Development and testing of an artificial intelligence tool for predicting end-stage kidney disease in patients with immunoglobulin A nephropathy

Francesco Paolo Schena[1,2], Vito Walter Anelli[3], Joseph Trotta[3], Tommaso Di Noia[3], Carlo Manno[1], Giovanni Tripepi[4], Graziella D'Arrigo[4], Nicholas C. Chesnaye[5], Maria Luisa Russo[6], Maria Stangou[7], Aikaterini Papagianni[7], Carmine Zoccali[4], Vladimir Tesar[8], Rosanna Coppo[6], and members of the VALIGA study[9]

[1]Department of Emergency and Organ Transplant, University of Bari, Bari, Italy; [2]Research Laboratory, Fondazione Schena, Valenzano, Bari, Italy; [3]Department of Electrical and Information Engineering, Polytechnic of Bari, Bari, Italy; [4]CNR Institute of Clinical Physiology, Reggio Calabria, Italy; [5]Department of Medical Informatics, Public Health Research Institute, University of Amsterdam, Amsterdam, The Netherlands; [6]Research Laboratory, Fondazione Ricerca Molinette, Torino, Italy; [7]Department of Nephrology, Hippokration General Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece; and [8]Department of Nephrology, 1st Faculty of Medicine and General University Hospital, Charles University, Prague, Czech Republic

We have developed an artificial neural network prediction model for end-stage kidney disease (ESKD) in patients with primary immunoglobulin A nephropathy (IgAN) using a retrospective cohort of 948 patients with IgAN. Our tool is based on a two-step procedure of a classifier model that predicts ESKD, and a regression model that predicts development of ESKD over time. The classifier model showed a performance value of 0.82 (area under the receiver operating characteristic curve) in patients with a follow-up of five years, which improved to 0.89 at the ten-year follow-up. Both models had a higher recall rate, which indicated the practicality of the tool. The regression model showed a mean absolute error of 1.78 years and a root mean square error of 2.15 years. Testing in an independent cohort of 167patients with IgAN found successful results for 91% of the patients. Comparison of our system with other mathematical models showed the highest discriminant Harrell C index at five- and ten-years follow-up (81% and 86%, respectively), paralleling the lowest Akaike information criterion values (355.01 and 269.56, respectively). Moreover, our system was the best calibrated model indicating that the predicted and observed outcome probabilities did not significantly differ. Finally, the dynamic discrimination indexes of our artificial neural network, expressed as the weighted average of time-dependent areas under the curve calculated at one and two years, were 0.80 and 0.79, respectively. Similar results were observed over a 25-year follow-up period. Thus, our tool identified individuals who were at a high risk of developing ESKD due to IgAN and predicted the time-to-event endpoint. Accurate prediction is an important step toward introduction of a therapeutic strategy for improving clinical outcomes.

Kidney International (2021) **99,** 1179–1188; https://doi.org/10.1016/j.kint.2020.07.046

KEYWORDS: artificial neural networks; clinical decision support system; end-stage kidney disease; IgA nephropathy; joint models; machine learning

Today's physicians create large amounts of health data using electronic health records. Artificial intelligence and machine learning combined with human intelligence can help physicians to provide better care for their patients and to improve individual health outcomes. Machine learning is a robust methodology used to extract information from large datasets based on experience. Several applications in health care,[1–4] especially those based on deep learning,[5,6] have been proposed.

Artificial neural networks (ANNs) represent one of the most notable advances in artificial intelligence; they are advanced models using a multivariate analysis. Inspired by the structure of the human brain, ANNs are composed of computational units called neurons that are interconnected and that estimate functions that depend on a large number of inputs. Each connection between the neurons is based on a weight representing the influence of one neuron on another neuron. Each ANN can learn from a training set of data by adapting its own weight, and then, if the learning phase is successful, the model can be used to predict the outputs. ANNs are relatively flexible because they tolerate missing data and noise in single variables well and translate multivariate nonlinear relationships into continuous functions without the need to understand the underlying relationships between the

Correspondence: *Francesco Paolo Schena, Department of Emergency and Organ Transplant, University of Bari, Bari, Puglia, Italy. E-mail: paolo.schena@uniba.it*

[9]Members of the VALIGA study are listed in the Appendix.

Received 29 January 2020; revised 23 July 2020; accepted 30 July 2020; published online 2 September 2020

variables.[7,8] ANNs as nonlinear, flexible, and general tools are capable of declining with any sort of arbitrary function.

IgA nephropathy (IgAN) is the most common biopsy-proven primary glomerulonephritis in the world[9] and is characterized by a progressive development of end-stage kidney disease (ESKD) in 30% to 40% of individuals after 20 years from kidney biopsy.[10] The heterogeneity of the clinical phenotype may influence the outcome of the disease. Many investigators have tried to search for the risk factors of ESKD (Supplementary Table S1), and others have developed various scoring systems[11–17] (Supplementary Table S2). More recently, a risk predictor tool based on a conventional Cox regression analysis has been used by Barbour *et al.*[18] to predict ESKD in a large multiethnic cohort of patients with IgAN. However, there are many constraints and difficulties in collecting data because some outpatients visit the hospital irregularly. Moreover, the nonlinearity of some predictors and the effects of drugs complicate the interpretation of the data and their use for clinical practice. Artificial intelligence research has developed ANNs and deep learning. By jointly exploiting the research results in the 2 fields, today we may take advantage of deep neural networks to refine risk prediction.

The aims of our study are (i) to identify risk factors and covariates affecting ESKD in patients with IgAN using the traditional regression model; (ii) to develop a neural network classifier, the clinical decision support system (CDSS), that may be able to predict ESKD at the time of the diagnosis based on clinical findings and histologic data of the renal biopsy; and (iii) to use joint model analysis of repeated ANN measures and time-to-event data to provide the adequate risk discrimination of our tool.

RESULTS

Characteristics of the study cohort

The demographic characteristics and laboratory and histologic findings of 948 patients with IgAN are shown in Table 1. Patients had a mean age of 40.6 ±14.0 years at the time of renal biopsy. The male-to-female ratio was 2.6:1. Hypertension was observed in 30.3% of the patients. The median values of laboratory findings were serum creatinine 1.20 (range, 0.96–1.70) mg/dl, estimated glomerular filtration rate (eGFR) 67.30 (range, 44.89–89.85) ml/min per 1.73 m^2 and daily proteinuria 1.30 (range, 0.60–2.50) g/d. Patients were categorized into chronic kidney disease stages according to the Kidney Disease Outcomes Quality Initiative classification: stage I (25.0%), stage II (32.7%), stage III (32.3 %), stage IV (8.5%), and stage V (1.5%). Mild proteinuria (<1 g/d) was present in 37.9% of the patients, moderate proteinuria (1–3 g/d) in 43.0%, and severe proteinuria (>3 g/d) in 19.1% of the patients with IgAN at the time of kidney biopsy. The duration of follow-up was 89.0 (50.0–134.0) months and 7952 years–man (mean follow-up per number of patients). The primary outcome was ESKD (eGFR <15 ml/min per 1.73 m^2), dialysis, or transplantation. After the kidney biopsy 60.9% of patients received renin-angiotensin system blockers (RASBs), whereas corticosteroids with or without cytotoxic agents were used in 27.2% of patients.

