Full citation: Marshall, M.R., Song, Q., Ma, T.M., MacDonell, S.G., & Kasabov, N.K. (2005) Evolving connectionist system versus algebraic formulas for prediction of renal function from serum creatinine, Kidney International 67(5), pp.1944-1954.
doi:10.1111/j.1523-1755.2005.00293.x

Evolving connectionist system versus algebraic formulas for prediction of renal function from serum creatinine

Mark R. Marshall, Qun Song, Tian M. Ma, Stephen G. MacDonell, and Nikola K. Kasabov

Department of Renal Medicine, Middlemore Hospital, Auckland, New Zealand;
KEDRI, Auckland University of Technology, Auckland, New Zealand

mrmarshall@middlemore.co.nz

Abstract

Background. *In clinical trials, equation 7 from the Modification of Diet in Renal Disease (MDRD) Study is the most accurate formula for the prediction of glomerular filtration rate (GFR) from serum creatinine. An alternative approach has been developed using evolving connectionist systems (ECOS), which are novel computing structures that can be trained to generate accurate output from a given set of input variables. This study aims to compare the prediction errors associated with each method, using data that reproduce routine clinical practice as opposed to the artificial setting of clinical trials.* ***Methods.*** *The methods were compared using 441 radioisotope measurements of GFR in 178 chronic kidney disease patients from 12 centers in Australia and New Zealand. All clinical and laboratory measurements were obtained from the patients' center rather than central laboratories, as would be the case in routine clinical practice. Both the MDRD formula and ECOS used the same predictive variables, and both were optimized to the study cohort by stepwise regression and training, respectively.* ***Results.*** *Mean measured GFR in the cohort was 22.6 mL/min/1.73 m^2. The bias and precision of the MDRD formula were -3.5 mL/min/1.73 m^2 and 34.5%, respectively, improving to -1.2 mL/min/1.73 m^2 and 31.1% after maximal optimization of the formula to study data. The bias and precision of the ECOS were 0.7 mL/min/1.73 m^2 and 32.6%, respectively, improving to -0.1 mL/min/1.73 m^2 and 16.6% after maximal optimization of the system to study data. The prediction of GFR using ECOS was improved by accounting for the center from where clinical and laboratory measurements originated within the connectionist model.* ***Conclusion.*** *Algebraic formulas will be associated with greater prediction error in routine clinical practice than in the original trials, and machine intelligence is more likely to predict GFR accurately in this setting.*

Keywords: glomerular filtration rate, creatinine clearance, artificial intelligence, connectionist systems

1. BACKGROUND

Accurate evaluation of renal function is fundamental to sound nephrologic practice. Early detection of renal disease allows for the institution of appropriate diagnostic and therapeutic measures, and potentially maximizes preservation of intact nephrons. As important, the evaluation of greater degrees of renal disease is useful to provide additional information to determine the onset of end-stage renal failure and facilitate the timely initiation of dialysis.

Glomerular filtration rate (GFR) is recognized as the best index of renal function[1]. GFR is most accurately measured by the renal clearance of inulin during continuous infusion. This procedure is technically demanding, and alternative methods using radioisotope tracers produce results of comparable accuracy, and have now become the gold standard for clinical research[2]. These are still too cumbersome and costly for routine clinical use, and most clinicians over the years have relied upon the clearance of creatinine as a convenient and inexpensive surrogate for GFR despite methodologic and systematic inaccuracies.

More recently, formulas developed for the Modification of Diet in Renal Disease (MDRD) Study have been shown to predict GFR from serum creatinine with greater accuracy than creatinine clearance[3]. In the original study, over 90% of predicted values by equation 7 were within 30% of measured GFR. These formulas, however, have not been rigorously evaluated in a manner that would reflect their use in routine clinical practice, namely in a clinically diverse patient cohort using data measurements originating from their respective centers. This is especially applicable for patients with GFR <30 mL/min/1.73 m^2, where data have been relatively scarce and studies conflicting as to whether any formulas are sufficiently accurate to support good clinical decision-making in the predialysis setting[4,5].

Artificial neural networks are an alternative approach to algebraic formulas for problem-solving in medical research and routine clinical practice[6]. These are computing architectures that can be trained to generate accurate output from a given set of input variables. Artificial neural networks have several advantages in

comparison to traditional statistical models and the algebraic formulas derived from them. Noisy data are well tolerated, and no limiting assumptions regarding distribution are required. They outperform classic predictive tools in situations where input variables are interrelated, and are increasingly used for nonlinear modeling dealing with complex and chaotic problems[7].

Evolving connectionist systems (ECOS) are yet a further advance in computing architecture. They allow a departure from the usual paradigm of so-called "global models," in which statistical functions or mathematical equations are developed and applied uniformly to the entire "problem space." A typical example of a global model is the MDRD formula, which implicitly assumes that relationships between predictive variables and GFR are the same for every patient within a given cohort. In contrast, ECOS involve a framework of multiple so-called "local models," in which different statistical models or mathematical equations are developed and applied in different clusters within the problem space. ECOS can be provided with a self-mapping function by which new data are allocated to whichever cluster or clusters are closest in terms of the associated predictive variables. From there, the local models unique to the allocated cluster or clusters are applied to the new data in a weighted fashion.

When training data are provided to such an ECOS (that is, containing both predictive variables *and* the output), then the ECOS will optimize both clustering and local models within the clusters until error cannot be reduced further, so-called "adaptive modeling"[8,9].

There are two potential benefits to this ECOS framework for medical applications. First, new training patient data will result in incremental and autonomous machine learning in a rapid continuous manner through on-line changes to the connectionist structure and function, without the necessity for down-time or complete de novo system retraining on a new enhanced data set. Second, the use of multiple local models has the potential for less predictive error than global models, by optimizing accuracy within each patient subset of the total cohort rather than relying on the application of a single model or equation designed to provide the greatest accuracy to the greatest number. ECOS appear particularly suitable for the prediction of GFR in chronic kidney disease, where the complex interrelation of patient factors and markers for GFR make the estimation of renal function very difficult. This article has two aims: (*1*) to evaluate the accuracy of algebraic formulas for the prediction of GFR across a range of centers using center-specific data measurements, and (*2*) to compare the performance of ECOS using local modeling with these formulas on the same data set.

2. METHODS

Study design

Reference GFR measurements by radioisotope tracer clearance were compared with predicted GFR values by alternative methods in a sample of patients from Australia and New Zealand. The predictions were based upon clinical and laboratory data from the day of GFR measurement, using the algebraic formulas and the ECOS as detailed below. The evolving nature of the latter method was evaluated through the degree of improvement in ECOS performance using progressively more complex testing protocols in several discrete ECOS modeling phases.

Data source

The EPO AUS-14 study was a prospective multicenter randomized study conducted from 1998 to 2002 to determine if maintenance of serum hemoglobin between 120 and 130 g/L prevented and/or delayed the development of left ventricular hypertrophy in patients with advanced kidney disease. The coordinating center did the original selection of 12 centers in Australia and New Zealand, and all incident patients fulfilling the criteria for study were screened for participation. These criteria were (*1*) age between 18 and 75 years, (*2*) GFR between 15 and 50 mL/min, and (*3*) demonstrated historic decline in hemoglobin concentration to 110 to 130 g/L for males and 100 to 120 g/L for females. Full details of the methods and results of the study have been reported elsewhere[10]. EPO AUS-14 was approved by ethical review committees at respective institutions and informed consent was obtained from all patients in accordance with the guidelines proposed in the Declaration of Helsinki[11].

