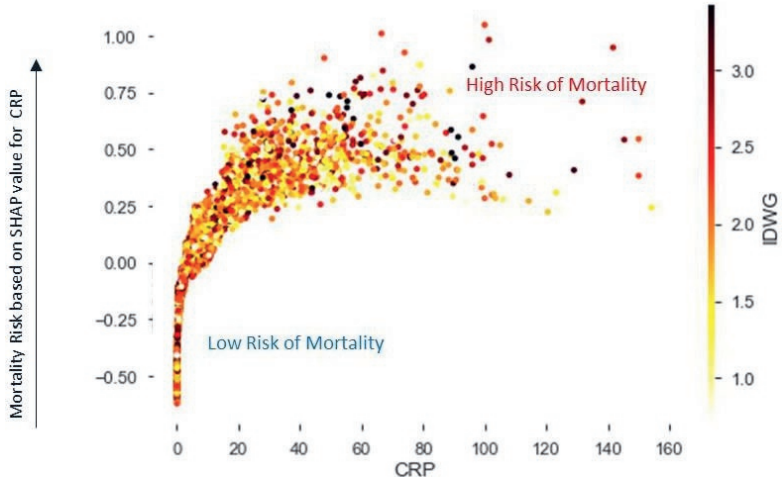


Figure 8.7B Partial dependence plots for hydration related parameters.

The first PDP plot shows the association between a selected nutrition related parameter (albumin) and an anemia related parameter (HCT). The graph shows the associations between variables, as well as the impact on mortality. Patients with lower albumin and lower HCT have the highest risk of mortality predicted as shown by the dark red dots on the top left corner of the first subplot **Figure 8.7A**. Patients with higher albumin and higher HCT have the lowest impact on mortality as shown by the cluster of light-yellow colored dots on the bottom right corner of the plot.

The second PDP plots show the association between a selected inflammation related parameter (CRP) and a hydration related parameter (IDWG) on mortality. Higher CRP and higher IDWG have a higher impact on mortality as indicated by the dark red color clusters in the top righthand corner of the graph. Patients with lower CRP have lower risk of mortality regardless of the IDWG.

Results from the logistic regression model

Table 8.3 shows the parameters from the stepwise logistic regression model in a descending level of significance along with the odds ratio estimate and 95% confidence interval. The results show most of the input parameters other than LTM, FTM, TBW, HCT and pre-weight have a statistically significant ($p<0.05$) impact on the output of the logistic regression model. Male gender, higher age, higher levels of CRP, ferritin, phosphorous, IDWG, WBC, RRF, LTM, OH and BMI associates with an increased risk of mortality in the follow-up period. The assessment between the directionality provided by the odds ratio estimates and the SHAP values is presented by the color coding in Table 8.3. The odds ratio estimates shown in red font indicate the risk of mortality increases with the increasing value of the input parameter in a consistent manner with the results shown in the SHAP values from the XGBoost model. The odds ratio estimates shown in a blue font indicate the risk of mortality decreases with increasing value of the input parameter in a consistent manner with the results shown by SHAP values. The odds ratio estimates without any color indicate that the directionality of mortality risk in the SHAP value and the logistic regression are not the same.

Discussion

In the internationally representative MONDO Initiative cohort of hemodialysis patients from dialysis providers throughout 37 countries in 5 continents[10], we describe the development of a ML model and a traditional statistical model that was suitable for classification of a prevalent HD patient's three- year risk of death. The performance of both techniques (AUROC 0.80 for ML model 0.75 for traditional model) was high and much better than chance (i.e., AUROC 0.50). Furthermore, due to the inherent ability

of the ML modeling technique to account for collinearity between input variables and handle missing data, we were able to identify important predictors for death events, such as HCT – the third most important prognostic risk factor. Importantly, the prediction model used a set of input parameters that is frequently collected in the standard care of HD patients by most dialysis providers worldwide. Thus, our proof-of-concept model and findings are generalizable, and the model may be scalable for prognostic decision support in the care of dialysis patients.

There have been other mortality prediction models developed in dialysis patients, however, most models have overestimated the probability of death.[18] **Table 8.4** references several of these studies along with the size and the geographic distribution of the cohort. The AUROC, or the C-statistic, of most models described is lower than the AUROC of the advanced ML model described in this paper. There are a couple reports on mortality models where the AUROC was comparable with our ML model, however, they were tested on considerably smaller cohorts.[19,20] A strength of our model is it was trained and tested on a large cohort of patients geographically distributed across the globe who were treated by multiple providers.[10]

It is worthwhile to detail the ML modeling technique we used is not based on linear relationships for the input variables. This can be advantageous since the method does not assume linearity and can avoid bias from u-, j-, and other irregular-shaped associations with an independent variable and an outcome, yet the interpretability of the coefficients can be difficult since they are reported in log odds. Unlike the ML model in this study, traditional modeling models (e.g., logistic regression, Cox proportional hazards) frequently assume linearity in relationships and are unable to handle missing data. Hence, imputation of missing data becomes necessary like we performed for the logistic regression model assessed in our study. Traditional models can be easier to interpret with odds or hazards ratios providing the proportion of risk per unit of change in the parameter based off an assigned reference point. ML models are best suited for complex datasets especially with missing data.

Table 8.3 shows how the important variables identified by each modeling technique. Four out of the top five predictors from both models are common and display similar directionality. Patients with higher age, lower albumin, higher CRP, and lower creatinine have higher risk of mortality in the following three-years. Other studies have shown age, albumin and CRP are significant factors in predicting mortality in HD patients.[21] HCT did not appear significant in the logistic regression model yet was the third most important risk factor in the ML model. Anemia is a strong predictor of worse clinical and patient-reported outcomes, yet anemia correction has not been shown to improve clinical outcomes, except at severely low hemoglobin levels for currently available interventions.[22,23] As anemia interventions, such as iron deficiency

management and erythropoietic stimulating agent prescription, may be linked to clinical outcomes regardless of hemoglobin levels, observational analysis of achieved hemoglobin levels may lead to important heterogeneity.[24-26] Thus, it may be that the interaction(s) with other parameters in the traditional models have obscured this association.

Table8.3 Summary of step-wise selection and Odds Ratio estimate.

Step	Entered	Chi-Square	Significance	Odds Ratio	95% Confidence Limits	
1	Age	2743.08	<.0001	**1.050**	1.048	1.053
2	CRP	901.22	<.0001	**1.017**	1.015	1.019
3	Albumin	560.65	<.00001	**0.486**	0.452	0.523
4	Creat	367.45	<0.0001	**0.835**	0.822	0.849
5	Cdiabetes	431.76	<0.0001	**0.555**	0.517	0.596
6	Phosph	190.47	<0.0001	**1.129**	1.106	1.154
7	IDWG	129.52	<0.0001	**1.255**	1.204	1.308
8	Ferritin	59.04	<0.0001	**1.000**	1.000	1.001
9	MaleE	43.28	<0.0001	**1.18**	1.124	1.239
10	NPCR	25.82	<0.0001	**0.406**	0.323	0.511
11	Bicarb	12.84	0.000	**0.970**	0.955	0.986
12	BMI	11.23	0.000	**1.030**	1.020	1.041
13	Postweight	33.08	<0.0001	**0.993**	0.990	0.995
14	WBC	7.52	0.006	**1.018**	1.006	1.031
15	OH	7.2225	0.0072	**1.094**	1.038	1.154
16	Potassium	4.0704	0.0436	**0.949**	0.903	0.998
17	Calcium	3.9427	0.0471	**0.95**	0.902	1.001
18	RRF	3.7162	0.0539	**1.016**	0.999	1.033
19	BCM_TBW	1.9393	0.1637	**0.956**	0.913	1.001
20	BCM_LTM	1.5422	0.2143	**1.046**	1.008	1.085

Red: Risk of mortality increases with the increasing value of the input parameter in a consistent manner with the results shown in the SHAP values from the XGBoost model; **Blue**: Risk of mortality decreases with the increasing value of the input parameter in a consistent manner with the results shown in the SHAP values from the XGBoost model; **Black**: The direction of the risk of mortality between the logistic regression model and the XGBoost model is not consistent.