Table 1 | Baseline characteristics of patients with IgAN enrolled in the cohorts at the time of kidney biospy[a]

Characteristics	Study cohort	Test cohort
Patients (n)	948	167
Age at biopsy, yr	40.6 ± 14.0	40.1 ± 15.5
Sex (male/female)	685/263	120/47
Systolic blood pressure, mm Hg	134.1 ± 18.6	135.5 ± 17.8
Diastolic blood pressure, mm Hg	83.2 ± 10.9	81.2 ± 11.0
Mean arterial pressure, mm Hg	100.2 ± 12.4	99.3 ± 12.4
Hypertension, n (%)	287 (30.3)	58 (35.0)
Serum creatinine, mg/dl	1.20 (0.96–1.70)	1.12 (0.88–1.50)
eGFR MDRD, ml/min per 1.73 m^2	67.30 (44.89–89.85)	69.00 (48.80–91.00)
KDOQI stage, ml/min per 1.73 m^2, **n (%)**		
I >90	236 (25.0)	42 (25.1)
II 60–89	310 (32.7)	64 (38.3)
III 30–59	306 (32.3)	43 (25.7)
IV 15–29	81 (8.5)	13 (7.8)
V <15	15 (1.5)	5 (3.0)
Proteinuria, g/d, n (%)	1.30 (0.60–2.50)	1.00 (0.50–1.80)
Mild <1	359 (37.9)	80 (47.9)
Moderate 1–3	408 (43.0)	68 (40.7)
Severe >3	181 (19.1)	19 (11.4)
Renal biopsy, n/n (%)		
Mesangial (M) 1	307/948 (32.4)	99/164 (60.4)
Endocapillary (E) 1	108/948 (11.4)	19/164 (11.6)
Glomerular sclerosis (S) 1	710/948 (74.9)	106/164 (64.6)
Tubulointerstitial damage (T) 1	194/948 (20.5)	52/164 (31.7)
Tubulointerstitial damage (T) 2	44/948 (4.6)	15/164 (9.1)
Crescent (C) 1	86/948 (9.1)	35/120 (29.2)
Therapy, n (%)		
RASBs	577 (60.9)	110 (65.9)
Corticosteroids/cytotoxic agents	258 (27.2)	65 (38.9)
Follow-up, mo	89.0 (50.0–134.0)	76.1 (45.5–143.2)
Year-man, yr	7952	1411
Clinical outcome		
ESKD or eGFR < 15 ml/min per 1.73 m^2, n (%)	210 (22.2)	23 (13.8)

eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease; IgAN, IgA nephropathy; KDOQI, Kidney Disease Outcomes Quality Initiative; MDRD, Modification of Diet in Renal Disease; RASBs, renin-angiotensin system blockers.
[a]Therapy, follow-up, and clinical outcome are included.
Data are expressed as mean ± SD, median (interquartile range), absolute, and percent frequency.

Risk factors and renal survival

During a median follow-up of 89.0 months, 210 patients (22.2%) had reached the endpoint of ESKD, dialysis, or transplantation. The 5-, 10-, 15-, and 20-year renal survival rates in all patients from the time of the renal biopsy were 88.9% (patients at risk, 666); 77.2% (patients at risk, 297); 67.7% (patients at risk, 115); and 56.3% (patients at risk, 35), respectively.

Unadjusted and adjusted hazard ratios (HRs) that can estimate the association between the different risk factors and ESKD are reported in Table 2. The risk of ESKD significantly increased for every 1.0-mg/dL increase in the serum creatinine level (adjusted HR 1.53, 95% confidence interval [CI] 1.35–1.73, $P < 0.001$) and increased for every 1.0 g/d in the

Table 2 | Unadjusted and adjusted risk estimates by Cox proportional hazard models for ESKD in 948 patients with IgA nephropathy in the study cohort

Risk factor	Unit of increase or reference	Unadjusted risk (HR) (95% CI)	*P* value	Adjusted risk (HR) (95% CI)	*P* value
Age, yr	1	1.02 (1.01–1.03)	<0.001	1.00 (0.99–1.01)	0.726
Gender (female/male)	Female	1.38 (0.99–1.92)	0.050	0.78 (0.54–1.09)	0.137
Arterial hypertension	0 = absent 1 = present	1.98 (1.51-2.60)	<0.001	1.25 (0.93–1.67)	0.131
Systolic arterial presssure, mm Hg	1	1.02 (1.01–1.03)	<0.001	Not selected	
Diastolic arterial presssure, mm Hg	1	1.03 (1.02–1.05)	<0.001	Not selected	
Mean arterial presssure, mm Hg	1	1.03 (1.02–1.04)	<0.001	Not selected	
Proteinuria, g/d	1	1.17 (1.13-1.22)	<0.001	1.16 (1.11–1.22)	<0.001
Serum creatinine, mg/dl	1	1.72 (1.64–1.90)	<0.001	1.53 (1.35–1.73)	<0.001
eGFR (MDRD), ml/min per 1.73 m^2	1	0.96 (0.95–0.97)	<0.001	Not selected	
Mesangial (M)	0 = absent 1 = present	1.69 (1.29-2.23)	<0.001	1.06 (0.79–1.43)	0.676
Endocapillary (E)	0 = absent 1 = present	1.22 (0.80–1.85)	0.359	1.14 (0.74–1.76)	0.556
Glomerular stenosis (S)	0 = absent 1=present	3.27 (2.12–5.06)	<0.001	1.87 (1.19–2.96)	0.007
Tubulointerstitial damage (T)	T0				
	T1	5.30 (3.93–7.14)	<0.001	3.19 (2.26–4.51)	<0.001
	T2	9.50 (6.16–14.65)	<0.001	4.74 (2.94–7.63)	<0.001
Crescent (C)	0 = absent 1 = present	0.94 (0.56–1.59)	0.826	0.71 (0.43–1.22)	0.212
Renin-angiotensin system blockers	0 = no; 1 = yes	1.71 (1.28–2.28)	<0.001	1.40 (1.03–1.88)	0.030
Corticosteroids/cytotoxic drugs	0 = no; 1 = yes	1.13 (0.82–1.54)	0.462	0.58 (0.41–0.82)	0.002

CI, confidence interval; eGFR, estimated glomerular filtration rate; HR, hazard ratio; MDRD, Modification of Diet in Renal Disease.
Basic (initial) – 2 log L = 2544.540; Lower (lowest) – 2 log L = 2286.248; $P < 0.0001$.

daily proteinuria (adjusted HR 1.16, 95% CI 1.11–1.22, $P < 0.001$). Furthermore, there was a strong association between the risk of ESKD and the presence of tubulointerstitial lesions and glomerular sclerosis at renal biopsy (tubulointerstitial damage [T] 1: adjusted HR 3.19, 95% CI 2.26–4.51, $P < 0.001$; T2: adjusted HR 4.74, 95% CI 2.94–7.63, $P < 0.001$; glomerular sclerosis [S] 1: adjusted HR 1.87, 95% CI 1.19–2.96, $P = 0.007$). Finally, therapy with corticosteroids or cytotoxic drugs significantly reduced the risk of ESKD (adjusted HR 0.58, 95% CI 0.41–0.82, $P = 0.002$).

Performance of the ANNs

The combined panel of 7 clinical variables: age, sex, hypertension, serum creatinine, proteinuria, MEST-C (mesangial, endocapillary, glomerular stenosis, tubulointerstitial damage, crescent) classification for histologic lesions, and therapy, selected by Cox regression analysis was used for our ANN model.