Patients

A sample of patients was drawn from EPO AUS-14 for this study. In the original study, 296 patients were consented and screened for randomization. We excluded patients from study if the date of GFR measurement by the reference method did not coincide with the date of laboratory testing, or if the protocol employed for this GFR measurement differed from that stated below. A total of 178 patients from the original cohort were included in this study. The demographic and clinical characteristics of these patients are provided in Table 1.

Measurement of GFR by chromium-51-ethylenediaminetetraacetic acid (51Cr-EDTA) clearance

Reference GFR measurements were made for all patients at baseline and then yearly intervals for the duration of the study. GFR was measured as the plasma clearance of ^{51}Cr-EDTA corrected for body surface area (GFR^{EDTA}). Clearance was determined by either two or three point sampling at variable intervals between 0.5 and 4.5 hours after tracer injection, with or without a correction for the monoexponential assumption. Samples were processed in the nuclear medicine laboratories in each of the respective centers. Median intratest and intertest coefficients of variation within and between these centres were not studied and are therefore unavailable. A total of 441 GFR^{EDTA} measurements were available for this study, with an average of 2.5 measurements per patient (range 1 to 4).

Parameter	Number (% of total)	Mean	Standard Deviation
Number of patients	178	—	—
Number of gomerular filtration rate measurements	441	—	—
Male	93 (52%)	—	—
Female	85 (48%)	—	—
White race	160 (89%)	—	—
Asian race	3 (2%)	—	—
Black race	15 (9%)	—	—
Angiotensin-converting enzyme inhibitor use	130 (74%)	—	—
3-hydroxy-3-methylglutaryl coenzymeA reductase inhibitor use	69 (39%)	—	—
Loop diuretic use	76 (43.3%)	—	—
Diabetes mellitus	47 (26.4%)	—	—
Hypertension	165 (93.0%)	—	—
Congestive heart failure	4 (2.2%)	—	—
Ischemic heart disease	30 (16.9%)	—	—
Age		53.2	13.7
Weight *kg*		77.1	15.9
Height *cm*		167.8	9.6
Systolic blood pressure		141.2	21.5
Diastolic blood pressure		79.5	11.7
Serum creatinine *mmol/L*		0.35	0.21
Serum urea *mmol/L*		20.4	7.6
Serum albumin *g/L*		38.8	4.7
Hemoglobin		111.7	10.0
Left ventricular ejection fraction		65	7.4
Left ventricular mass *g*		162.8	57.7
Left ventricular mass index g/m^2		87.0	26.7

Table 1. Baseline clinical characteristics of patients

Prediction of GFR from serum creatinine by algebraic formulas

GFR were predicted using the following algebraic formulas, with clinical and laboratory input variables obtained on the day of GFR^{EDTA} testing: (*1*) equation 7 as described by the MDRD investigators in the original article[3] (analytes other than serum albumin in mg/dL, serum albumin in g/dL, age in years):

$$\begin{aligned} GFR^{MDRD} = {} & 170 \times (\text{serum creatinine})^{-0.999} \\ & \times (\text{age})^{-0.167} \times 0.762\,(\text{if sex is female}) \\ & \times 1.180\,(\text{if race is black}) \\ & \times (\text{serum urea nitrogen})^{-0.170} \\ & \times (\text{serum albu min})^{+0.318} \end{aligned}$$

and (*2*) modified MDRD equation containing the same variables, but different regression coefficients and multiplicative constants developed using multiple regression analyses on the EPO AUS-14 data set (GFR^{mMDRD}). Successive GFR^{mMDRD} equations were derived for each of the ECOS modeling phases described below.

The rationale for modifying the original MDRD equation 7 is as follows: equation 7 was developed in a sample of the United States population to predict GFR as measured by renal clearance of ^{125}I-iothalamate. In this study, patients were sampled from an Australian and New Zealand population and GFR was measured by plasma clearance of ^{51}Cr-EDTA. The original MDRD equation 7 cannot therefore be expected to perform as well in this data set as the original, due to patient related factors and also the intertest variability between the two techniques for radioisotope GFR measurement. A meaningful comparison between the MDRD equation and the ECOS developed in the new data set requires that the original MDRD equation 7 be remodeled to optimize accuracy under the new conditions. The modified MDRD equation and ECOS will therefore be products of the same data set, and neither will be disadvantaged by being developed under one set of conditions and tested under another.

Prediction of GFR by ECOS

GFR were predicted using the dynamic evolving neuro-fuzzy inference system (DENFIS)[9], an ECOS that optimizes its generated output by learning from training data using multiple local models. In this case, the generated output was GFR^{DENFIS}, and the training data were comprised of the target output (GFR^{EDTA}) and the clinical and laboratory variables to be associated with this target output and therefore to be used for computational modeling. DENFIS was engineered to report GFR^{DENFIS} as the average of ten internal modeling experiments for both training and testing data. Background information on DENFIS structure and function is available on the website http://gfr-ecos.kedri.org.

Modeling phases for ECOS and algebraic formulas

Three phases of modeling were performed. The purpose of the first phase of modeling was conventional validation of both the ECOS and modified MDRD formula. Variables used for the training of DENFIS and modification of the MDRD formula were the same six as were used in the original MDRD equation 7. The EPO AUS-14 data set was randomly divided into training and testing subdata sets, comprising 70% (309 renal function measurements) and 30% (132 renal function

measurements), respectively, of the total. The GFR^{mMDRD} equation was derived from the training data set using stepwise multiple regression analyses. The model for GFR^{DENFIS} was derived from the training data set within the ECOS as previously described[9].

The purpose of the second phase of modeling was to evaluate the effect of adaptive properties of DENFIS in clinical practice. As previously, the EPO AUS-14 data set was randomly divided into training and testing subdata sets, but this time comprising 80% (353 renal function measurements) and 20% (88 renal function measurements), respectively, of the total. The GFR^{mMDRD} equation was again derived from the training dataset using stepwise multiple regression analyses and the usual six variables. The modeling of DENFIS was performed in a manner to closest reproduce its use in clinical practice. The likeliest clinical scenario is that centers would be sequentially recruited to the ECOS over time, to join other centers already using the trained system. The recruitment of the new center would involve provision of some center-specific training data to the ECOS, after which one could expect accurate prediction of GFR for the new patients.

The "leave one out" method is the modeling protocol that best reflects this clinical scenario. This protocol involved dividing the EPO AUS-14 dataset into 12 subdata sets according to the center of origin of the renal function measurement. For a given center of interest, GFR^{DENFIS} was initially modeled by training on the other 11 centers. GFR^{DENFIS} was then further modeled in the center of interest by retraining on a random sample comprising 80% of renal function measurements from that center. This protocol was applied for each of the 12 centers. The overall prediction error was then calculated as the average error across the 12 centers from testing in the remaining 20% of the measurements from each center. This modeling protocol provides the most realistic reflection of ECOS performance with sequential recruitment of centers to the system over time.