Fluid related parameters identified by the ML model were IDWG, OH and TBW. OH and IDWG were statistically significant in the conventional logistic regression model and had similar direction as the advanced ML model. Patients with higher IDWG and higher OH had higher risk of mortality, thus emphasizing the essential role of fluid control in dialysis. The traditional model also suggests that the TBW is associated with lower mortality, however, it is important to note that this is not an important predictor (p=0.1637) in the logistic regression model. The ML model on the other hand shows TBW is the 22nd most important factor suggesting that it may be important for a select group of patients where the data is available. So having a TBW measurement itself may decrease the risk due to a more targeted therapy being delivered with the BIS. We see similar results for all other BIS parameters, where high OH slightly increases the risk of

death, but low OH is highly protective for those who have the data. For all these BIS measures, missing values are in the center suggesting they do not affect the risk and are missing at random.

Our ML model also showed patients with higher BMI had lower risk of mortality, which is consistent with previous findings.[27,28] Surprisingly, our logistic model found opposite signals with a small positive association in the risk of mortality with higher BMI. There were also contrasts between modeling techniques with higher RRF, where the ML model showed higher RRF was protective as expected, while the traditional model showed the risk of mortality slightly increased with increasing RRF. This result is likely the secondary to the imputation of missing values in traditional models. Advanced imputation techniques may have resulted in a different outcome. The inclusion of multiple covariates in a model can lead to biases if each associated parameter is interpreted causally often referred to as Table 8.2 fallacy.[29] Thus, caution is warranted in the interpretation of specific associations in our models, since the goal of our analysis is predictive, rather than causal.[30] For instance, inclusion in the model of any downstream variable in a causal pathway starting with RRF can potentially lead to collider bias, which can reverse the direction of expected associations.[31]

Table8.4 Summary of other mortality prediction models in HD patients.

Author	ML model	Number of patients/region	AUROC on test data
Siga et al.[35]	Bayesian Network	9010 (Europe)	0.78
Jung et al.[36]	Cox Proportional Hazards Regression Analysis	3309 (Korea)	0.74
Zhu et al.[37]	Cox Model	173 (China)	0.79
Wang et al.[38]	Long short Term Memory Autoencoder	1200 (China)	0.57
Tapak et al.[39]	Randon Survival Forests	785 (Iran)	0.80
Thijssen et al.[40]	Logistic Regression	6838 (North America)	0.72
Hemke et al.[41]	Cox Regression Analysis	1835 (Netherlands)	0.78
Doi et al.[20]	Logisitc Regression	688 (Japan)	0.83
Floege et al.[42]	Cox Model	11508 (Europe)	0.73
Cohen et al.[19]	Cox Model	512 (North America)	0.80
Holme et al.[43]	Cox Model	868 (Region Not Available)	0.72

The partial dependence plot created using the SHAP values show if bi-modal interactions between parameters exist in more than one domain. Although this paper presents only two of these relations amongst the important domains from the top predictors common to both the models, other interactions between input parameters from other domains could be evaluated. Another limitation of the model is that it did not study the impact and interaction of the cardiovascular parameters such as blood pressure measurements. There may be many other potentially important parameters including comorbidity burden that could be evaluated, and optimization for specific parameters more readily available at specific providers could be considered. This model

has only been tested on HD patients and may not be applicable to patients who are on peritoneal dialysis or other forms of home or in-center dialysis. Albeit death events were recorded during the provision of dialysis care worldwide in MONDO, the database did not capture data on patients who transitioned to other kidney replacement modalities (e.g. peritoneal dialysis, transplant), left the provider, or became lost to follow up and this could have the potential to introduce bias in the classification of survival. The model currently predicts death in the 3 years following the baseline period, however, the model could be designed to predict death at various other time periods. It is also important to note that in this analysis we have not considered interactions between the input parameters when developing the traditional models, it is possible that the performance of the traditional may have improved if we considered these interactions. Advanced ML model automatically disentangles these complexities of the relationship between input parameters. Many traditional models, when compared to ML models, assume linearity, and avoid known interactions. Thus, we may not be comparing the best traditional method with ML technique.[32] ML models are often considered a black box that are difficult to interpret how the model arrives at its prediction.[33] Since the goal of this paper is to study the risk factors associated with mortality and to compare the outputs of the two modeling techniques, the ML model was not optimized for performance using hyperparameter tuning and this could yield further improvements.[34] If fully optimized and implemented in clinical practice, ML models could provide a suitable method to help identify a patient's prognosis and drive pre-emptive interventions and/or care planning efforts. Additionally, MONDO database is undergoing an update with most recent data and additional data points being captured. This proof-of-concept model can be scaled to adapt to additional input parameters or different follow-up time periods.

Conclusion

This paper demonstrates the successful use of ML and traditional statistical techniques in predicting mortality on a large cohort of international HD patients using clinical parameters collected consistently across regions. It shows how the risk factors compare between the ML techniques and conventional modeling techniques. If implemented in clinical practice, such models could be used to provide pre- emptive care for HD patients.