To provide a good performance of the prediction model, our study cohort dataset was randomly divided into 2 subsets (Supplementary Table S3). The training subset was formed by 758 patients (80%) and the test subset by 190 patients (20%). The first subset was used to perform 10-fold cross-validation to select the best ANN model. The second subset was used to evaluate the technologies described in the Supplementary Methods section. The 7 variables selected by regression analysis were used in both subsets.

The performance value was 0.82 (area under the receiver operating characteristic [ROC] curve [AUC]) in patients with IgAN with a median follow-up of 5 years (Figure 1a). The ROC curve had 0.92 sensitivity (recall), 0.80 accuracy, and 0.83 precision. In the model of 10-year median follow-up, the AUC was 0.89 with a sensitivity (recall) of 0.89, accuracy of 0.83, and precision of 0.81 (Figure 1b). We have shown here that both models have a high recall rate, indicating the practicality of the ANN.

Dynamic prediction of ESKD by ANN

The dynamic prediction of ANN as assessed at 1 and 2 years to predict ESKD over 5 and 25 years of follow-up was investigated in 932 patients—that is, in the combined cohort of training and test subsets. Over a 5-year follow-up period, the dynamic discriminations (AUCs) of 1- and 2-year ANN were 0.80 and 0.79, respectively, and similar results were obtained over the extended (25-year) follow-up period (1-year ANN, AUC 0.82; 2-year ANN, AUC 0.78).

Regression model

The second step was based on the development of a regression model trained to analyze the event-time interval of ESKD prediction. To test our second model, we considered the set of study cohort patients who had reached ESKD during the follow-up of 10 or more years. The mean absolute error and root mean square error were used as the metrics to evaluate the performance of our model and were applied to 2 independent cohorts: the training subset and the test subset, which was composed of randomly selected 20% of the patients. The regression model showed mean absolute error and root mean square error values of 1.78 years and 2.15, respectively.

Test cohort

We collected clinical and laboratory findings of an independent cohort of 178 patients with biopsy-proven IgAN from 6 renal units. We removed 11 patients because of missing data; thus, the reproducibility of our model was evaluated in 167 patients. As shown in Table 1, at the time of the kidney biopsy, the mean patient age was 40.1 ± 15.5 years; the median values of serum creatinine were 1.12 (0.88–1.50) mg/dl; eGFR

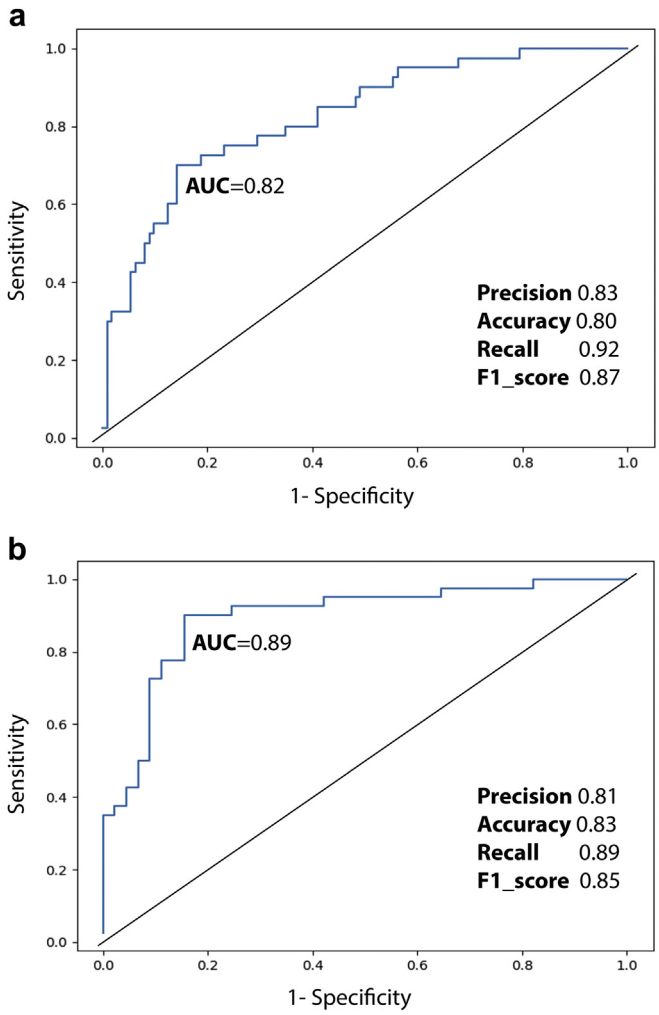

Figure 1 | Receiver operating characteristic curves of the artificial neural network models at (a) 5 years and (b) 10 years of end-stage kidney disease risk. AUC, area under the curve.

69.00 (48.80–91.00) ml/min per 1.73 m^2; and daily proteinuria 1.00 (0.50–1.80) g/d. Hypertension was present in 35% of patients. The median duration of follow-up was 76.1 (45.5–143.2) months. As shown in Figure 2, our model predicted the development of ESKD in 44 of 167 patients (26.3%). Sixteen individuals received dialysis or renal transplant. In 20 subjects with eGFR higher than 45 ml/min per 1.73 m^2, proteinuria higher than 1.0 g/d, and moderate to severe renal lesions, their renal function improved after RASBs and corticosteroid therapy (pulse in 8 patients, oral in 112 patients for 6 months with gradual daily dose reduction). ESKD was erroneously predicted in 8 patients who maintained normal renal function. The tool predicted no ESKD in 123 of 167 patients (73.6%), 6 of whom reached ESKD. In conclusion, a prediction error was observed in 14 to 167 patients (8.4%) of the test cohort.

Comparisons of different models

Many investigators have proposed different risk score models, based on mathematical models and statistical analyses, to predict ESKD in patients with IgAN (Supplementary Table S2). We blindly compared our CDSS model with those of Okonogi *et al.*,[13] Berthoux *et al.*,[14] Tanaka *et al.*,[16] and Barbour *et al.*[18] Therefore, we established a cohort of 150 patients with IgAN with 5-year follow-up randomly selected in the study cohort (baseline characteristics of the patients are shown in Supplementary Table S4). Then, all scores, save those for Barbour *et al.*,[18] were also tested to predict ESKD in a cohort of 83 patients with IgAN with 10-year follow-up randomly selected in the study cohort (Supplementary Table S4).

As shown in Tables 3 and 4, the short-term (81%) and long-term (86%) discriminatory power of CDSS for predicting ESKD was superior to the scores provided by Okonogi *et al.*[13] (76% and 74%), Berthoux *et al.*[14] (58% and 55%), and Tanaka *et al.*[16] (78% and 76%). At 5 years, the discriminatory power of the CDSS was almost identical to that of Barbour *et al.*[18] (81% vs. 82%). The CDSS also showed the lowest Akaike information criteria (the lower the Akaike information criterion, the better the prognostic estimates) in both the short-term and long-term follow-up periods (Tables 3 and 4). In more detail, the calculation of Akaike weights showed that the model including the ANN had 83.1% chances to be the best model to predict the 5-year risk of ESKD followed by the scores of Barbour *et al.*[18] (13.7%), Tanaka *et al.*[16] (3.0%), Okonogi *et al.*[13] (0.16%), and Berthoux *et al.*[14] (0.04%). Of note, at 10 years the probability of the ANN to perform the best among the 4 candidate risk prediction rules was 99.9%.