The purpose of the third and final phase of modeling was to develop the most accurate ECOS and algebraic formula possible, and compare the limits of optimization for both frameworks. It should be noted that virtually all algebraic formulas in common clinical use, including the original MDRD equation 7 as published, have been optimized by using the entire respective data sets for concurrent training and testing[3,12,13,14,15,16,17]. The third phase of modeling in this study was similarly undertaken using the entire EPO AUS-14 data set for both training and testing of both the ECOS and algebraic formula. The variables used in the training of DENFIS and modification of the MDRD formula were the same six as were used in the original MDRD equation 7.

Modeling for both algebraic formulas and ECOS was performed using Matlab® version 6 software (Natick MA, USA).

Statistical analysis

The accuracy of predicted GFR values (GFR^{MDRD}, GFR^{mMDRD}, and GFR^{DENFIS}) was determined by their bias and precision in relation to reference GFR measurements (GFR^{EDTA}). Absolute agreement or bias was assessed by the mean difference between the predicted GFR values and GFR^{EDTA}, which is the systematic difference between the methods. Relative agreement or precision was assessed by the fluctuation of these differences around the mean. The standard deviation of these differences can be quantified as the root mean square error (RMSE), which can be expressed in mL/min/1.73 m^2 or as a percentage of GFR. The Bland-Altman procedure was also used which defines range of agreement. This is the mean difference ± 1.96 standard deviations, and represents how far apart predicted GFR values are likely to be from reference GFR measurements for 95% of cases[18,19]. Analyses were made using Analyze-It® version 1.62 software (Leeds, UK), and presented as scatter and bias plots.

3. RESULTS

Results are presented as mean ±standard deviation (range) unless otherwise specified. GFR^{EDTA} in the cohort was 22.6 ±10.7 (0.2 to 70) mL/min/1.73 m^2. GFR^{MDRD} was 19.1 ±9.3 (3.3 to 46.9) mL/min/1.73 m^2. GFR^{mMDRD} was 21.0 ±8.0 (4.2 to 40.8) mL/min/1.73 m^2 after the first phase of modeling, 22.3 ±8.0 (3.0 to 45.4) mL/min/1.73 m^2 after the second, and 21.4 ±7.8 (6.4 to 41.2) mL/min/1.73 m^2 after the third. The modified MDRD formula for Australians and New Zealanders generated using the entire EPO AUS-14 data set from the third phase of modeling was (analytes other than serum albumin in mg/dL, serum albumin in g/dL, and age in years):

$$\begin{aligned} GFR^{mMDRD} = {} & 120.4 \times (\text{serum creatinine})^{-0.825} \\ & \times (\text{age})^{-0.159} \times 0.837\,(\text{if sex is female}) \\ & \times 0.913\,(\text{if race is black}) \\ & \times (\text{serum urea nitrogen})^{-0.0114} \\ & \times (\text{serum albu min})^{+0.0651} \end{aligned}$$

GFR^{DENFIS} was 23.2 ±8.6 (5.0 to 47.6) mL/min/1.73 m^2 after the first phase of modeling, 22.6 ±8.7 (0.0 – 48.7) mL/min/1.73 m^2 after the second, and 22.5 ±9.9 (5.0 to 64.6) mL/min/1.73 m^2 after the third.

Statistical assessments of bias and precision of predicted GFR values are presented in Table 2 and Figures 1 to 4. The prediction error of GFR^{DENFIS} versus GFR^{mMDRD} from the second phase of modeling for each of the 12 centers is shown in Figure 5. It can be seen that the ECOS outperformed the algebraic formula in only certain centers. This finding can be further explored considering Center 2 as a case study. Patients from Center 2 had a marginally higher serum creatinine (0.40 ±0.10 mmol/L) but a markedly lower GFR^{EDTA} (12.1 ±6.7 mL/min/1.73 m^2) when compared to the other centers. The relationship between these two variables was therefore different in patients from Center 2, explaining the improved prediction with local modelling via DENFIS in comparison to global modeling via the GFR^{mMDRD}. There are several possible hypotheses to explain this

observation. Perhaps the patients from Center 2 were biologically different with lower rates of creatinine production. Indeed, patients from Center 2 did tend to be female (60% of patients), older (mean age 60 years), and none were black. Alternatively, laboratory assays for serum creatinine or measurements of GFR^{EDTA} may be systematically lower in Center 2 than other centers. Irrespective of the reason, improved ECOS performance in this second phase of modeling is due to additional clustering and local model optimization, and allows for improved prediction for patients within centers by accounting for such center disposition.

Versus GFR^{EDTA}	Bias [95% CI] *mL/min/1.73 m²*	RMSE *mL/min/1.73 m²*	RMSE *% of GFR*	95% limits of agreement *mL/min/1.73 m²*	
				Lower	Upper
GFR^{MDRD}	−3.5 [−4.2, −2.9]	7.75	34.5%	−17.2	10.1
Modeling phase 1					
GFR^{mMDRD}	−1.6 [−2.3, −0.9]	7.59	33.6%	−16.1	13.0
GFR^{DENFIS}	0.7 [0.0. 1.3]	7.36	32.6%	−13.7	15.0
Modeling phase 2					
GFR^{mMDRD}	−0.3 [−0.9, 0.4]	7.08	31.3%	−14.2	13.6
GFR^{DENFIS}	0.1 [−0.6, 0.6]	6.75	29.9%	−13.2	13.3
Modeling phase 3					
GFR^{mMDRD}	−1.2 [−1.8, −0.6]	7.03	31.1%	−14.8	12.4
GFR^{DENFIS}	−0.1 [−0.4, 0.3]	3.73	16.6%	−7.4	7.2

RMSE is root mean square error, CI, confidence interval. Accuracy is reported for the testing subdata sets for the modeling phases 1 and 2 and for the entire data set for GFR^{MDRD} and modeling phase 3.

Table 2. Agreement between predicted glomerular filtration rate (GFR) values and reference GFR measurements

Fig. 1. Agreement of glomerular filtration rate according to the Modification of Diet in Renal Disease formula (GFR^{MDRD}) with ethylenediaminetetraacetic acid (GFR^{EDTA}) clearance (mL/min/1.73 m²). In the scatter plot (*A*), the dotted line (·····) represents the line of identity between methods. In the bias plot (*B*), dotted lines represent the bias between methods, broken lines (- ·– ·) the range of agreement, and the solid line the line of regression indicating bias according to level of GFR.