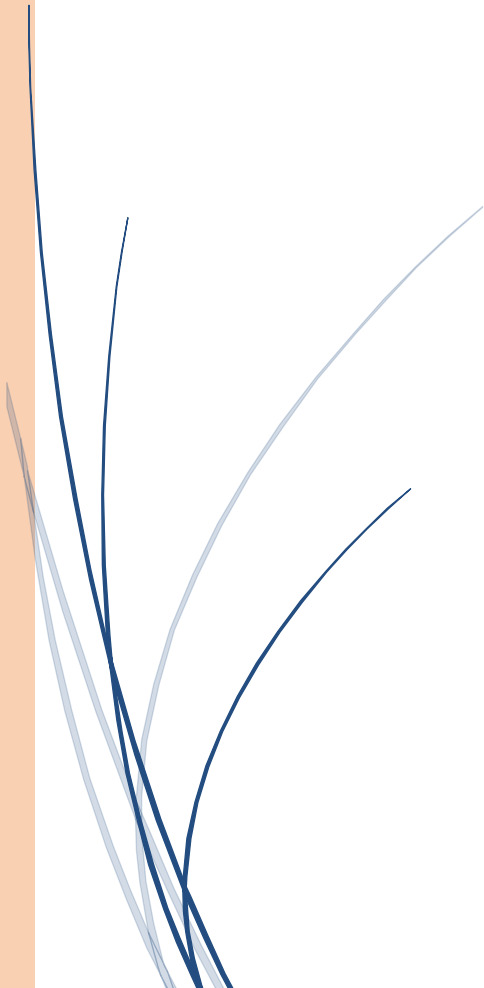
References

1. Marcelli D, Usvyat LA, Kotanko P, Bayh I, Canaud B, Etter M, et al. Body composition and survival in dialysis patients: results from an international cohort study. Clin J Am Soc Nephrol. 2015;10(7): 1192-1200.
2. Chen JB, Cheng BC, Yang CH, Hua MS. An association between time-varying serum albumin level and the mortality rate in maintenance haemodialysis patients: a five-year clinical cohort study. BMC Nephrol. 2016;17(1):117.
3. Dekker MJ, Marcelli D, Canaud BJ, Carioni P, Wang Y, Grassmann A, et al. Impact of fluid status and inflammation and their interaction on survival: a study in an international hemodialysis patient cohort. Kidney Int. 2017;91(5):1214-1223.
4. Tabinor M, Elphick E, Dudson M, Kwok CS, Lambie M, Davies SJ. Bioimpedance-defined overhydration predicts survival in end stage kidney failure (ESKF): systematic review and subgroup meta-analysis. Sci Rep. 2018;8(1):4441.
5. Broers NJH, Canaud B, Dekker MJE, van der Sande FM, Stuard S, Wabel P, et al. Three compartment bioimpedance spectroscopy in the nutritional assessment and the outcome of patients with advanced or end stage kidney disease: What have we learned so far? Hemodial Int. 2020;24(2):148-161.
6. Dekker M, Konings C, Canaud B, Carioni P, Guinsburg A, Madero M, et al. Pre-dialysis fluid status, pre-dialysis systolic blood pressure and outcome in prevalent haemodialysis patients: results of an international cohort study on behalf of the MONDO initiative. Nephrol Dial Transplant. 2018; 33(11):2027-2034.
7. Ye X, Kooman JP, van der Sande FM, Raimann JG, Usvyat LA, Wang Y, et al. Relationship between serum phosphate levels and survival in chronic hemodialysis patients: interactions with age, malnutrition and inflammation. Clin Kidney J. 2021;14(1):348-357.
8. Chaudhuri S, Long A, Zhang H, Monaghan C, Larkin JW, Kotanko P, et al. Artificial intelligence enabled applications in kidney disease. Semin Dial. 2021;34(1):5-16.
9. Usvyat LA, Haviv YS, Etter M, Kooman J, Marcelli D, Marelli C, et al. The MONitoring Dialysis Outcomes (MONDO) initiative. Blood Purif. 2013;35(1-3):37-48.
10. von Gersdorff GD, Usvyat L, Marcelli D, Grassmann A, Marelli C, Etter M, et al. Monitoring dialysis outcomes across the world--the MONDO Global Database Consortium. Blood Purif. 2013;36(3-4): 165-172.
11. Chen T GC. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining San Francisco, California, USA, Association for Computing Machinery. 2016:785–794.
12. Caitlin K. Monaghan JWL, Sheetal Chaudhuri, et.al. Machine Learning for Prediction of Patients on Hemodialysis with an Undetected SARS-CoV-2 Infection. Kidney 360. 2021;https://doi.org/10.34067/KID.0003802020
13. Arunajadai SG. Stepwise logistic regression. Anesth Analg. 2009;109(1):285; author reply 85-6.
14. Hoo ZH, Candlish J, Teare D. What is an ROC curve? Emerg Med J. 2017;34(6):357-359.
15. Štrumbelj E. KI. Explaining prediction models and individual predictions with feature contributions. J Knowledge and Information Systems. 2013;41:647-665.
16. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From Local Explanations to Global Understanding with Explainable AI for Trees. Nat Mach Intell. 2020;2(1):56-67.
17. Szumilas M. Explaining odds ratios. J Can Acad Child Adolesc Psychiatry. 2010;19(3):227-229.
18. Ramspek CL, Voskamp PW, van Ittersum FJ, Krediet RT, Dekker FW, van Diepen M. Prediction models for the mortality risk in chronic dialysis patients: a systematic review and independent external validation study. Clin Epidemiol. 2017;9:451-464.
19. Cohen LM, Ruthazer R, Moss AH, Germain MJ. Predicting six-month mortality for patients who are on maintenance hemodialysis. Clin J Am Soc Nephrol. 2010;5(1):72-79.
20. Doi T, Yamamoto S, Morinaga T, Sada KE, Kurita N, Onishi Y. Risk Score to Predict 1-Year Mortality after Haemodialysis Initiation in Patients with Stage 5 Chronic Kidney Disease under Predialysis Nephrology Care. PLoS One. 2015;10(6):e0129180.

21. Msaad R, Essadik R, Mohtadi K, Meftah H, Lebrazi H, Taki H, et al. Predictors of mortality in hemodialysis patients. Pan Afr Med J. 2019;33:61.
22. Palmer SC, Navaneethan SD, Craig JC, Johnson DW, Tonelli M, Garg AX, et al. Meta-analysis: erythropoiesis-stimulating agents in patients with chronic kidney disease. Ann Intern Med. 2010; 153(1):23-33.
23. Fishbane S, El-Shahawy MA, Pecoits-Filho R, Van BP, Houser MT, Frison L, et al. Roxadustat for Treating Anemia in Patients with CKD Not on Dialysis: Results from a Randomized Phase 3 Study. J Am Soc Nephrol. 2021;32(3):737-755.
24. Agarwal R. Mechanisms and mediators of hypertension induced by erythropoietin and related molecules. Nephrol Dial Transplant. 2018;33(10):1690-1698.
25. Guedes M, Guetter CR, Erbano LHO, Palone AG, Zee J, Robinson BM, et al. Physical health- related quality of life at higher achieved hemoglobin levels among chronic kidney disease patients: a systematic review and meta-analysis. BMC Nephrol. 2020;21(1):259.
26. Guedes M, Muenz DG, Zee J, Bieber B, Stengel B, Massy ZA, et al. Serum Biomarkers of Iron Stores Are Associated with Increased Risk of All-Cause Mortality and Cardiovascular Events in Nondialysis CKD Patients, with or without Anemia. J Am Soc Nephrol. 2021;32(8):2020-2030.
27. Kotanko P, Thijssen S, Kitzler T, Wystrychowski G, Sarkar SR, Zhu F, et al. Size matters: body composition and outcomes in maintenance hemodialysis patients. Blood Purif. 2007;25(1):27-30.
28. Toida T, Sato Y, Ogata S, Wada A, Masakane I, Fujimoto S. Synergic Impact of Body Mass Index, Diabetes, and Age on Long-Term Mortality in Japanese Incident Hemodialysis Patients: A Cohort Study on a Large National Dialysis Registry. J Ren Nutr. 2020;30(4):333-340.
29. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. Am J Epidemiol. 2013;177(4):292-298.
30. Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. CHANCE. 2019 2019/01/02;32(1):42-49.
31. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004; 15(5):615-625.
32. Harrell F. Road Map for Choosing Between Statistical Modeling and Machine Learning. Staistical Thinking. 2021;[Accessed Feb 2022].
33. Lee H, Yun D, Yoo J, Yoo K, Kim YC, Kim DK, et al. Deep Learning Model for Real-Time Prediction of Intradialytic Hypotension. Clin J Am Soc Nephrol. 2021;16(3):396-406.
34. Dernoncourt F, Nemati S, Kassis EB, Ghassemi MM. Hyperparameter Selection. Secondary Analysis of Electronic Health Records. Cham (CH): 2016: 419-427.
35. Siga MM, Ducher M, Florens N, Roth H, Mahloul N, Fouque D, et al. Prediction of all-cause mortality in haemodialysis patients using a Bayesian network. Nephrol Dial Transplant. 2020;35(8):1420-1425.
36. Jung HY, Kim SH, Jang HM, Lee S, Kim YS, Kang SW, et al. Individualized prediction of mortality using multiple inflammatory markers in patients on dialysis. PLoS One. 2018;13(3):e0193511.
37. Zhu J, Tang C, Ouyang H, Shen H, You T, Hu J. Prediction of All-Cause Mortality Using an Echocardiography-Based Risk Score in Hemodialysis Patients. Cardiorenal Med. 2021;11(1):33-43.
38. Wang Y, Zhu Y, Lou G, Zhang P, Chen J, Li J. A maintenance hemodialysis mortality prediction model based on anomaly detection using longitudinal hemodialysis data. J Biomed Inform. 2021;123:103930.
39. Tapak L, Sheikh V, Jenabi E, Khazaei S. Predictors of mortality among hemodialysis patients in Hamadan province using random survival forests. J Prev Med Hyg. 2020;61(3):E482-E88.
40. Thijssen S, Usvyat L, Kotanko P. Prediction of mortality in the first two years of hemodialysis: results from a validation study. Blood Purif. 2012;33(1-3):165-170.
41. Hemke AC, Heemskerk MB, van Diepen M, Dekker FW, Hoitsma AJ. Improved Mortality Prediction in Dialysis Patients Using Specific Clinical and Laboratory Data. Am J Nephrol. 2015;42(2):158-167.
42/ Floege J, Gillespie IA, Kronenberg F, Anker SD, Gioni I, Richards S, et al. Development and validation of a predictive mortality risk score from a European hemodialysis cohort. Kidney Int. 2015;87(5):996-1008.
43. Holme I, Fellstrom BC, Jardin AG, Schmieder RE, Zannad F, Holdaas H. Prognostic model for total mortality in patients with haemodialysis from the Assessments of Survival and Cardiovascular Events (AURORA) study. J Intern Med. 2012;271(5):463-471.