The CDSS-based model was also the best calibrated rule at 5-year follow-up because the *P* values of the formal comparison between the observed and predicted risk of ESKD, as provided by the CDSS, were higher (i.e., further from the statistical significance) than those provided by the 4 traditional scores. At 10 years, the calibration ability of the CDSS was satisfactory.

CDSS description

Supplementary Figures S1 and S2 illustrate the structure of the ANN tool that may be widely used by physicians for predicting the ESKD at the 5- and 10-year follow-ups for patients with IgAN at the time of kidney biopsy diagnosis. The CDSS is available at https://igan.poliba.it/, and the website can be accessed by cell phone. Physicians may use the

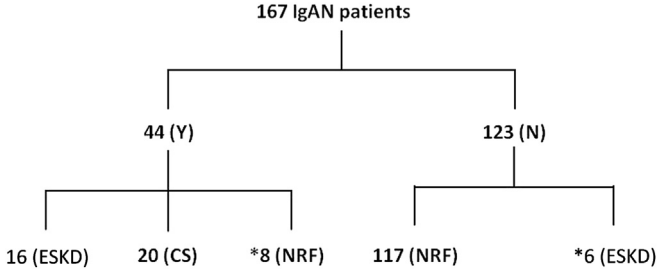

Figure 2 | Tool results in the test cohort. CS, corticosteroids; ESKD, end-stage kidney disease; IgAN, IgA nephropathy; N, no ESKD; NRF normal renal function; Y, yes ESKD. *Error prediction.

Table 3 | Short-term (5-year) prognostic accuracies of the 4 predictive scores compared with ANN tool

Authors	Harrell C index	Akaike information criterion	May-Hosmer test
Okonogi *et al.*[13]	76%	367.48	$\chi^2 = 5.36$ ($P = 0.15$)
Berthoux *et al.*[14]	58%	390.08	$\chi^2 = 2.26$ ($P = 0.52$)
Tanaka *et al.*[16]	78%	361.64	$\chi^2 = 2.82$ ($P = 0.42$)
Barbour *et al.*[18]	82%	358.62	$\chi^2 = 4.05$ ($P = 0.26$)
ANN model	81%	355.01	$\chi^2 = 1.21$ ($P = 0.76$)

ANN, artificial neural network.

model by entering the features at the time of the kidney biopsy. The tool predicts the ESKD (yes or no), and in the presence of the potential development of ESKD, the tool indicates the time to reach the outcome (5–10 years). The physician may test the effect of RASBs or immunosuppressive therapy on the outcome of the patient. We have reset the CDSS tool 1year after the kidney biopsy as baseline in a subgroup of 314 patients with IgAN who continued RASB therapy. As shown in Supplementary Figure S3, our tool still predicted no deterioration of the renal function in 216 patients and ESKD in 50 subjects, demonstrating the correct prediction of our tool in 266 of 314 patients (84.7%) and the benefit of RASB therapy in 216 of 239 subjects (90.4%) as recently suggested by the last Kidney Disease Improving Global Outcomes clinical practice guidelines on Glomerular Diseases.[19] The reset was also done for 61 patients of the test cohort treated with RASBs; the prediction of no deterioration of renal function was confirmed in 48 of 52 (92.3%) and ESKD in other patients. In conclusion, these results strengthen the data of the study cohort.

DISCUSSION

We developed a CDSS that can predict ESKD in patients with IgAN with a median follow-up of 5 and 10 years. The CDSS is simple to use because it acquires the bulk of the data at the time of kidney biopsy, a time when it is important for the physician to know if the patient will reach the final outcome of ESKD.

Our CDSS has been built on links between clinical data and good standards for accuracy. In the first phase, we assembled a panel of risk factors by reviewing the clinical and histologic studies published over the past 30 years, as shown in the Supplementary Table S1. Then we chose 7 variables selected by the Cox regression analysis in the study cohort.

Table 4 | Long-term (10-year) prognostic accuracies of the 3 predictive scores compared with ANN tool

Authors	Harrell C index	Akaike information criterion	May-Hosmer test
Okonogi *et al.*[13]	74%	304.32	$\chi^2 = 12.67$ ($P = 0.005$)
Berthoux *et al.*[14]	55%	330.87	$\chi^2 = 2.68$ ($P = 0.44$)
Tanaka *et al.*[15]	76%	297.40	$\chi^2 = 8.96$ ($P = 0.03$)
ANN model	86%	269.56	$\chi^2 = 4.73$ ($P = 0.19$)

ANN, artificial neural network.

The combination of variables for assessing the risk of ESKD in the general population was first reported by Hallan *et al.*[20] Many investigators have used combined clinical variables recorded at the time of kidney biopsy to improve the accuracy of prediction of renal outcome in patients with IgAN (Supplementary Table S2). The combination of clinical parameters has been based on statistical methods, but these studies[11–18] have shown several limitations. First, the different histologic classifications, not the MEST-C classification, were used in the studies in references 11 through 17. Second, the risk scores were not tested in independent cohorts in those studies. We tested our tool using an external cohort of 167 patients with biopsy-proven IgAN. Finally, the comparison with other mathematical models has shown that our prediction tool is more precise than others (Tables 3 and 4). Further, the calibration ability of our tool is acceptable at 10 years. Recently Barbour *et al.*[18] developed a risk prediction tool in which the MEST classification (without the crescent [C] lesion) was used and where the time prediction was 60 to 80 months. Comparisons between these 2 tools are shown in Supplementary Table S5. Principally, Barbour *et al.*[18] use an endpoint composed of eGFR $<$ 15 ml/min per 1.73 m^2, ESKD, transplantation, and permanent reduction in eGFR to less than 50% of the baseline value. The last endpoint is questionable. When hard clinical outcomes are too expensive to study in randomized clinical trials in nephrology, an alternative is to target surrogate endpoints, such as permanent reduction of more than 50% of the baseline eGFR, is used. This outcome should not be used in long-term retrospective studies. Moreover, IgAN is a slow, progressive indolent long-term disease that may maintain a permanent reduction of more than 50% of baseline eGFR for more than 80 months, mainly in patients with chronic kidney disease stage 2 and 3. Therefore, we choose the endpoint ESKD (yes or no) and the time-to-event of 60 to 120 months.

Machine learning can assemble large clinical databases and generate tools for making decisions in various fields of human health.[21] ANNs have the advantage of detecting complex nonlinear relationships between independent and dependent variables and all the possible interactions between the predictor variables, but an ANN cannot identify possible causal relationships that can be studied by computational resources. The methodology used in our study was based on deep machine learning, using ANNs that can (i) learn extremely complex relationships between the heterogeneous kind of data generated by clinical care and (ii) exceed human abilities in performing tasks.[6] Recently new methods for the time-to-event prediction were proposed by extending the Cox proportion hazard model using neural networks.[22]

Our previous tool was based on a classic ANN formed by 2 hidden layers.[23] After exploiting the new ANN techniques, we used 4 hidden layers with 100 neurons in each layer for the classification model and 3 hidden layers with 125 neurons for each layer for the regression model. Moreover, we introduced additional advanced techniques, such as batch

normalization, that reduce the internal covariate shift in the classification model. Finally, we introduced a new proxy AUC function.

Furthermore, we used the exponential linear unit (ELU) function, which is more flexible than the rectified linear unit, for the classification model and scaled ELU for the regression task. The ELU avoids mean and bias shifts to improve learning, and the scaled ELU is a slightly modified formula of ELU. For the regression model, we introduced dropout layers.