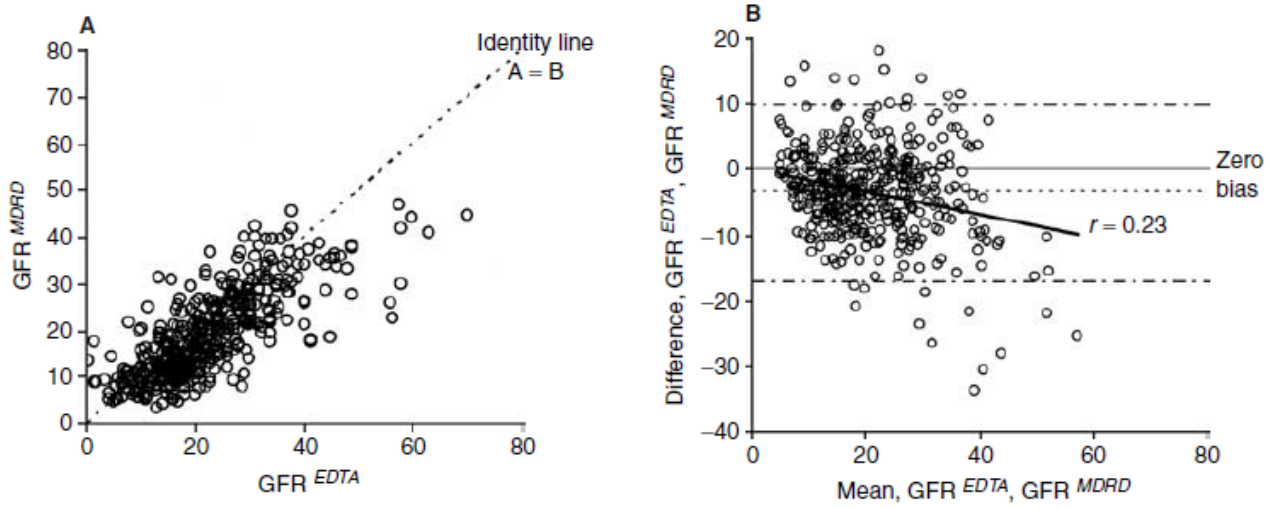


Figure 6 shows the ECOS interface, with one of the fuzzy rules generated by the trained DENFIS. Each rule represents a local model associating predictive variables with the generated output within a given cluster. All rules together represent the equivalent of a global model that can be applied for the prediction of GFR for any new patient.

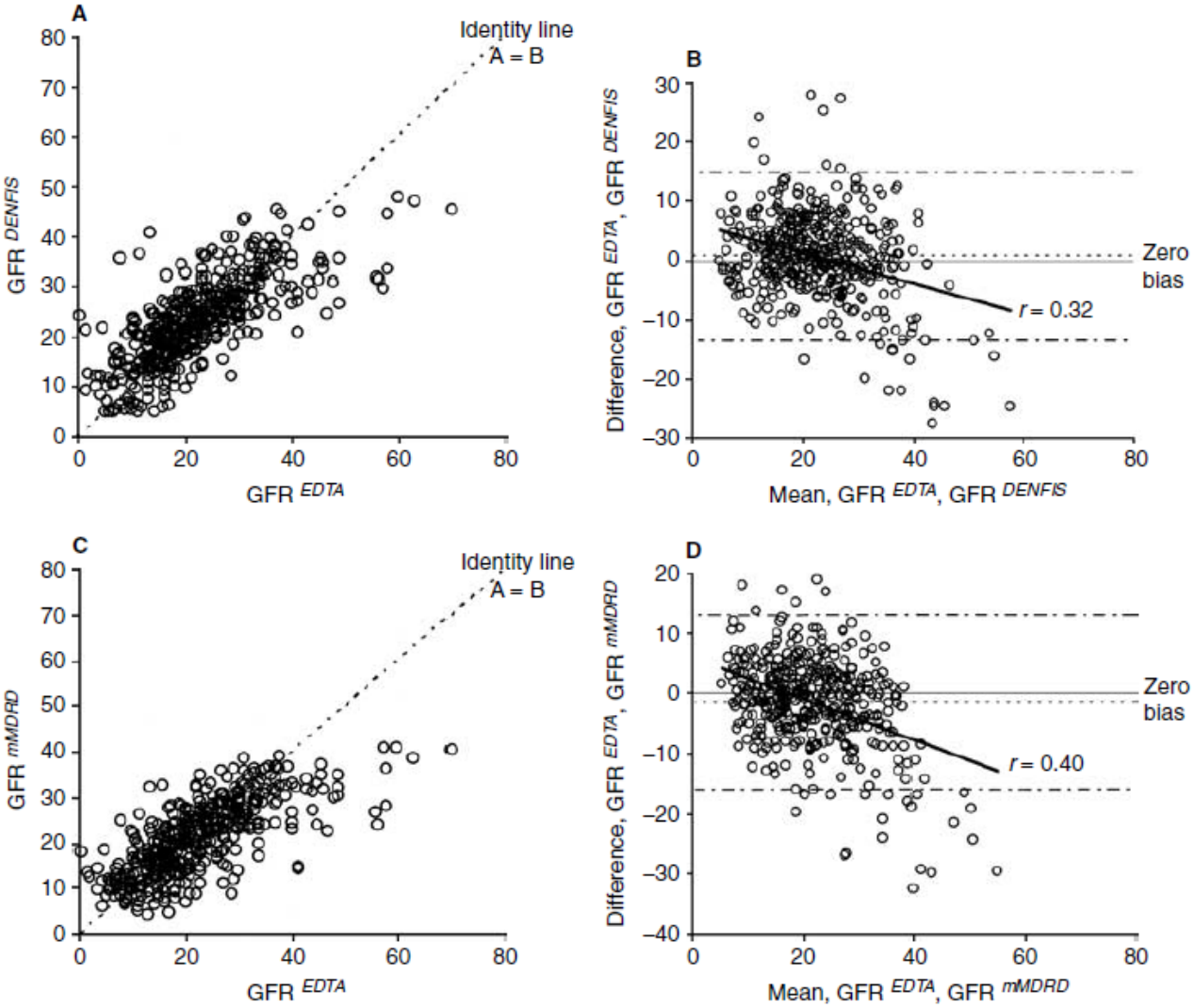
4. DISCUSSION

There seems little doubt that most clinicians will continue to rely on estimates of renal function from serum creatinine to assist with clinical decision making, and an array of algebraic formulas have been developed using regression techniques to predict GFR from standard clinical variables[20]. In the setting of clinical trials, the most accurate of these formulas are those from the MDRD Study[21,22,23].

The data from this study indicate that these formulas will be less accurate than expected in routine clinical practice. However, modification of the original MDRD equation 7 by multiple regression analyses within the EPO AUS-14 data set did achieve some improvement in absolute prediction error (bias) from -3.5 to -1.2 mL/min/1.73 m^2, and in relative prediction error (precision) from 34.5% to 31.1%. This represents the best accuracy that can be achieved in the study cohort by an algebraic formula

using the MDRD equation 7 template. It should be noted that the modified MDRD formulas generated for these analyses are not suitable for clinical use, as this study constitutes insufficient validation in terms of (*1*) patient numbers and (*2*) the range of GFR measurements over which the modified formulas were tested.

Fig. 2. Agreement of glomerular filtration rate according to the modified Modification of Diet in Renal Disease formula (GFR^{mMDRD}) and dynamic evolving neuro-fuzzy inference system (GFR^{DENFIS}) with ethylenediaminetetraacetic acid (GFR^{EDTA}) clearance (mL/min/1.73 m^2) from modeling phase 1. In the scatter plots (*A*, *C*), the dotted lines (·····) represent the line of identity between methods. In the bias plots (*B*, *D*), dotted lines represent the bias between methods, broken lines (- ·– ·) the range of agreement, and solid lines the line of regression indicating bias according to level of GFR.