Chapter 9

Discussion



Discussion

This thesis demonstrated several applications of how Artificial Intelligence (AI) and digital technologies can be used in clinical care of End Stage Kidney Disease (ESKD) patients. Advances in computing, mathematics and statistics have resulted in the evolution of AI methods. With advancing technology and Electronic Medical Records (EMR) for ESKD patients, large volume of clinical and lab data is being collected. AI techniques can allow us to derive small signals from large volumes of data.

We introduce some high-level concepts related to AI in **chapter 2**, we also look at several examples of how AI has been implemented in caring for patients with kidney disease. In **chapter 3**, we demonstrated that a Machine Leaning (ML) driven decision support model used to direct care in outpatient dialysis clinics participating in a pilot program Dialysis Hospital Reduction Program (DHRP) was associated with lower annualized all-cause hospitalization rates compared to matched control clinics. ML model used historical patient data to identify patients who were at high risk of at least 6 hospitalizations in the following 12 months. There are no other examples of AI/ML based hospitalization risk models being used to direct care in quality improvement efforts in dialysis care. In general, the clinical application of AI in dialysis is scarce with only one report identified in a recent bibliometric study on the global evolution of AI in healthcare.[1,2] The all-cause hospitalization risk models implemented in the DHRP pilot assisted care teams in identifying the subset of individual patients in their clinic at the highest risk of hospitalization and thus prioritizing personalized care for these patients to improve outcomes. Other areas of healthcare frequently demonstrate the use of risk-based prediction models to identify patients needing intervention.[3,4] A major challenge however is to integrate the risk score in regular clinical workflow of clinical staff performing interventions on the output of the model.

In **chapter 4** and **chapter 5** we presented examples of advancing digital technologies to make ML based decision support tools available on demand at the point of care. In **chapter 4** we propose a new proof of concept architecture to run ML models in real time. The findings from our proof-of-concept analysis suggest the potential for real-time reporting and prediction of treatment blood volume profiles that are associated with an increased risk of intra-dialytic symptoms and would subsequently be amenable to intervention. Furthermore, the architectural framework demonstrated in this paper can be used for making real-time predictions of other events during dialysis treatments. Making real-time predictions can provide proactive decision support to clinicians and nurses at the point of care during dialysis treatment. A practical implication for the present would be that, if nurses and clinicians are alerted to the risk of a drop in the blood volume 15 minutes prior to the RBV decreasing at rate of at least -6.5% per hour during a dialysis treatment, they would have sufficient time to intervene and adjust the

ultrafiltration rate in order to prevent that patient from entering the risk zone for intradialytic symptoms like intra dialytic hypotension (IDH).[5] Prior studies have been attempted to monitor hematocrit and reduce intradialytic symptoms, however, they were not used in standard practice because of the difficulty in interpreting the Optical Sensing Device (OSD) outputs updated every 10 seconds.[6,7] The ML model presented in this analysis enhances the findings and delivers them in a comprehendible way. Near real time analytics is crucial for care in emergency rooms and intensive care units (ICU) where decisions need to be made faster. Other attempts to develop real time analytics architecture have been made in the field of ICU care using a different analytics platform.[8]

In **chapter 5** we studied the clinical outcome of application of the digital technology on a cohort of patients treated with peritoneal dialysis (PD). In a large population of patients on PD who registered online for the Patient Hub Remote Treatment Monitoring (RTM) application, we found that higher RTM use was associated with lower hospitalization and sustained technique failure rates. These results further substantiate prior findings suggesting that employing RTM systems in PD patients may reduce hospitalization rates[9,10], and reveal frequent RTM use may also have the potential to increase sustained use of PD as a modality. RTM has also shown that it can reduce healthcare utilization and associated costs in a simulation study for PD patients.[11] Given PD patients are typically younger, it would be expected that there might be a larger proportion of PD patients with access to the internet. Under this assumption, RTM might have the potential to be used in a larger proportion of the PD population, and as smartphone and computer technology advances and becomes more universally affordable, it could become a conducive option for treatment monitoring for most patients. There are several examples in other fields of healthcare where RTM has shown to improve patient compliance and outcomes.[12,13]

In **chapter 6 and chapter 7**, we studied the application of ML concepts for ESKD patients during the Coronavirus Disease (COVID) pandemic. We first observed unique temporal trends in various clinical and laboratory parameters among Hemodialysis (HD) patients who tested positive versus negative for SARS-CoV-2 infection. Ultimately, these trends helped to define the physiological disturbances that characterize the onset and course of COVID-19 in HD patients. The disturbances in various parameters commonly started several days before presentation and were more pronounced in patients who tested positive for COVID-19. We identified statistically distinct daily changes in vital signs, inflammatory, and nutritional markers across the weeks before presentation among patients who tested COVID-19 positive versus negative, as well as unique differences between groups on the date of suspicion/testing. We used these findings to successfully develop a ML prediction model using retrospective data that appears to have suitable performance in identifying HD patients at risk of having an

undetected SARS-CoV-2 infection that is identified in the following ≥3 days. The top predictors of a patient having a SARS-CoV-2 infection were the change in interdialytic weight gain from the previous month, mean pre-HD body temperature in the prior week, and change in post-HD pulse from the previous month. Our analysis augmented previous findings by showing quantitative trends in key parameters during the period before and after presentation with COVID-19. Digital technologies have played a critical role in diagnosing, monitoring, and managing COVID-19 remotely.[14,15] Individual predictions can be further used to identify the risk level for dialysis clinics based on proportion of patients classified with an undetected SARS-CoV-2 infection. We proposed a conceptual workflow for the application of the ML model predictions to assist with directing care to individual patients and assisting with directing resource allocations to clinics. It is important to note that this analysis was done in the early stages of the pandemic before the various SARS-CoV-2 variants existed and vaccinations were available. However, the framework for intervention presented in the paper can be expanded to current situation.