Our CDSS can identify patients at low and high risk of ESKD development in the next 5 to 10 years for whom therapy may be crucial, as shown in the 20 patients with IgAN of the test cohort who maintained renal function (eGFR $>$ 60 ml/min per 1.73 m^2) after corticosteroid therapy. When combining different variables, the AUC increased from 0.82 to 0.89, suggesting that our model has better diagnostic accuracy in predicting the long-term risk of potential ESKD development in patients with IgAN.

Accurate prediction of the IgAN outcomes based on the individual characteristics of the patient is an important step toward developing personalized medicine, and it may potentially improve the clinical course of the disease, as reported in other diseases.[24] The prediction of ESKD in individual cases is important for the following indications:

- Screening of patients with IgAN who should or should not receive regular examinations during the clinical course of the disease, which can translate into saving time and money
- Therapeutic strategy, such as whether to use antiproteinuric drugs (high doses of RASB) or immunosuppressive drugs after the kidney biopsy
- Evaluating the response to therapy
- Monitoring disease progression

Our study has several strong points. (i) ANN is not expressed by a conventional Cox regression analysis, and it includes, for the first time, the complete MEST-C classification used for histologic diagnosis of kidney biopsies. To facilitate the application, we developed a tool that is easy to use for consultation and to predict the time-to-event interval to achieve ESKD. (ii) By using this approach, a reliable and sustainable therapeutic approach can be achieved more easily and quickly. This approach can improve patient outcomes. (iii) A specific cost analysis should be performed to evaluate the economic input of such tools. The potential future application in clinical practice could be evaluated over the next few years. (iv) The patients with IgAN included in our study were not restricted to a single center. (v) The joint model analysis for longitudinal and time-to-event data showed an adequate time-dependent prognostic ability of our ANN, indicating that the tool provides adequate risk discrimination when measured repeatedly over time.

However, our study also has some limitations. (i) This model has been developed and tested in retrospective cohorts of patients with IgAN; therefore, multicenter prospective cohort studies are needed to evaluate the validity of this tool in other patients with IgAN, ranging from Caucasian to non-Caucasian races. (ii) We need to confirm whether the therapeutic interventions may reduce the number of ESKD events predicted by our tool.

In conclusion, our CDSS is a reliable tool for predicting ESKD in patients with IgAN. It is simple to use because it is based on the input of 7 indicators and can predict the approximate number of years in which ESKD will develop. To further explore the potential value and sustainability of this tool in the management of the disease, a prospective multicenter clinical study enrolling a large number of patients would be helpful.

METHODS

Patient characteristics

Study cohort. In our cohort, we enrolled 2 groups of patients with biopsy-proven IgAN. The first one was the retrospective VALIGA cohort study[25] that included 1147 patients from 13 participating European countries. The second group was composed of 84 patients recruited from the Thessaloniki Renal Unit, Greece.

Figure 3 illustrates the selection process of patients enrolled in our observational study. Patients younger than 18 years (n = 174) were excluded because we focused on IgAN in adults. Then 109 patients were excluded because of missing data. In total, 948 patients with IgAN were included in this retrospective observational longitudinal study.

The diagnosis of IgAN was based on the histologic and immunofluorescence study of the kidney biopsy. Patients with IgAN secondary to Henoch-Schönlein purpura, lupus nephritis, chronic liver diseases, and other immunologic disorders were excluded.

The study cohort of patients with IgAN does not represent the natural history of the disease because after kidney biopsy, 60.9% of the patients were treated with RASBs, and 27.2% received corticosteroids or immunosuppressive drugs. The demographic characteristics and baseline laboratory data of the patients at the time of kidney biopsy are shown in Table 1.

Hypertensive patients were considered when at the time of the renal biopsy, their systolic blood pressure was 140 mm Hg or higher, their diastolic blood pressure was 90 mm Hg or higher, or both, or the patients received antihypertensive drugs. Serum creatinine values were expressed in milligrams per deciliter. The glomerular filtration rate was estimated using the 4 variables in the Modification of Diet in Renal Disease formula.[26] Daily proteinuria was measured by grams per day. Renal biopsy was categorized according to the Oxford classification,[27,28] which considers the 4 lesions as M (mesangial), E (endocapillary), S (glomerular sclerosis), and T (tubulointerstitial damage). The renal lesions were quantified as M (0,1), E (0,1), S (0,1), and T (0,1,2). Then the renal biopsy reports were revised for the capillary lesion (crescent, C) and quantified as either 0 (absent) or 1 (C1 or C2). In conclusion, the kidney biopsy reports were updated in accordance with the MEST-C classification, which was recently published by the International IgAN Classification Working Group.[29]

The primary variable endpoint of the renal outcome was ESKD defined as eGFR less than 15 ml/min/173^2 or the initiation of periodic hemodialysis or kidney transplantation. The renal survival time was calculated from the biopsy to the last follow-up visit.

Test cohort. An independent external cohort of 167 patients with biopsy-proven IgAN from 6 European renal units was used for our ANN model. The demographic characteristics and baseline laboratory data of the patients at the time of kidney biopsy are shown

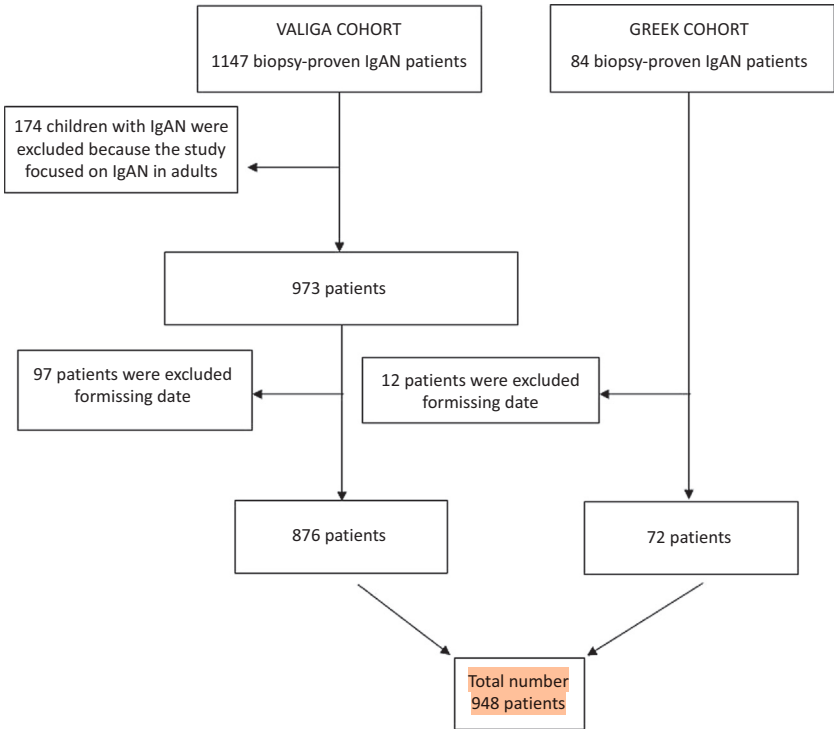

Figure 3 | Flowchart of the patients with IgA nephropathy (IgAN) from the VALIGA and Greek cohorts included in the study.

in Table 1. The study was conducted according to the principles of the Declaration of Helsinki and was approved by the local institutional ethics review boards.