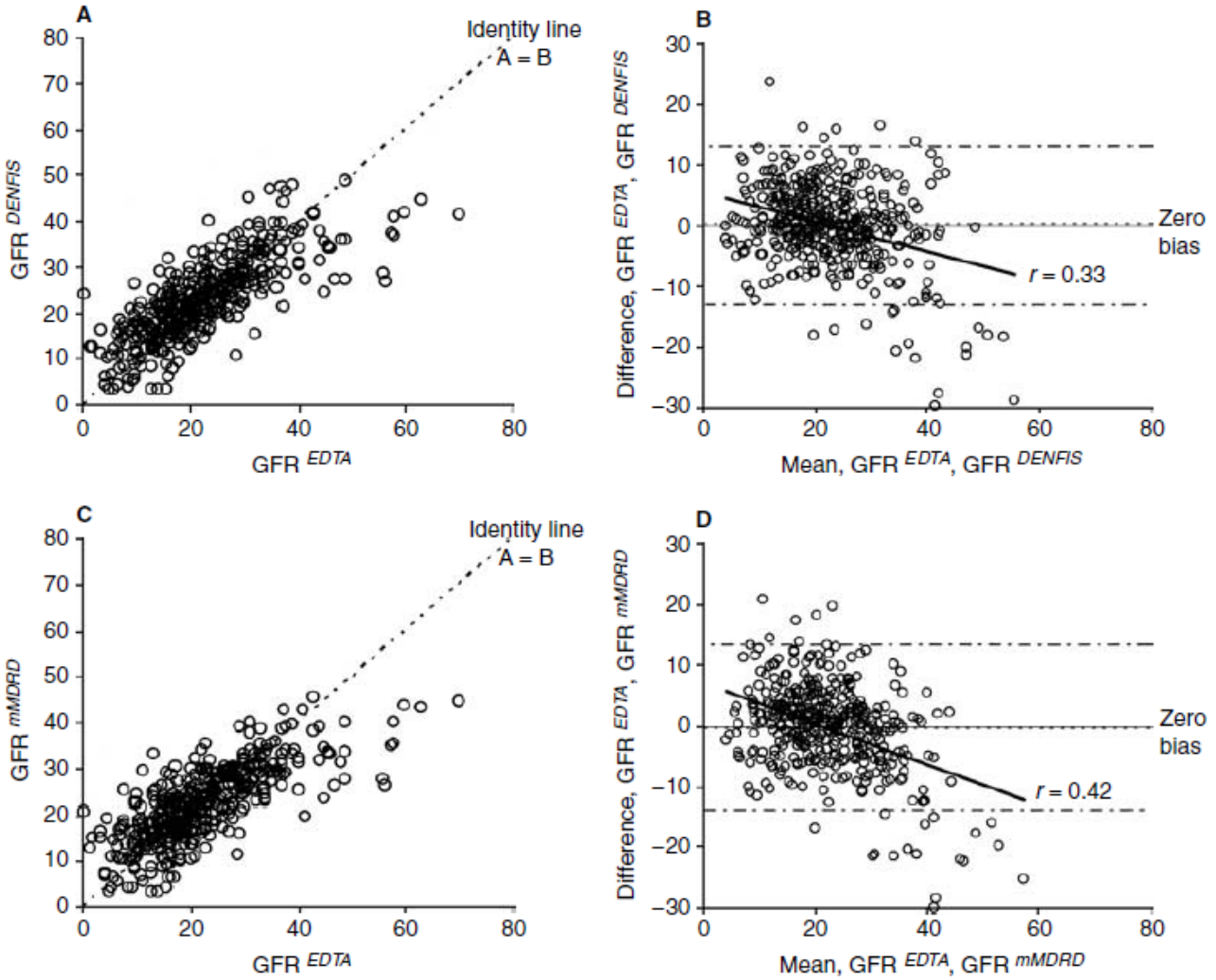


A fundamental methodologic feature of this study is the prediction of GFR in multiple centers using center-specific clinical and laboratory measurements. This study design reproduces the use of these formulas in routine clinical practice. To our knowledge, there are no similar studies in the literature for comparison. Previous studies have usually been undertaken in single centers, or in multiple centers but using central laboratories. The prediction error from each of these individual studies cannot be simply averaged for comparison with the data presented here, although raw data could be pooled and reanalyzed.

There are several factors that led to the greater than expected bias and imprecision of algebraic formulas in this study. The minimum prediction error for GFR that might be achievable by any method will be no less than the measurement error of the reference method. The use of central laboratories for the reference GFR measurement reduces error since it is dependent on intra-test error only (variation between the reported clearances of two forms of the same marker administered to the same patient simultaneously). In this study and also in routine clinical practice, the variety of reference methods that might be used in different laboratories leads to additional interest error (variation between the reported clearances of different marker standards administered to the same

patient simultaneously). There are few published data that definitively quantify these errors for GFR measurement, although the most widely quoted estimates are ~5% to 10%[2,24,25].

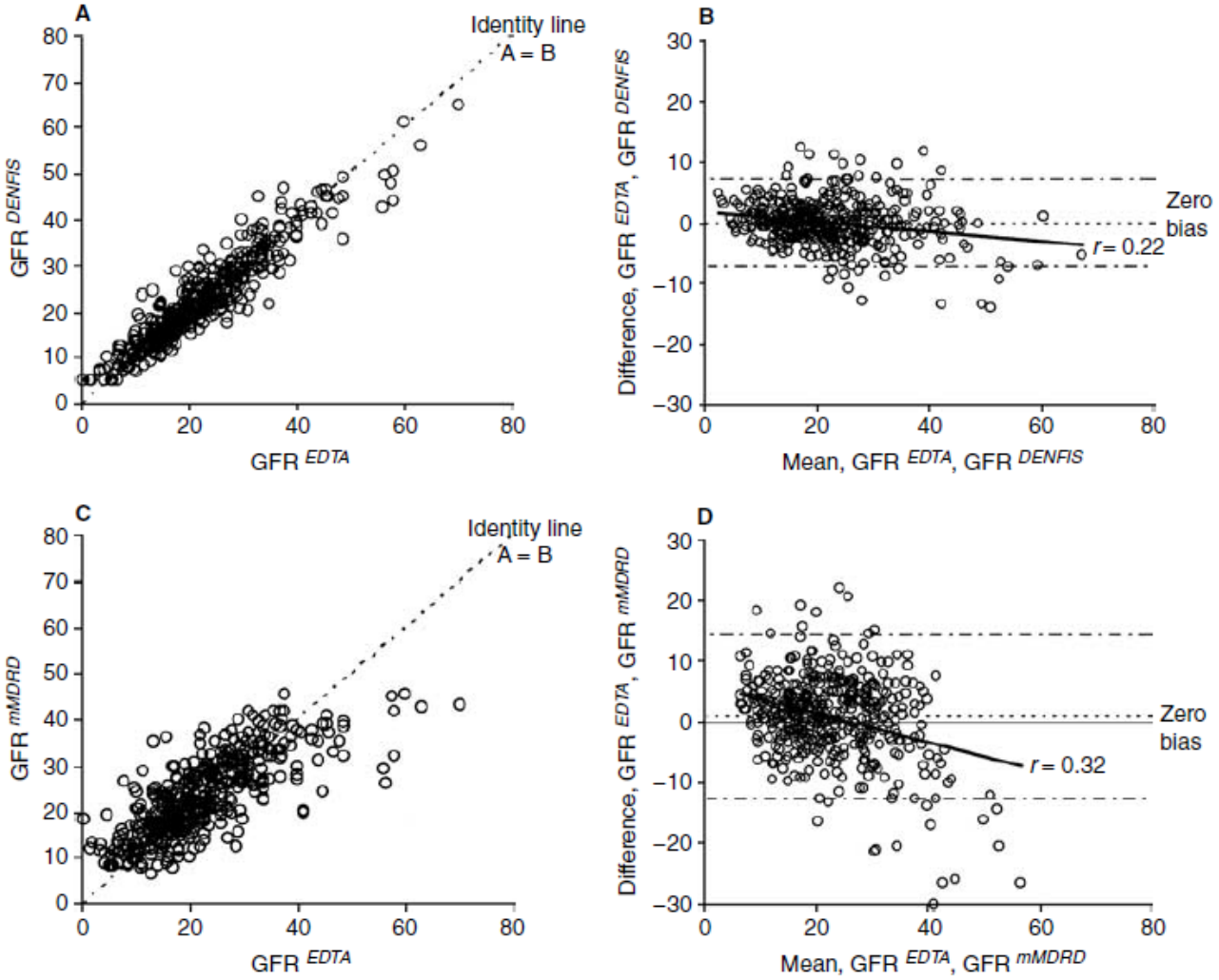
Fig. 3. Agreement of glomerular filtration rate according to the modified Modification of Diet in Renal Disease formula (GFR^{mMDRD}) and dynamic evolving neuro-fuzzy inference system (GFR^{DENFIS}) with ethylenediaminetetraacetic acid (GFR^{EDTA}) clearance (mL/min/1.73 m^2) from modeling phase 2. In the scatter plots (*A*, *C*), the dotted lines (·····) represent the line of identity between methods. In the bias plots (*B*, *D*), dotted lines represent the bias between methods, broken lines (- ·– ·) the range of agreement, and solid lines the line of regression indicating bias according to level of GFR.