In **chapter 8**, we developed a mortality prediction model based on a large internationally representative cohort of hemodialysis patients from dialysis providers throughout 37 countries in 5 continents.[16] Importantly, the prediction model developed here used a set of input parameters that is frequently collected in the standard care of HD patients by most dialysis providers worldwide. We describe the development of a ML model and a traditional model that was suitable for classification of a prevalent HD patient's three-year risk of death. We assessed the risk factors from both techniques and compared the output of the advanced ML model to output from the traditional model that is generally well understood in the clinical community. Traditional statistical and advanced ML techniques could be used in a complementary way. It is a widely accepted practice to initially build a baseline model using traditional statistical techniques and subsequently use that model to benchmark against advanced ML models to compare performance.

XGBoost is a widely adopted advanced modeling technique and is also used in many ML models presented in this thesis.[17] It is based on building several decision trees with a random sample training data and a series of thresholds that split the parameters to maximize information gain. New decision trees are created iteratively to minimize prior prediction errors. It can handle collinearity between input parameters and handle missing data for such parameters by recognizing the missingness when determining the splits.

Figure 9.1 shows a vision of how clinical care in chronic kidney disease (CKD) and ESKD can transition from a reactive care to a proactive care. Traditionally, patients seek care from nephrologist when they see a drop in their kidney function or when they

unexpectedly have medical conditions that require dialysis. The clinicians make several decisions for the patients along the care continuum on a need basis. The patients who transition to dialysis perform dialysis mostly three times a week and use resources such as a hospital setting when needed.

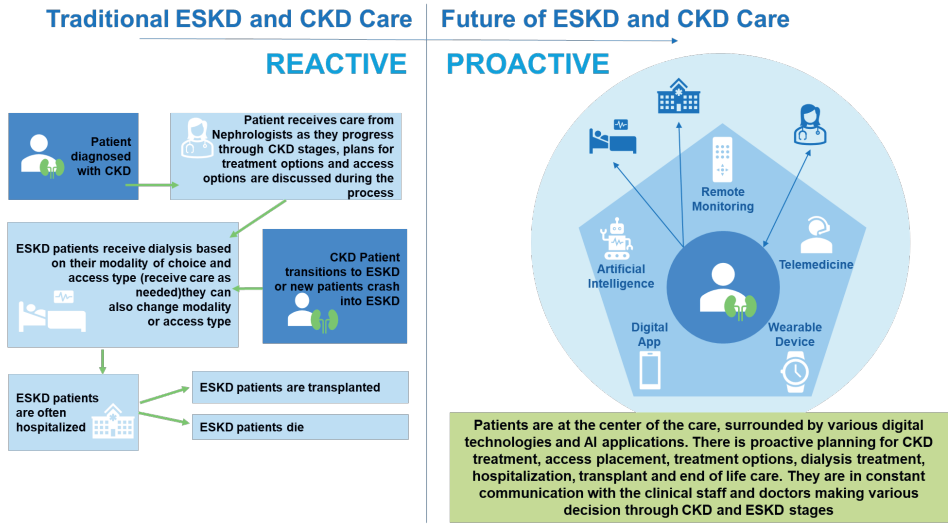


Figure 9.1 Future of ESKD and CKD Care.

In a proactive care setting the patient is at the center of the care and has full control of his or her health. The patient is surrounded by digital applications, remote monitoring technologies and artificial intelligence applications helping the clinicians make proactive decision for the patient and stay in constant contact with the patient through their CKD and ESKD journey.

AI and digital applications in other fields of healthcare have made some significant advancements in measuring outcomes of clinical applications in the last couple of years. For example, a review paper looked at 21 studies using AI based systems and robotic applications in managing advanced stages of Parkinson's disease. The studies used AI and digital technologies for autonomous management of pharmacologic therapy, home-based telemedicine systems and robot-assisted gait training systems. It was shown that there was significant evidence demonstrating that current AI-based technologies are feasible for improving quality of care and reducing cost of patients with advanced stages of Parkinson's disease.[18]

Similarly, for hospitalizations related to Acute Myocardial Infarction, a non-randomized controlled trial was conducted at 4 US hospitals between 2015 to 2019. Digital health interventions such as medication reminders and activity tracking after hospitalization was integrated into smartphones, smart watch, and blood pressure monitors. It was shown that such digital health intervention may be associated lower risk of all-cause unplanned 30-day readmissions.[19]

The COVID-19 pandemic has shown us that remote technologies and digital applications are of high importance in caring for patients especially the elderly and vulnerable population. One important example of telemonitoring in cardiology include the use of implanted devices such as pacemakers, defibrillators, and wearable sensors for arrythmia detection and to remotely monitor patients with heart failure.[20,21]

In another randomized controlled trial, patients with inflammatory bowel disease were randomly assigned to care via a telemedicine system that monitors and registers disease activity. It was observed that, telemedicine reduced outpatient visits and hospital admissions compared with standard care.[22]

These examples support the fact that AI and digital technologies have the potential to change care for patients in the field of nephrology.

Limitations and Strengths

Limitations

In all applications presented, it is important to note that the clinicians should use their medical judgement to make decisions with assistance from ML tools. It is also important for clinical teams to understand the outputs generated and consider the limitations of the models in the design of directed assessments and interventions. ML based risk classification and other clinical decision support tools are never perfect, there will be instances when the ML model will predict incorrectly, so teams developing interventions using ML models need to be aware of this limitation. Furthermore, incorrect predictions should have a feedback mechanism to ML models to improve their accuracy. The effectiveness of the prediction models depends heavily on the ability to use insights to make clinical interventions. Also, in cases where the model has predicted incorrectly, a precedent of accountability needs to be established. ML models should be transparent and traceable.

True performance of ML models can only be demonstrated after conducting randomized clinical trials (RCT), whereas none of the applications presented in this

thesis were studied with randomized clinical trials. Creative design of pilots that include cluster randomization during rollout may be able to strengthen results in a randomized way in practice of medicine as opposed to hypothesis driven tests performed in formal RCTs.[23]

AI and digital solutions must also follow ethical guidelines and consider at the time of conception whether software programs are medical devices that require formal regulatory pathways and trials.[24,25] Ethical and legal governing bodies should be setup to evaluate and monitor the impact of the AI and digital applications on the patient and the society at large.[26] We should consider what aspects of care are important for the patient and how it impacts them psycho-socially when implementing AI and digital solutions.[27] Data security is also a major concern with data collected via digital applications. There should be checks in place to ensure health data is not sold to third parties.

It is critical that the predictive models use data collected routinely in standard of care or it will likely produce models that are biased 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.

We should be cautious about generalizing the outcomes for ESKD patients who use digital technologies and applications at home as it may not be true representative of the entire population as they may be healthier and hence technology savvy.

When comparing traditional methods with advanced ML techniques, it is essential to consider interactions between the input parameters. It is possible that the performance of the traditional model may improve if we considered these interactions, we may not be comparing the best traditional method with advanced ML technique.

When models are implemented in clinical practice it is important to evaluate if the model is making a difference on the clinical outcome of the patients. Most ML models are built on retrospective data and such models must be tested prospectively and re-trained on newer data periodically.

Lastly, knowledge about applications of AI and digital applications in nephrology is limited, we must continue to bridge the gap between medical and analytical insights through educational programs for clinicians.

Strengths

Most of the AI and digital applications presented here were developed and tested using large cohorts of ESKD patients that are mostly representative of international dialysis populations. All applications use advanced models to gain insights about ESKD patients.

ML models that have the potential to provide an intelligent and timely triage of additional resources to proactively assist in directing medical decision making and personalized care for ESKD patients and thus improve quality of life/clinical outcomes of patients. With access to large amounts of data collected in EMR and from the dialysis machines, ML models can harness this vast amount of data to personalizing dialysis treatments for ESKD patients.