ANN models

In a previous article,[30] we compared ANNs with other machine learning techniques, such as neuro-fuzzy systems, support vector machines, and decision trees, for predicting ESKD in IgAN. ANNs have the potential to analyze multidimensional and nonlinear data, hence providing a correct interpretation that is difficult to obtain by standard statistical analyses. Advances in deep neural networks in medicine have suggested ways for improving our previous model by exploiting new ANN techniques.[31–34] In this study, we tested many techniques described in the Supplementary Methods, and we then developed 2 models, of which the classifier model could predict the ESKD and the second, the regression model, could predict the time to reach ESKD.

The classifier model is composed of 4 hidden layers with 100 neurons in each layer. Variables selected by a multivariate model of adequate statistical power constructed by means of Cox regression analysis were collected at the time of the kidney biopsy (age, sex, hypertension, proteinuria, serum creatinine, histologic renal lesions, and therapy), whereas the outcome (ESKD yes or no) represented the output data.

The regression model consisted of 3 layers containing 125 neurons each. We included patients with IgAN of the study cohort with follow-up between 5 and 10 years or more than 10 years in the dataset. In the ANN scaled exponential linear unit, the activation function was used to better exploit fine differences in the feature values.[35] The mean absolute error, expressed in years in our model, and root mean square error were used as performance metrics for the regression model. Finally, the best ANN architectures and parameters were selected by the K-fold cross-validation for both the classifier and regression models.

The performance of the models was evaluated by calculating the sensitivity (recall), positive predictive value, negative predictive value (specificity), accuracy, and AUC. The ROC curve is an important criterion for evaluating the performance of a classifier. However, because the AUC is not directly optimizable, we designed a new cost function that works as a proxy function for the AUC maximization task. The cost function is directly derived from proofs, and demonstrations are already available in work by Rendle *et al.*[36]

Our ANN model was then compared with other published mathematical models (listed in the Appendix, Table 2). Okonogi *et al.*[13] considered 2 parameters (proteinuria and eGFR). Berthoux *et al.*[14] considered 3 parameters (hypertension, proteinuria, and histological grade). Tanaka *et al.*[16] considered 5 parameters (age, hypertension, eGFR, proteinuria, and histologic grade). Barbour *et al.*[18] considered 7 parameters (race, age, hypertension, eGFR, proteinuria, histologic MEST score, and medication used before the renal biopsy). The Barbour model was not compared at 10 years because it was not specifically developed to predict ESKD for more than 80 months.

Statistical analysis

Baseline sociodemographic and clinical characteristics were calculated and expressed as the mean plus or minus the SD or median (interquartile range 25th–75th percentile) for continuous variables and as the absolute and percent frequency for categorical variables, as appropriate. Renal survival time from the ESKD endpoint was calculated from the biopsy to the last follow-up. Cumulative renal survival was analyzed by Kaplan-Meier curves for censored data. The potential nonlinear effects of the exposure factors were explored as appropriate and reported if identified.

Univariate and multivariate analyses based on the Cox regression proportional hazard model[37] were used to assess the relative risk of ESKD based on the influence of baseline prognostic factors. Variables that could significantly predict ESKD in a univariate analysis ($P < 0.05$) or if they were clinically relevant, were used to construct a multivariate model of adequate statistical power by means of backward or forward stepwise approaches. Risk estimates were presented as unadjusted and adjusted HRs and their 95% CIs, which were calculated by using an estimated regression coefficient and its standard error.

To assess the time-dependent prognostic ability of the ANN, measured repeatedly over time, we applied joint models for longitudinal and time-to-event data. The discriminative capability (AUC) of the ANN score was assessed using a dynamic discrimination index, defined as the weighted average of time-dependent AUCs, calculated over 1- and 2-year time horizons, across the entire follow-up period.[38]

The CDSS for predicting ESKD was compared with 4 predictive scores[13,14,16,18] by using the Harrell C index (an index of discrimination),[39] the Akaike information criterion,[40] and the May-Hosmer test (an index of calibration).[41] In brief, discrimination measures how well a prognostic model distinguishes (discriminates) patients with and without the outcome of interest. Discrimination was measured by the Harrell C index, which at a variance of the ROC curve analysis, takes into account the survival time. This index can take values ranging from 0.5 (no discrimination) to 1.0 (perfect discrimination), and its interpretation is similar to that of the ROC curve analysis. The higher the Harrell C index, the higher the accuracy of the model for predicting the event of interest. In our study, calibration measures how much the prognostic estimate of a specific predictive model matches the real probability of the outcome (i.e., the observed proportion of an event in a given time period). In the calibration analysis, the predicted and observed probabilities of the event are compared with the May-Hosmer test. A nonsignificant May-Hosmer test indicates that the predicted and observed probabilities of the event do not differ, and thus the calibration of the model is satisfactory. The Akaike information criterion provides an objective way of determining which model among a set of candidate models is the best for predicting the occurrence of a given event. The lower the AIC, the better the prognostic model is for predicting the event of interest. All analyses were performed using SPSS for Windows version 17.0 (SPSS Inc., North Sydney, Australia) and Stata/IC release 13.1 for Windows (StataCorp, College Station, TX). A P value less than 0.05 was considered statistically significant.

