Sonographic Features of Chronic Kidney Disease in Agricultural Community in Sri Lanka

Muditha S. Bandara¹, Buddika Gurunayaka², Gamage Pemanatha Lakraj³, Aruna Pallewatte⁴, Sisira Siribaddana⁵, Janaka Wansapura⁶,⁷

¹Department of Physics, University of Colombo, Colombo, ²Department of Radiology, Anuradhapura Teaching Hospital, Anuradhapura, ³Department of Statistic, University of Colombo, Colombo, ⁴Department of Neuroradiology, National Hospital of Sri Lanka, Colombo, ⁵Department of Medicine, Rajarata University of Sri Lanka, Anuradhapura, ⁶Department of Physics, University of Colombo, Colombo, Sri Lanka, ⁷Advanced Imaging Research Center, University of Texas Southwestern, Texas, USA.

INTRODUCTION

Chronic kidney disease (CKD) is a global health concern with increasing prevalence in most countries. Estimated glomerular filtration rate (eGFR), <60 ml/min/1.73 m², increased albumin excretion, or both lasting for more than 3 months, is defined as CKD. Hypertension, diabetes mellitus, and older age are among the classic risk factors of CKD. However, in the past few decades, a new form of CKD known as either CKDu or Chronic Interstitial Nephropathy among Agricultural Communities (CINAC) has emerged among paddy farmers in the dry zone of Sri Lanka. CINAC is defined as a form of CKD that affects mainly young men,
occasionally women. The etiology of CINAC is thought to be mainly toxicological, occupational, and environmental. Chemical compounds used in routine agricultural procedures such as pesticides and fertilizers including parquat, glyphosate, and methamidophos were identified as the toxicological contributors in CINAC. Further, the harsh agricultural employment conditions, drinking well water, family history of the disease, Ayurvedic treatments, and snake bites are considered as CINAC risk factors in Sri Lanka. CKDu/CINAC does not present with commonly known risk factors for CKD such as diabetes, hypertension, glomerulopathies, or renal stone disease. In 2013, Jayatilake et al. found the prevalence of CKDu/CINAC to be more than 15% in some districts of Sri Lanka. CINAC is now found to be in several tropical countries including El Salvador, Nicaragua, Guatemala, and Costa Rica in Central America, Sri Lanka, and India in Asia, and Egypt in Africa. The emergence of CINAC warrants a comparative study of the kidney morphological changes in both CKD with classic risk factors and CKD of unknown etiology (CINAC).

US imaging, considered the first-line diagnostic imaging procedure in CKD, provides several structural measurements of kidney, including length (LEN), width (WDTH), volume (VOL), and cortical thickness (THK). Although it is widely accepted that the kidneys shrink in size as the disease progress and that the kidney size correlates with anthropometric parameters, there is no consent among investigators on the kidney features that should be used to represent the kidney size. According to Lucisano et al., parenchymal THK and kidney LEN normalized to body height (HGHT) show the highest correlation with kidney function in CKD. However, in a larger population study of healthy kidneys, Emamian et al. found kidney VOL to be the most exact kidney measurement; with the highest correlation to the body surface area (BSA) (correlation coefficients = 0.595 and 0.600 for the left and right kidneys), whereas the normal kidney LEN correlated to the HGHT with correlation coefficients of 0.461 and 0.416 for the left and right kidneys. In a study of healthy subjects and CKD patients, Jovanović et al. found that while kidney morphological features (e.g., WDTH, depth, VOL, etc.) correlated well with anthropometric parameters in healthy subjects, only VOL correlated significantly with the anthropometric parameters in CKD patients.

Since existing studies do not consent to a single morphological feature and an appropriate correction for the body size, we analyzed feature vectors consisting of multitudes of morphological features normalized to different anthropometric parameters such as BSA, body mass index (BMI), HGHT, and waist circumference (CIRC). Feature vectors were constructed from the most influential normalized features that discriminate given two groups, based on a random forest (RF) procedure.

**RF classification**

RF is a classification and regression methodology based on the assembly of Classification and Regression Trees (CART) which can be used as a supervised learning algorithm. RF fits many decorrelated trees on bootstrapped training sets and takes the most commonly occurring class among the fitted trees as the overall prediction in the classification context. In addition, RF uses mean decrease accuracy as a metric to measure the importance of features in building a classifier. The importance of a feature is determined by how the accuracy decreases when the feature is excluded from the RF. Features with high mean decrease accuracy are the most influential predictors for the classifier.