Cloud computing resources and digital applications provide seamless tools to build, analyze, and integrate real-time predictive models without investing in many hardware and software resources on premise. This allows for a secure and cost-effective way of building predictive models when resources are limited. These applications can also be scaled on-demand, where support can be expanded from tens to hundreds of clinics seamlessly. However, such cloud resources should be compliant with local regulations like GDPR and HIPAA. Data stored on such cloud platforms should be anonymized and independent security/privacy risk assessments may be warranted to confirm anonymization strategies.

The flexible ability of the RTM like digital applications allow patients to enter records when it works the best for them during/around their treatment. Connected health technologies also provide the ability for clinical staff to review patients' entries around their daily workflows and intervene if there are any exceptions. Digital applications allow patients to take control of their care, becoming an active participant and creating synergies with their clinical care teams.

In summary, applications of AI and digital technologies in caring for dialysis patients have several advantages and disadvantages. Following **Table 9.1** provides a summary of the advantages and disadvantages.

Conclusion

Advancements in AI and digital applications have the potential to change the way care is provided for patients in nephrology. It should be viewed as a decision support tool to extend human insight and not something that will replace human medical decision making.

Table 9.1 Advantages and Disadvantages of using Digital Technologies and AI Enabled Applications.

Advantages	Disadvantages
Can be used to develop proactive care for patients	Not all patients may have access to or understand how to use and interpret advanced digital technology and applications
Patient is at the center of the treatment and can help empower patients to make decisions	Not all patients may be open to embracing change in how they are cared
Can help with better adherence and clinical to patient communication	Can be perceived as a threat to the health care professionals
Digital can AI applications can help with personalized treatment plans	If AI applications are not trained optimally, it could result in some bias towards certain population
Can help reduce cost of care in the long run	Requires lots of historical data and the initial cost to setup may be high
Can help with real time analysis of the data captured	Some diagnosis may be wrong
Advanced digital technologies can integrate analytics at the point of care in real time	With advance in technology there are security concerns with transmitting and storing data digitally

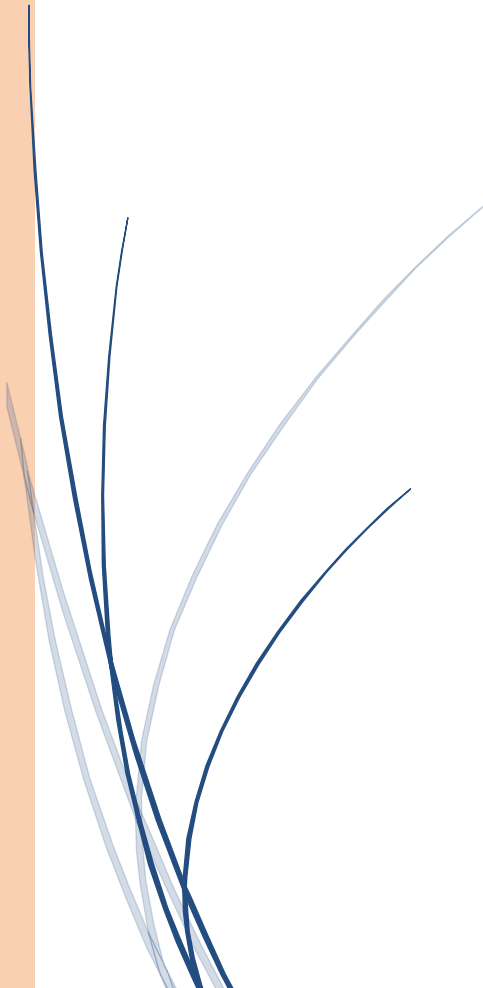
References

1. Tran BX, Vu GT, Ha GH, et al. Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. J Clin Med 2019;8:360.
2. 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.
3. Durojaiye OC, Kritsotakis EI, Johnston P, et al. Developing a risk prediction model for 30-day unplanned hospitalization in patients receiving outpatient parenteral antimicrobial therapy. Clin Microbiol Infect 2019;25:905 e901-905 e907.
4. Pugh K, Granger D, Lusk J, et al. Targeted Clinical Interventions for Reducing Pediatric Readmissions. Hosp Pediatr 2021;11:1151-1163.
5. America FMCN. Crit-Line IV Monitor. 2020; (Accessed May 2020): https://fmcna.com/products/fluid-management/crit-line-iv-monitor/.
6. Steuer RR, Leypoldt JK, Cheung AK, et al. Reducing symptoms during hemodialysis by continuously monitoring the hematocrit. Am J Kidney Dis 1996;27:525-532.
7. Al-Hilali N, Al-Humoud HM, Ninan VT, et al. Profiled hemodialysis reduces intradialytic symptoms. Transplant Proc 2004;36:1827-1828.
8. van de Sande D, Van Genderen ME, Huiskens J, et al. Generating insights in uncharted territories: real-time learning from data in critically ill patients-an implementer report. BMJ Health Care Inform 2021; 28: e100447.
9. Sanabria M, Buitrago G, Lindholm B, et al. Remote Patient Monitoring Program in Automated Peritoneal Dialysis: Impact on Hospitalizations. Perit Dial Int 2019;39:472-478.
10. Lew SQ, Sikka N, Thompson C, et al. Impact of remote biometric monitoring on cost and hospitalization outcomes in peritoneal dialysis. J Telemed Telecare 2018:1357633X18784417.
11. Makhija D, Alscher MD, Becker S, et al. Remote Monitoring of Automated Peritoneal Dialysis Patients: Assessing Clinical and Economic Value. Telemed J E Health 2018;24:315-323.
12. Natarajan MK, Sheth TN, Wijeysundera HC, et al. Remote ECG monitoring to reduce complications following transcatheter aortic valve implantations: the Redirect TAVI study. Europace 2022;24: 1475-1483.
13. Mir O, Ferrua M, Fourcade A, et al. Digital remote monitoring plus usual care versus usual care in patients treated with oral anticancer agents: the randomized phase 3 CAPRI trial. Nat Med 2022;28: 1224-1231.
14. Lukas H, Xu C, Yu Y, et al. Emerging Telemedicine Tools for Remote COVID-19 Diagnosis, Monitoring, and Management. ACS Nano 2020;14:16180-16193.
15. Amft O, Lopera L, Lukowicz P, et al. Wearables to Fight COVID-19: From Symptom Tracking to Contact Tracing. IEEE Pervasive Comput 2020;19:53-60.
16. von Gersdorff GD, Usvyat L, Marcelli D, et al. Monitoring dialysis outcomes across the world--the MONDO Global Database Consortium. Blood Purif 2013;36:165-172.
17. Chen T GC. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining San Francisco, California, USA, Association for Computing Machinery 2016:785–794.
18. Perju-Dumbrava L, Barsan M, Leucuta DC, et al. Artificial intelligence applications and robotic systems in Parkinson's disease (Review). Exp Ther Med 2022;23:153.
19. Marvel FA, Spaulding EM, Lee MA, et al. Digital Health Intervention in Acute Myocardial Infarction. Circ Cardiovasc Qual Outcomes 2021;14:e007741.
20. Kedzierski K, Radziejewska J, Slawuta A, et al. Telemedicine in Cardiology: Modern Technologies to Improve Cardiovascular Patients' Outcomes-A Narrative Review. Medicina (Kaunas) 2022;58:210.
21. Leclercq C, Witt H, Hindricks G, et al. Wearables, telemedicine, and artificial intelligence in arrhythmias and heart failure: Proceedings of the European Society of Cardiology Cardiovascular Round Table. Europace 2022;24:1372-1383.
22. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet 2017;390:959-968.