Another factor to consider is that the MDRD formulas were developed and validated within a study cohort where only 3% had diabetes mellitus, and validated further in the African American Study of Hypertension and Kidney Disease study cohort where 100% were African Americans and 0% had diabetes mellitus[26]. Similarly unrepresentative patient samples have been used for the development and validation of other popular formulas, such as that by Cockcroft and Gault[14]. It is possible and even likely that different ethnic populations such as Asians, Hispanics, and Polynesians and also patients with different comorbid medical burden will have different biologic and therefore algebraic relationships between GFR and its predictive variables such as serum creatinine.

Finally, analytical differences between laboratories in measurement of analytes such as serum creatinine contribute to the prediction error of these algebraic formulas. Interlaboratory error has been evaluated in a number of studies and may be as high as 15%, as a result of calibration differences to a great extent. It should be noted that this error has progressively less impact with decreasing GFR, and may be clinically insignificant in the predialysis setting[27,28]. Minimization of this error in the routine clinical practice is possible through improved calibration in laboratories to a single external standard, although there will still be some error attributable to laboratory methodology until it becomes more precise.

Fig. 4. Agreement of glomerular filtration rate according to the modified Modification of Diet in Renal Disease formula (GFR^{mMDRD}) and dynamic evolving neuro-fuzzy inference system (GFR^{DENFIS}) with ethylenediaminetetraacetic acid (GFR^{EDTA}) clearance (mL/min/1.73 m^2) from modeling phase 3. In the scatter plots (*A*, *C*), the dotted lines (·····) represent the line of identity between methods. In the bias plots (*B*, *D*), dotted lines represent the bias between methods, broken lines (- ·– ·) the range of agreement, and solid lines the line of regression indicating bias according to level of GFR.



Given the difficulties in predicting GFR from serum creatinine, should the clinician abandon such methods and rely exclusively on the measurement of tracer clearance? Radioisotope methods are accurate, and are known to have a high degree of reproducibility across different centers as long as computing algorithms are similar[29,30,31,32]. These methods have been recommended as routine by some authors[33], although their expense and logistics limit utilization in most centers. The most accurate and clinically accessible methods for measurement of GFR by tracer clearance are first measured creatinine clearance with cimetidine blockade[1], and second averaged measured urea and creatinine clearance[3]. The main disadvantage of these methods is the requirement for accurate 24-hour urine collection, and the potential for collection error which can result in spurious day to day variation in GFR of up to 70%[1].

Fig. 5. Prediction error for glomerular filtration rate according to the dynamic evolving neuro-fuzzy inference system (GFR^{DENFIS}) (□) and the modified Modification of Diet in Renal Disease (GFR^{mMDRD}) (▲) in each of the 12 centers from modeling phase 2.

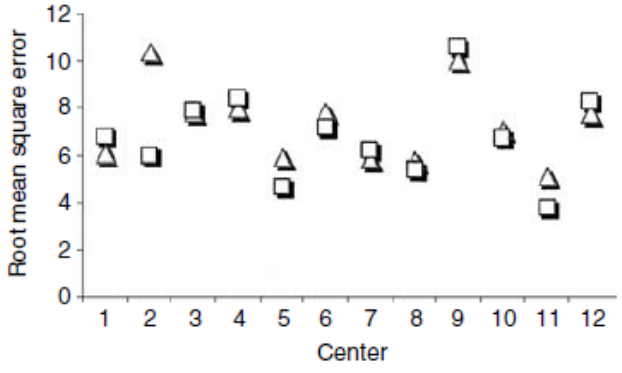
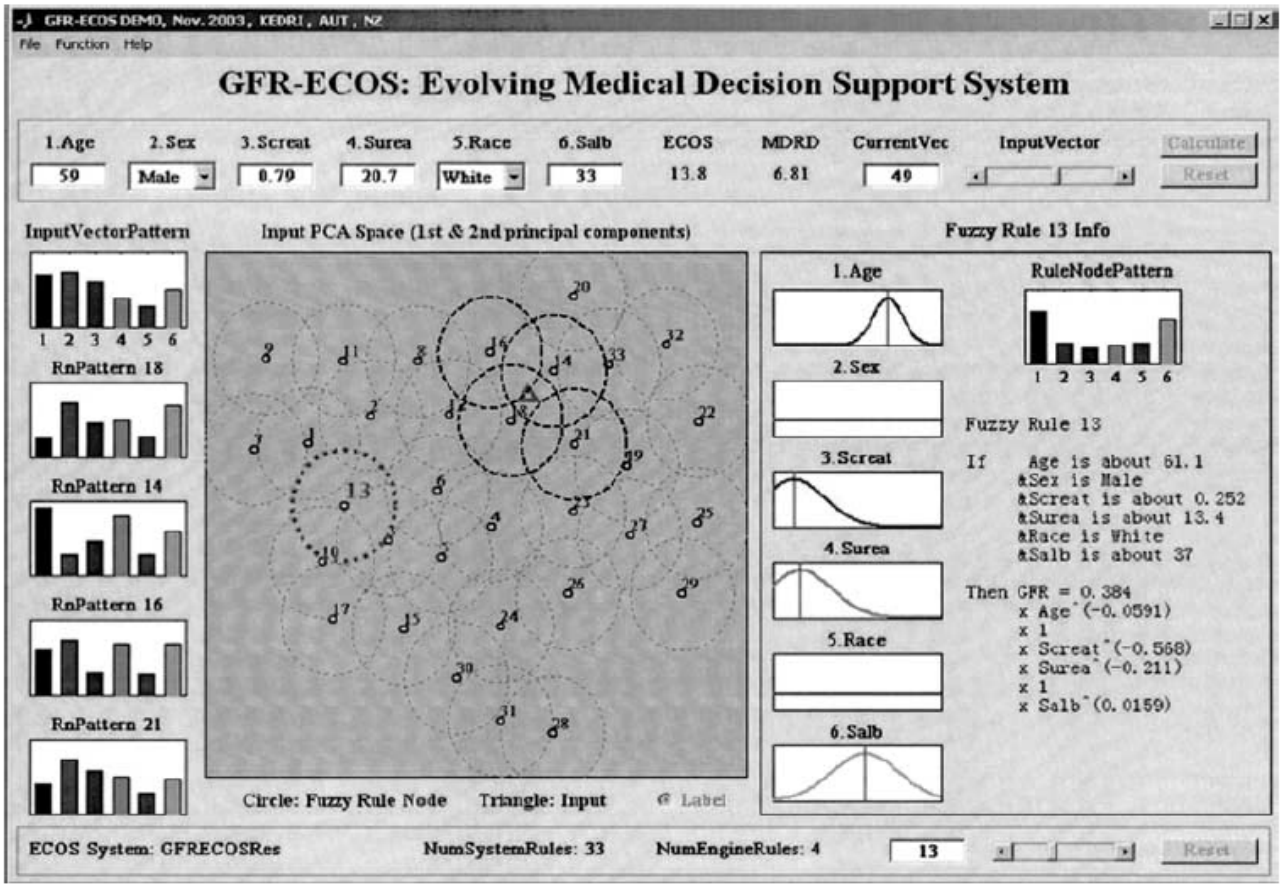