APPENDIX

Nephrology units of the VALIGA study

We are deeply grateful to the nephrology units of the VALIGA study (asterisks mark the centers that sent the update by 2016): V. Tesar, D. Maixnerova (Nephrology, First Faculty of Medicine and General University Hospital, Prague, Czech Republic)*; S. Lundberg (Nephrology, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden)*; L. Gesualdo (Nephrology, Emergency and Organ Transplantation, University of Bari "Aldo Moro," Foggia-Bari, Italy)*; F. Emma, L. Fuiano (Nephrology, Pediatrico Bambino Gesù Hospital, Rome, Italy)*; G. Beltrame, C. Rollino (Nephrology, San Giovanni Bosco Hospital, Turin, Italy)*; R. Coppo, A. Amore, R. Camilla, L. Peruzzi (Nephrology, Regina Margherita Children's Hospital, Turin, Italy)*; M. Praga (Nephrology, Hospital 12 de Octubre, Madrid, Spain)*; S. Feriozzi, R. Polci, (Nephrology, Belcolle Hospital, Viterbo, Italy)*; G. Segoloni, L. Colla (Nephrology, S. Giovanni Battista University Hospital, Turin, Italy)*; A. Pani, A. Angioi, L. Piras (Nephrology, G. Brotzu Hospital, Cagliari, Italy)*; J. Feehally (John Walls Renal Unit, Leicester General Hospital, Leicester, UK)*; G. Cancarini, S. Ravera (Nephrology, Spedali Civili University Hospital, Brescia, Italy); M. Durlik (Transplantation Medicine and Nephrology, Warsaw Medical University, Warsaw, Poland)*; E. Moggia (Nephrology, Santa Croce Hospital, Cuneo, Italy)*; J. Ballarin (Nephrology, Fundacion Puigvert, Barcelona, Spain)*; S. Di Giulio (Nephrology, San Camillo Forlanini Hospital, Rome, Italy); F. Pugliese, I. Serriello (Nephrology, Policlinico Umberto I University Hospital, Rome, Italy)*; Y. Caliskan, M. Sever, I. Kilicaslan (Nephrology, Internal Medicine, Istanbul Faculty of Medicine, Istanbul, Turkey)*; F. Locatelli, L. Del Vecchio (Nephrology, A. Manzoni Hospital, Lecco, Italy)*; J.F.M. Wetzels, H. Peters (Nephrology and Pathology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands)*; U. Berg (Pediatrics, Department of Clinical Science, Intervention and Technology, Huddinge, Sweden)*; F. Carvalho, A.C. da Costa Ferreira (Nephrology, Hospital de Curry Cabral, Lisbon, Portugal)*; M. Maggio (Nephrology, Hospital Maggiore di Lodi, Lodi, Italy)*; A. Wiecek (Nephrology, Endocrinology and Metabolic Diseases, Silesian University of Medicine, Katowice, Poland); M. Ots-Rosenberg (Nephrology, Tartu University Clinics, Tartu, Estonia)*; R. Magistroni (Nephrology, Policlinic of Modena and Reggio Emilia; Modena, Italy); R. Topaloglu, Y. Bilginer (Pediatric Nephrology and Rheumatology, Hacettepe University, Ankara, Turkey)*; M. D'Amico (Nephrology, S. Anna Hospital, Como, Italy)*; M. Stangou (Nephrology, Hippokration General Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece)*; F. Giacchino (Nephrology, Ivrea Hospital, Ivrea, Italy)*; D. Goumenos, M. Papasotiriou, (Nephrology, University Hospital of Patras, Patras, Greece)*; K. Galesic, L. Toric (Nephrology, University Hospital Dubrava, Zagreb, Croatia)*; C. Geddes (Renal Unit, Western Infirmary Glasgow, Glasgow, UK)*; K. Siamopoulos, O. Balafa (Nephrology, Medical School University of Ioanina, Ioannina, Greece)*; M. Galliani (Nephrology, S. Pertini Hospital, Rome, Italy); P. Stratta, M. Quaglia (Nephrology, Maggiore della Carità Hospital, Piemonte Orientale University, Novara, Italy)*; R. Bergia, R. Cravero (Nephrology, DeglilNfermi Hospital, Biella, Italy)*; M. Salvadori, L. Cirami (Nephrology, Careggi Hospital, Florence, Italy)*; B. Fellstrom, H. Kloster Smerud (Renal Department, University of Uppsala, Uppsala, Sweden)*; F. Ferrario, T. Stellato (Nephropathology, San Gerardo Hospital, Monza, Italy); J. Egido, C. Martin (Nephrology, Fundación Jiménez Diaz, Madrid, Spain)*; J. Floege, F. Eitner, T. Rauen (Nephrology and Immunology, Medizinische Klinik II, University of Aachen, Aachen, Germany)*; A. Lupo, P. Bernich (Nephrology, University of Verona, Verona, Italy); P. Menè (Nephrology, S. Andrea Hospital, Rome, Italy); M. Morosetti (Nephrology, Grassi Hospital, Ostia, Italy); C. van Kooten, T. Rabelink, M.E.J. Reinders (Nephrology, Leiden University Medical Centre, Leiden, The Netherlands)*; J.M. Boria Grinyo (Nephrology, Hospital Bellvitge, Barcelona, Spain); S. Cusinato, L. Benozzi (Nephrology, Borgomanero Hospital, Borgomanero, Italy)*; S. Savoldi, C. Licata (Nephrology, Civile Hospital, Ciriè, Italy)*; M. Mizerska-Wasiak, M. Roszkowska-Blaim (Pediatrics, Medical University of Warsaw, Warsaw, Poland); G. Martina, A. Messuerotti (Nephrology, Chivasso Hospital, Chivasso, Italy)*; A. Dal Canton, C. Esposito, C. Migotto (Nephrology Units, S. Matteo Hospital and Maugeri Foundation, Pavia, Italy); G. Triolo, F. Mariano (Nephrology CTO, Turin, Italy)*; C. Pozzi (Nephrology, Bassini Hospital, Cinisello Balsamo, Italy)*; R. Boero (Nephrology, Martini Hospital, Turin, Italy)*.

List of pathologists included in the VALIGA study

G. Mazzucco (Turin, Italy); C. Giannakakis (Rome, Italy); E. Honsova (Prague, Czech Republic); B. Sundelin (Stockholm, Sweden); A.M. Di Palma (Foggia-Bari, Italy); F. Ferrario (Monza, Italy); E. Gutiérrez (Madrid, Spain); A.M. Asunis (Cagliari, Italy); J. Barratt (Leicester, UK); R. Tardanico (Brescia, Italy); A. Perkowska-Ptasinska (Warsaw, Poland); J. Arce Terroba (Barcelona, Spain); M. Fortunato (Cuneo, Italy); A. Pantzaki (Thessaloniki, Greece); Y. Ozluk (Istanbul, Turkey); E. Steenbergen (Nijmegen, The Netherlands); M. Soderberg (Huddinge, Sweden); Z. Riispere (Tartu, Estonia); L. Furci (Modena, Italy); D. Orhan (Ankara, Turkey); D. Kipgen (Glasgow, UK); D. Casartelli (Lecco, Italy); D. Galesic Ljubanovic (Zagreb, Croatia); H. Gakiopoulou (Athens, Greece), E. Bertoni (Florence, Italy); P. Cannata Ortiz (Madrid, Spain); H. Karkoszka (Katowice, Poland), H.J. Groene (Heidelberg, Germany); A. Stoppacciaro (Rome, Italy); I. Bajema, J. Bruijn (Leiden, The Netherlands); X. Fulladosa Oliveras (Barcelona, Spain); J. Maldyk (Warsaw, Poland); and E. Ioachim (Ioannina, Greece).

The authors are grateful to Stéphan Troyanov for his work on the first analysis of the original VALIGA database. We thank Daniela Abbrescia, PhD, who filled in the files for the joint model analysis, and the nephrologists of the 6 renal units who collected the findings of the IgAN patients for the test cohort: Nikoleta Kouri, Maria Stangou, Aikaterini Papagianni, Department of Nephrology, Hippokration General Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece; Francesco Scolari, Elisa Delbarba, Nephrology Unit, University of Brescia; Mario Bonomini, Luca Piscitani, Nephrology Unit, University of Chieti; Giovanni Stallone, Barbara Infante, Giulia Godeas, Desiree

Madio, Nephrology Unit, University of Foggia; Luigi Biancone, Marco Campagna, Nephrology Unit, University of Torino; and Gianluigi Zaza, Isabella Squarzoni, Concetta Cangemi, Nephrology Unit, University of Verona.

DISCLOSURE

All the authors declared no competing interests

AUTHOR CONTRIBUTIONS

FPS and TD developed the concept and designed the study; RC, VT, MS, AP, GD, and MLR collected and evaluated the data; FPS, VWA, JT, CM, and GT analyzed the data and interpreted the results; NCC analyzed data and developed the joint models; FPS, VWA, CM, and GT wrote the manuscript; RC, VT, TD, and CZ critically revised the manuscript for intellectual content; all authors revised and approved the final manuscript.

DATA SHARING

Collected data of the patients were anonymously transferred for the analyses.

ACKNOWLEDGMENTS

The work was financially supported by the following grants: ERA-EDTA 2009 (RC); PON REC 02_00134/2011 (FPS), Italy. The funders had no role in the design and conduct of the study.

Data of the pilot study were presented at the American Society of Nephrology Congress, October 31–November 5, 2017, New Orleans, LA, USA.