23. Lorenz E, Kopke S, Pfaff H, et al. Cluster-Randomized Studies. Dtsch Arztebl Int 2018;115:163-168.
24. Team RaA. Ethics guidelines for trustworthy AI European Commission 2019; (Accessed December 2019): https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.
25. Administration UFaD: Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD). In, edited by Services DoHaH, 2019
26. Van Biesen W, Decruyenaere J, Sideri K, et al. Remote digital monitoring of medication intake: methodological, medical, ethical and legal reflections. Acta Clin Belg 2021;76:209-216.
27. van Biesen W, Van Der Straeten C, Sterckx S, et al. The concept of justifiable healthcare and how big data can help us to achieve it. BMC Med Inform Decis Mak 2021;21:87.

Addendum

Summary



Summary

Healthcare around the world is under tremendous pressure, with rising costs of care and limited resources. This is also true in providing care for End Stage Kidney Disease (ESKD) patients. Coronavirus Disease (COVID) pandemic has added further burden to these vulnerable patients and clinical care teams caring for these patients. It is therefore particularly timely to look for ways to cut the cost and improve care with the help of technology driven solutions. With rise in healthcare data collected in Electronic Medical Records (EMR), we can use advancing technologies to harness this data. Traditionally computer-based algorithms in health care include a set of rules encoding expert knowledge on medical decisions. These rules are subsequently applied to draw conclusions about specific clinical scenarios. Artificial Intelligence (AI) algorithms, however, strive to learn from the data without concrete rules. **In Chapter 2** We summarized high level concepts related to AI and presented several areas where AI has been applied in the field of Nephrology.

While there are many articles in literature that present how AI can be applied in small cohort of patients, there are very few articles where outcomes are measured retrospectively in a large clinical application. Dialysis Hospitalization Risk Pilot (DHRP) application presented in **Chapter 3** is one such example. In this retrospective analysis of the clinical application, we found ML directed assessment and personalized interventions in the DHRP, were associated with lower all-cause hospitalization rates compared to control clinics. The DHRP findings detail how such applications can reduce the cost of care and should be considered by payors, providers, and clinicians.

In the future, AI application will be more impactful if such models are run in real time as the patient is receiving dialysis. This proof-of-concept analysis demonstrated in **Chapter 4** shows the potential of the creation and deployment of a real-time predictive model based on patient and dialysis treatment data.

The success of AI in the future will also depend on the data beyond the data collected in traditional EMR. Data collected from wearable devices, genomic data and other digital applications will be very important in deriving additional insights. Our findings in **Chapter 5** provide an example of one such digital application that can be used by Peritoneal Dialysis (PD) patients at home. We conclude frequent use of a Remote Treatment Monitoring (RTM) application associates with less hospital admissions, shorter hospital length of stay, and lower rates of sustained technique failure.

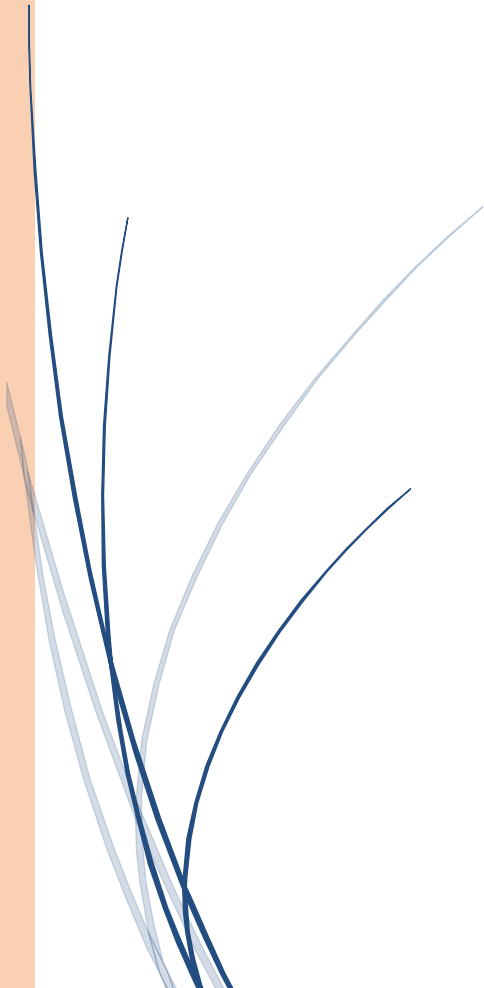
Similar to identifying patients at high risk of hospitalization, AI can also help with identification of high risk of other events such as SARS-CoV-2 infection and mortality. In **Chapter 6 and 7** we showed how such models can be developed using the data

collected in EMR and utilized during the Coronavirus Disease Pandemic. In **Chapter 8** we demonstrated the successful use of AI in predicting mortality on a large cohort of international HD patients.

AI has the potential to assist clinical teams in improving the quality of life of ESKD patients and reduce the overall cost of care. However, factors such as clinical effectiveness of an AI solution, and accountabilities in case of error, need to be carefully considered before implementing such solutions. Furthermore, policies and regulations need to evolve to guide the development of AI at an acceptable scale. Most importantly, the value of an AI solution to effectively deliver better outcomes needs to be demonstrated to patients, physicians, and providers to foster their trust.

Addendum

Valorization



Valorization

We showed several examples of how data collected in Electronic Medical Record (EMR) can be utilized to build applications for patients with kidney disease. Artificial Intelligence (AI) techniques can allow for large datasets to be leveraged with minimal efforts. We have shown several examples in this thesis of how such applications can be developed and implemented at the point of care for End Stage Kidney Disease (ESKD) patients. AI and digital applications can be developed and integrated within EMR caring for ESKD as well as Chronic Kidney Disease (CKD) patients.

From a patient's perspective, AI based applications can help provide proactive care by identifying problems before they occur throughout their journey as a CKD or ESKD patient. Digital applications keep the patients engaged in their personal care. Further, digital applications and AI models can be developed to incorporate genomic. Proactive and personalized care can improve the overall quality of life for ESKD and CKD patients. As shown in **chapters 3** and **5**, AI models can reduce health care resource utilization and remote monitoring applications can prolong the length of stay for peritoneal dialysis patients. Additionally, timing of when the predictions are made and how they are integrated in the clinical workflow is of utmost importance, such frameworks have been discussed in chapters 4 and 7.

From a dialysis provider's perspective, offering proactive and personalized care can help lower the overall cost of care. AI based risk models can be used to negotiate contracts with governments, private insurance companies and providers by identifying high risk ESKD patients and designing cost saving clinical interventions around such high-risk patients. In the United States, ESKD Seamless Care Organization (ESCO) was such an attempt to allow nephrologists, dialysis providers and other partners to test a new care delivery paradigm aimed at improving clinical outcomes and patient experience under a shared savings structure.