Fig. 6. Illustration of one of the dynamic evolving neuro-fuzzy inference system (DENFIS) computer interfaces. The problem space is visualized, as is the progressive partitioning of the space for the ongoing creation of fuzzy rules. At each moment, GFR^{DENFIS} is calculated through a fuzzy inference system based on the most activated fuzzy rules that are dynamically selected from the existing fuzzy rule set. New fuzzy rules are created and updated during the operation of the system. As an example, rule 13 is illustrated in the interface. Note that input variables are normalized between zero and one. The entire complement of rules for the modeling of GFR^{DENFIS} (modeling phase 3) are provided in the Web site http://gfr-ecos.kedri.org.



In this article, we present a genuine alternative to all these methods in a convenient and inexpensive form of machine intelligence. ECOS has been shown to predict GFR with greater accuracy than what is regarded as the best of the available algebraic equations, even when the latter was also optimized using conventional statistical modeling within the EPO AUS-14 data set. Moreover, the second phase of modeling in our study illustrates the potential beneficial of adaptive modeling with sequential recruitment of centers to the system. In our study, such improvement, however, was not achieved through the development of discrete models for each center. Within each center, up to 21 models were used to calculate GFR^{DENFIS}, with renal function measurements often allocated to areas of the problem space partitioned to several overlapping clusters. It is the weighted application of these local models within the ECOS framework that provides the clinical benefit over and above global models such as algebraic formulas.

The second phase of modeling also demonstrated an important limitation of ECOS in clinical practice. In the case study of Center 2, the ECOS was unable to distinguish whether the discrepancy between GFR^{mMDRD} and the corresponding GFR^{EDTA} arose from patient-related factors or measurement error in laboratory parameters or radioisotope tracer clearance. ECOS is still a tool based on association rather than causality. However, unlike conventional artificial neural networks it is still possible to examine relationships among input and output variables within the ECOS. The local models are in the form of fuzzy rules that can be extracted and studied. Such rules may allow for generation of hypotheses for further laboratory or clinical testing, and also have the potential to directly add to our understanding of underlying biologic processes.

A feature of all of our modified MDRD formulas is the factor of less than unity that is used to account for black race, as opposed to 1.18 in the original MDRD equation 7. In this study, the ethnic mix of "black" patients in Australia and New Zealand included Maori, Polynesian, and Aboriginal patients, who are quite distinct from African Americans. It should not be assumed that creatinine generation is higher in these ethnic groups as is the case in African Americans. In our study cohort, the mean GFR was 18.73 mL/min and the mean serum creatinine 0.42 mmol/L for patients classified as black. The corresponding parameters were 22.9 mL/min and 0.34 mmol/L for patients classified as white. If one equates GFR with creatinine clearance and assumes a steady state, the mean 24-hour creatinine generation is 11.3 mmol for both blacks and whites. The only other available published data support this finding. The

relationship between calculated creatinine clearance and urine creatinine was not different in Maoris, Polynesians, and Europeans (P.A. Metcalf, personal communication, July 26, 2004)[34]. This issue highlights again the potential in applying a formula such as the original MDRD equation 7 in a population that is different from that in which it has been developed.

The engineering of machine intelligence into tools of medical practice is not difficult. Many medical devices already have such systems embedded in them such as arrhythmia detectors. Alternatively, the systems can be placed on a central server as an internet or intranet-based utility. If such computing resources were not available, these systems are amenable to rule extraction as described. Such rules may be imported in a non-evolving form into a hand-held device, although they would need updating whenever advances in predictive modeling were made.

This study has two limitations. The first of these is that multiple GFR measurements were included for each patient. This methodology has occasionally been a feature of previous studies of this nature[17], although is undesirable as formulas derived from regression analyses will be biased toward patients with more frequent measurements. We have compared demographic, clinical, and laboratory characteristics of the patients in this study with one or two GFR measurements, versus those with three or four measurements. There were no demonstrable differences in any of these parameters (data not shown), indicating that the average frequency of 2.4 GFR measurements per patient was unlikely to have confounded our results. Ultimately, any limitation of the data set was the same for both the algebraic formulas and the ECOS, and the comparison of the two methods at the core of this study still valid. The other limitation of this study is its sample size: the MDRD study used 1070 and 558 GFR measurements for training and validation, respectively, compared with 309 and 132 corresponding GFR measurements in this study. This will inevitably limit the power of the analyses presented in this study, although we believe that study of a larger cohort would not have produced different results.

5. CONCLUSION

This study strongly suggests that published algebraic formulas for the prediction of GFR will be less accurate than expected in routine clinical practice and confirms that their performance can be improved somewhat by additional regression analyses prior to clinical use in diverse populations. This study demonstrates machine intelligence to be workable with greater accuracy than such algebraic formulas. Furthermore, there is potential to enhance modeling further within the ECOS framework by the sequential inclusion of further clinical variables with training data in the final model in the future. A Web-based implementation of GFR^{DENFIS} has been developed by this group for further prospective multicenter study, and it is hoped that the computational models so developed may in turn shed light upon biologic processes that influence renal function and mitigate renal disease.

ACKNOWLEGEMENTS

The authors would like to thank Janssen-Cilag Pty for their cooperation, and the investigators from the original EPO AUS-14 study for their collaboration and generosity. The authors would also like to thank Associate Professor John F. Collins for his helpful comments. This research is supported by the research grant NERFAUTX02001, funded by the Foundation for Research Science and Technology New Zealand, and by an internal research grant at the Middlemore Hospital in Auckland.