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)

Supplementary Methods. Background technologies.

Table S1. Clinical and histologic risk factors for ESKD in patients with IgAN. Data from the literature.

Table S2. Scoring systems for predicting ESKD in patients with IgAN.

Table S3. Baseline characteristics of the patients with IgAN included in the training and test set of the study cohort. Therapy, follow-up, and clinical outcome are shown.

Table S4. Baseline characteristics of the patients with IgAN of the study cohort divided in 2 different subsets based on the mean follow-up. Therapy, follow-up, and clinical outcome are shown.

Table S5. Comparisons between two models for the prediction of ESKD risk.

Figure S1. The structure of the artificial neural network (ANN) tool.

Figure S2. The application screen of the IgAN CDSS.

Figure S3. Patients with IgAn of the study cohort treated with RASBs. The tool has been reset after 12 months from the kidney biopsy.

REFERENCES

1. Miotto R, Wang F, Wang S, et al. Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform*. 2018;19:1236–1246.
2. Litjens GJS, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88.
3. Cruz-Ramírez M, Hervás-Martínez C, Fernández JC, et al. Predicting patient survival after liver transplantation using evolutionary multi-objective artificial neural networks. *Artif Intell Med*. 2013;58:37–49.
4. Rajkomar A, Dear J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380:1347–1358.
5. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*. 2018;15:20170387.
6. Hinton G. Deep learning—a technology with the potential to transform health care. *JAMA*. 2018;320:1101–1102.
7. Almeida JS. Predictive non-linear modeling of complex data by artificial neural networks. *Curr Opin Biotechnol*. 2002;13:72–76.
8. Markey MK, Tourassi GD, Margolis M, DeLong DM. Impact of missing data in evaluating artificial neural networks trained on complete data. *Comput Biol Med*. 2006;36:516–525.
9. Schena FP, Nistor I. Epidemiology of immunoglobulin A nephropathy. A global perspective. *Semin Nephrol*. 2018;38:435–442.
10. Manno C, Strippoli GF, D'Altri C, et al. A novel simpler histological classification for renal survival in IgA nephropathy: a retrospective study. *Am J Kidney Dis*. 2007;49:763–775.
11. Wakai K, Kawamura T, Endoh M, et al. A scoring system to predict renal outcome in IgA nephropathy: from a nationwide prospective study. *Nephrol Dial Transplant*. 2006;21:2800–2808.
12. Goto M, Wakai K, Kawamura T, et al. A scoring system to predict renal outcome in IgA nephropathy: a nationwide 10-year prospective cohort study. *Nephrol Dial Transplant*. 2009;24:3068–3074.
13. Okonogi H, Utsunomiya Y, Miyazaki Y, et al. A predictive clinical grading system for immunoglobulin A nephropathy by combining proteinuria and estimated glomerular filtration rate. *Nephron Clin Pract*. 2011;118:c292–c300.
14. Berthoux F, Mohey H, Laurent B, et al. Predicting the risk for dialysis or death in IgA nephropathy. *J Am Soc Nephrol*. 2011;22:752–761.
15. Xie J, Kiryluk K, Wang W, et al. Predicting progression of IgA nephropathy: new clinical progression risk score. *PLoS One*. 2012;7:e38904.
16. Tanaka S, Ninomiya T, Katafuchi R, et al. Development and validation of a prediction rule using the Oxford classification in IgA nephropathy. *Clin J Am Soc Nephrol*. 2013;8:2082–2090.
17. Knoop T, Vågane AM, Vikse BE, et al. Addition of eGFR and age improves the prognostic absolute renal risk-model in 1,134 Norwegian patients with IgA nephropathy. *Am J Nephrol*. 2015;41:210–219.
18. Barbour SJ, Coppo R, Zhang H, et al. Evaluating a new international risk-reduction tool in IgA Nephropathy. *JAMA Intern Med*. 2019;179:942–952.
19. KDIGO Disease: Improving Global Outcomes (KDIGO). Clinical practice guidelines on glomerular diseases. June 2020. Available at: www.kdigo.org. Accessed July 18, 2020.
20. Hallan SI, Ritz E, Lydersen S, et al. Combining GFR and albuminuria to classify CKD improves prediction of ESRD. *J Am Soc Nephrol*. 2009;20:1069–1077.
21. Bean AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319:1317–1318.
22. Kvamme H, Borgan O, Scheel I. Time-to-event prediction with neural networks and Cox regression. *J Mach Learn Res*. 2019;20:1–30.
23. Pesce F, Diciolla M, Binetti G, et al. Clinical decision support system for end-stage kidney disease risk estimation in IgA nephropathy patients. *Nephrol Dial Transplant*. 2016;31:80–86.
24. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet*. 2019;393:1577–1579.
25. Coppo R, Troyanov S, Bellur S, et al. VALIGA study. Validation of the Oxford classification of IgA nephropathy in cohorts with different presentations and treatments. *Kidney Int*. 2014;86:828–836.
26. Levey AS, Bosch JP, Lewis JB, et al. more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med*. 1999;130:461–470.
27. Cattran DC, Coppo R, Cook HT, et al. The Oxford classification of IgA nephropathy: rationale, clinicopathological correlations, and classification. *Kidney Int*. 2009;76:534–545.
28. Roberts IS, Cook HT, Troyanov S, et al. The Oxford classification of IgA nephropathy: pathology definitions, correlations, and reproducibility. *Kidney Int*. 2009;76:546–556.
29. Trimarchi H, Barratt J, Cattran DC, et al. Oxford Classification of IgA nephropathy 2016: an update from the IgA Nephropathy Classification Working Group. *Kidney Int*. 2017;91:1014–1021.
30. Di Noia T, Ostuni VC, Pesce F, et al. An end stage kidney disease predictor based on an artificial neural networks ensemble. *Expert Syst. Appl*. 2013;40:4438–4445.
31. Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs). 2015. Available at: http://arxiv.org/abs/1511.07289. Accessed July 18, 2020.
32. Liu F, Li H, Ren C, et al. PEDLA: predicting enhancers with a deep learning–based algorithmic framework. *Sci Rep*. 2016;6:28517.
33. Wang Y, Zeng J. Predicting drug-target interactions using restricted Boltzmann machines. *Bioinformatics*. 2013;29:126–134.
34. Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn. Res*. 2014;15:1929–1958.
35. Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-normalizing neural networks. In: Guyon A, Luxburg UV, Bengio S, et al., eds. *Advances in Neural Information Processing Systems Proceedings 30*. Long Beach, CA: Neural Information Processing Systems; 2017:972–981.

36. Rendle S, Freudenthaler C, Gantner Z, et al. Bayesian Personalized Ranking from Implicit Feedback. In: Bilmes JA, Ng Ay, eds. *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. Montreal, QC, Canada: AUAI Press; 2009:452–461.
37. Cox DR. Regression models and life-tables. *J R Statist Soc B*. 1972;34:187–220.
38. Rizopoulos DJM. An R package for the joint modelling of longitudinal and time-to-event data. *J Stat Softw*. 2010;35:1–33.
39. Tripepi G, Heinze G, Jager KJ, et al. Risk prediction models. *Nephrol Dial Transplant*. 2013;28:1975–1980.
40. Akaike H. Information theory as an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. *Second international symposium on information theory*. Budapest: Akaemiai Kiado; 1973:267–281.
41. Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res*. 2016;25:1692–1706.