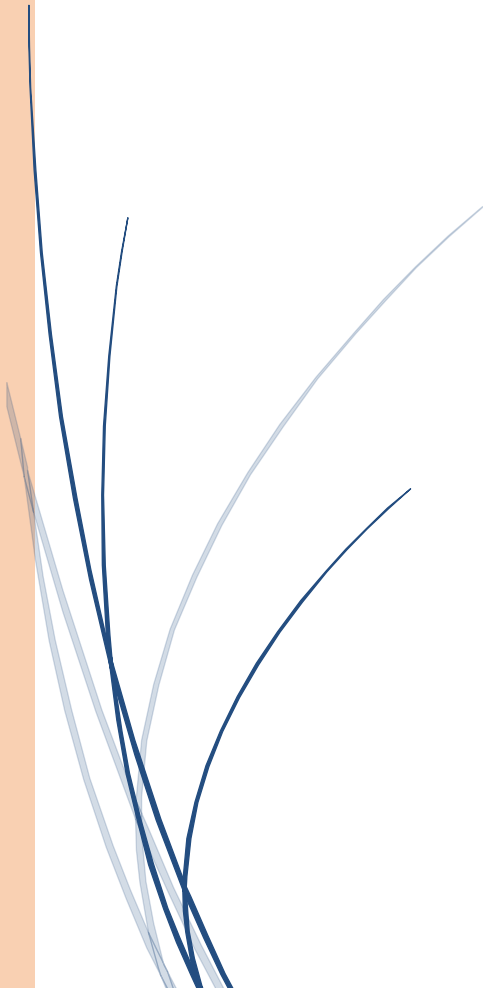
Lastly, from a clinician's perspective, the AI and digital applications can be viewed as a decision support tool that will help them guide their clinical decisions. It is not meant to replace the clinical judgement of a medical professional; however, it is meant to aid them in the process of providing care. AI applications can identify patterns in vast amounts of data and derive insights that are not often comprehendible to a human brain. If successfully adopted in medical practice, this can reduce the time it takes to care for patients. Thus, such applications have the potential to help the clinicians and reduce their overall burn out and turn over.

In summary, AI and digital applications harness vast amounts of clinical data that is currently being collected across EMRs. Mining clinical data using advanced method and

AI application can help improve communication, make smarter decisions, and improve the overall efficiency in providing care for patients suffering from kidney disease.

Addendum

Acknowledgements



Before I acknowledge my promoters, co-promoters and all those who have played such a critical role through this journey, I would like to pay a tribute and dedicate all this work to my late father Shrikant Naik. None of this would have ever been possible if he had not believed in me as a child. Being a teacher by profession, his sole aim was to impart knowledge and more so seeing girls prosper in life. Despite coming from humble beginnings with limited resources, he made sure I had all the resources I needed to gain knowledge. He would have been the proudest to see me defend this thesis.

I would also like to thank my mentor Dr. Len Usvyat and my husband Dr. Ayan Chaudhuri both have played a significant role in helping me decide if I wanted to embark on this journey with a full-time job and as a mother of two young children. They both encouraged and believed in me that I could successfully step into this journey and complete my thesis. I would also like to especially thank Dr. Usvyat for giving me this opportunity to advance my career in his team and mentoring me through every step of this process.

I would like to thank my promoter Dr. Jeroen Kooman who gave me this opportunity to work with Maastricht University. He helped me in thinking through all the studies outlined in the thesis and met with me several times in person and virtually, despite his busy schedule.

I would like to thank my co-promoter Dr. John Larkin for meeting with me every other week and helping me review all my manuscripts and final thesis in detail.

I would like to thank past and present members of the Clinical Advanced Analytics team especially Hao Han, Peter Yu, Dr. Andrew Long, Joanna Willetts, Jane Brzozowski, Dr. Yue Jiao, Dr. Caitlin Monaghan, Rachel Lasky and Suman Lama. They all have played a significant role in helping me along this journey.

I would like to thank Dr. Peter Kotanko and Dr. Hanjie Zhang from the Renal Research Institute, New York, who have played a key role in providing numerous inputs to my thesis.

I would like to thank Peter Waguespack, Zuwen Kuang, Kanti Singh, Srikanth Bellamkonda and Brain Shulman from the Digital Technology and Innovation group for unlimited guidance on the technical systems.

I would like to thank Dr. Yuedong Wang from University of Santa Barbara, California, for his statistical guidance on all projects listed in this thesis.

I would like to thank Diane Rondeau for helping me coordinate so many meetings with people in different time zones.

I would like to thank Glori Edewaard for her amazing graphic skills. She has helped me with several graphics listed in this thesis.

I would like to thank Stephanie Johnstone Steinberg, Dr. Dugan Maddux, Dr. Shashi Kalaskar, Dr. Frank M. van der Sande, Carlos Muchiutti, Jessica Ryter, Dr. Marta Reviriego-Mendoza, Dr. Jeffrey Hymes, Dr. Robert Kossmann, Dr. Dinesh Chatoth, Mike Black, David Sweet, Allison Vinson, Brad Bucci, Dr. Frank M. van der Sande, Murilo Guedes, Dr. Anke Winter, Dr. Luca Neri, Dr. Sangho Lee and Kathleen Belmonte for their valuable input on my thesis.

I would like to thank Dr Franklin W Maddux for his input in all the papers and giving me this wonderful opportunity to grow in the Fresenius Global Medical Office.

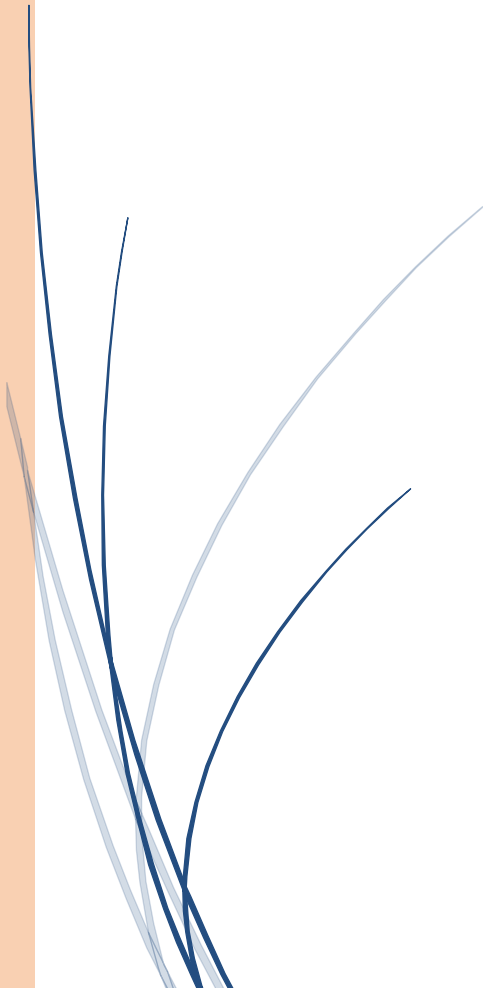
I would also like to thank my brother Gajanan Naik without whom I would have never pursued a degree or a career in the United States of America.

I would like to thank Leah Samanta and Georgia Kakoulaki for agreeing to be my paranymphs.

Last but not the least, I would like to acknowledge and thank my mother Mrs. Sheela Naik and my children Ansh Chaudhuri and Anya Chaudhuri for their unconditional love and support in this journey.

Addendum

Curriculum Vitae



Sheetal Chaudhuri has over 18 years of experience in designing data and analytical solutions in the healthcare industry. She received a bachelor's degree in computer engineering from Goa University, India in 2001. She moved to United States in 2002 to pursue her master's degree in computer and information Science from University of Massachusetts at Dartmouth. Sheetal started her career as a business intelligence consultant serving prestigious clients of SAP Inc. She worked on projects for the National Defense Academy and State Farm Insurance. Her introduction to healthcare analytics was through her role as a programmer for Risk Management Foundation of the Harvard Medical Institution, where she was responsible for developing analytical solutions related to malpractice claims associated with doctors in the Harvard network. Concurrently, Sheetal also pursued a second master's degree in liberal arts, general management from Harvard University. Sheetal has been working for Fresenius Medical Care for the past 10+ years in various analytical roles. In her current role as the Global Data Analytics and Engineering Lead, Sheetal is responsible for providing data and clinical analytics support for various stakeholders within Fresenius Medical Care. Her team is also responsible for project initiation, data extraction, operationalizing machine learning models and analyzing outcomes for advanced analytics initiatives within the company.