Reprint requests to Dr. Mark R. Marshall, FRACP, Department of Renal Medicine, Middlemore Hospital, Private Bag 93311, Auckland 1701, New Zealand. E-mail: mrmarshall@middlemore.co.nz

REFERENCES

1. WALSER M: Assessing renal function from creatinine measurements in adults with chronic renal failure. *Am J Kidney Dis* 32:23–31, 1998
2. LEVEY AS, GREENE T, SCHLUCHTER MD, *et al*: Glomerular filtration rate measurements in clinical trials. Modification of Diet in Renal Disease Study Group and the Diabetes Control and Complications Trial Research Group. *JAm Soc Nephrol* 4:1159–1171, 1993
3. LEVEY AS, BOSCH JP, LEWIS JB, *et al*: A 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* 130:461–470, 1999
4. BEDDHU S, SAMORE M, ROBERTS M, *et al*: Creatinine production, nutrition, and glomerular filtration rate estimation. *JAm Soc Nephrol* 14:1000–1005, 2003
5. LEVEY AS, GREENE T, BURKART J: Comprehensive assessment of the level of renal function at the initiation of dialysis in the MDRD study. *JAm Soc Nephrol* 9:153A, 1998
6. LARKIN M: Artificial intelligence slips cautiously into the clinic. *Lancet* 358:898, 2001
7. KASABOV N: *Foundations of Neural Networks, Fuzzy Systems and Knowledge Engineering*, Cambridge MA, MIT Press, 1996
8. KASABOV N: *Evolving Connectionist Systems—Methods and Applications in Bioinformatics, Brain Study and Intelligent Machines*, London-New York, Springer Verlag, 2002
9. KASABOV N, SONG Q: DENFIS: Dynamic, evolving neural-fuzzy inference systems and its application for time-series prediction. *IEEE Trans Fuzzy Sys* 10:144–154, 2002
10. ROGER S, MCMAHON L, CLARKSON A, *et al*: Impact of epoetin alfa on left ventricular mass in chronic kidney disease. *JAm Soc Nephrol* 15:148–156, 2003
11. 48TH WORLD MEDICAL ASSEMBLY: Declaration

of Helsinki: Recommendations guiding physicians in biomedical research involving human subjects. *JAMA* 277:925–926, 1997

12. HULL JH, HAK LJ, KOCH GG, *et al*: Influence of range of renal function and liver disease on predictability of creatinine clearance. *Clin Pharmacol Ther* 29:516–521, 1981

13. JELLIFFE RW: Estimation of creatinine clearance when urine cannot be collected. *Lancet* 1:975–976, 1971

14. COCKCROFT DW, GAULT MH: Prediction of creatinine clearance from serum creatinine. *Nephron* 16:31–41, 1976

15. MAWER GE, LUCAS SB, KNOWLES BR, *et al*: Computer-assisted prescribing of kanamycin for patients with renal insufficiency. *Lancet* 1:12–15, 1972

16. KAMPMANN J, SIERSBAEK-NIELSEN K, KRISTENSEN M, *et al*: Rapid evaluation of creatinine clearance. *Acta Med Scand* 196:517–520, 1974

17. WALSER M, DREW HH, GULDAN JL: Prediction of glomerular filtration rate from serum creatinine concentration in advanced chronic renal failure. *Kidney Int* 44:1145–1148, 1993

18. BLAND JM, ALTMAN DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307– 310, 1986

19. BLAND JM, ALTMAN DG: Comparing methods of measurement: Why plotting difference against standard method is misleading. *Lancet* 346:1085–1087, 1995

20. NATIONAL KIDNEY FOUNDATION KIDNEY DISEASE OUTCOME QUALITY INITIATIVE ADVISORY BOARD: K/DOQI clinical practice guidelines for chronic kidney disease: Evaluation, classification, and stratification. Kidney Disease Outcome Quality Initiative. *Am J Kidney Dis* 39 (Suppl 2):S1–S246, 2002

21. BEDROS F, KASISKE B: Estimating glomerular filtration rate from serum creatinine in renal transplant recipients. *JAm Soc Nephrol* 9:666A, 1998

22. LEWIS J, AGODOA L, CHEEK D, *et al*: Comparison of cross-sectional renal function measurements in African Americans with hypertensive nephrosclerosis and of primary formulas to estimate glomerular filtration rate. African-American Study of Hypertension and Kidney Disease. *Am J Kidney Dis* 38:744–753, 2001

23. LAMB EJ, WEBB MC, SIMPSON DE, *et al*: Estimation of glomerular filtration rate in older patients with chronic renal insufficiency: Is the modification of diet in renal disease formula an improvement? *JAm Geriatrics Soc* 51:1012–1017, 2003

24. WALSER M, DAVIDSON D, ORLOFF J: The renal clearance of alkali-stable inulin. *J Clin Invest* 24:1520–1523, 1955

25. PERRONE RD, STEINMAN TI, BECK GJ, *et al*: Utility of radioisotopic filtration markers in chronic renal insufficiency: Simultaneous comparison of 125I-iothalamate, 169Yb-DTPA, 99mTc-DTPA, and inulin. The Modification of Diet in Renal Disease Study. *Am J Kidney Dis* 16:224–235, 1990

26. LEWIS J, AGODOA L, CHEEK D, *et al*: Comparison of cross-sectional renal function measurements in African Americans with hypertensive nephrosclerosis and of primary formulas to estimate glomerular filtration rate. *Am J Kidney Dis* 38:744–753, 2001

27. ROSS JW, MILLER WG, MYERS GL, *et al*: The accuracy of laboratory measurements in clinical chemistry: A study of 11 routine chemistry analytes in the College of American Pathologists Chemistry Survey with fresh frozen serum, definitive methods, and reference methods. *Arch Pathol Lab Med* 122:587–608, 1998

28. CORESH J, ASTOR BC, MCQUILLAN G, *et al*: Calibration and random variation of the serum creatinine assay as critical elements of using equations to estimate glomerular filtration rate. *Am J Kidney Dis* 39:920–929, 2002

29. COSGRIFF PS, FLEMING JS, JARRITT PH, *et al*: UK audit of GFR measurment. *Nucl Med Commun* 23:286, 2002

30. WHITE DR, HOUSTON AS, SAMPSON WF, *et al*: Intra-and interoperator variations in region-of-interest drawing and their effect on the measurement of glomerular filtration rates. *Clin Nucl Med* 24:177–181, 1999

31. COSGRIFF PS, FLEMING JS, JARRITT PH, *et al*: UK audit of GFR measurement: An update. *Nucl Med Commun* 24:467, 2003

32. FLEMING JS, COSGRIFF PS, HOUSTON AS, *et al*: UK audit of relative renal function measurement using DMSA scintigraphy. *Nucl Med Commun* 19:989–997, 1998

33. MARIAT C, ALAMARTINE E, BARTHELEMY JC, *et al*: Assessing renal graft function in clinical trials: Can tests predicting glomerular filtration rate substitute for a reference method? *Kidney Int* 65:289–297, 2004

34. METCALF P, SCRAGG R, DRYSON E: Associations between body morphology and microalbuminuria in healthy middle-aged European, Maori and Pacific Island New Zealanders. *Int J Obes Relat Metab Disord* 21:203–210, 1997