**MATERIALS AND METHODS**

With the Institutional Ethics Review Committee approval, CKD patients (n = 79) undergoing routine US examination within a 6-month period at a regional hospital in Sri Lanka were randomly recruited for the study. All patients were clinically diagnosed with CKD according to the KDIGO criterion. The eGFR of patients, calculated based on the serum creatinine level, ranged from 14 ml/min/1.73 m\(^2\) to 87 ml/min/1.73 m\(^2\) with mean value of 49 ml/min/1.73 m\(^2\) at the time of the study. Patients with renal calculi, renal cysts, renal agenesis, and history of snake bites were excluded from the study. Total excluded patients were 29. Age matched healthy volunteers (n = 26) with no clinical symptoms and clinical history of kidney disease were recruited as a control group. Age, blood pressure, and anthropometric measurements (weight, HGHT, and waist CIRC) were collected from both CKD patients and healthy subjects before the US scan. Informed written consent was obtained before each examination from patients and healthy volunteers. As per literature, CKD patients diagnosed with hypertension, diabetes mellitus, or any other known cause were categorized as patients with known etiology (CKD\(_{\text{known}}\), n = 30) and the remaining kidney patients with no history of diabetes mellitus, chronic or severe hypertension, glomerular nephritis, or obstructive nephropathies were categorized as patients with unknown etiology (CKD\(_{\text{unknown}}\), n = 20).

Specialist radiologist performed US imaging. Toshiba Apio 500 (Toshiba, Japan) US scanner with 3.5 MHz curved linear probe was used. Coronal and transverse views of the both left and right kidneys were taken without speckle reduction. Custom written MATLAB (The MathWorks, Inc., Natick, Massachusetts, United States) application was used to measure kidney features. Kidney measurements were manually performed under the supervision of an experienced radiologist. Measurements included; maximum longitudinal LEN measured in the coronal plane, maximum transverse WDTH measured in the coronal plane at the HGHT of
the renal sinus, and the maximum THK measured in the transverse plane, as shown in Figure 1.

Kidney VOL calculations were based on the ellipsoid formula,\textsuperscript{24}

\[ \text{Volume} = 0.5223 \times \text{length} \times \text{width} \times \text{thickness (THK)} \]  

(1)

Cortical THK was measured as the distance from the outer border of the kidney cortex to the outer border of the medullary pyramid. Three such measurements were taken as shown in Figure 1c and average cortical THK value was calculated. BMI\textsuperscript{25} and the BSA\textsuperscript{26} were calculated for all the subjects. Kidney shape alternations were determined by the ratio of kidney LEN to kidney WDTH. Measured kidney features; LEN, WDTH, and calculated kidney VOL were normalized to anthropometric parameters; body CIRC, HGHT, BMI, and BSA. Hereinafter, normalized features will be represented in the following format: Feature/anthropometric parameter.

We used an RF procedure\textsuperscript{21} to classify given two groups. To determine which features were more influential within the RF framework, the mean decrease accuracy (accuracy index) was calculated. The top (4–7) features that could distinguish CKD\textsubscript{ALL} from healthy subjects in the descending order were VOL/CIRC, VOL/BSA, WDTH, VOL, VOL/BMI, LEN/CIRC, and WDTH/HGHT. Multivariate test for equality of mean vectors (permutation test) in the two groups showed a significant difference between healthy subjects and CKD\textsubscript{ALL} patients ($P < 0.001$). According to 95% bootstrapped confidence intervals (Bonferroni corrected), taken independently, all seven features were significantly low in CKD\textsubscript{ALL} compared to that of the healthy subjects (Table 2 for mean values). The 95% confidence interval of the mean difference between CKD\textsubscript{ALL} and healthy of each feature is given in Table 3.

**Healthy versus CKD\textsubscript{ALL}**

In this analysis, all CKD patients were grouped in to CKD\textsubscript{ALL}. According to the accuracy index [Figure 2a], the top seven features that could distinguish CKD\textsubscript{ALL} from healthy subjects in the descending order were VOL/CIRC, VOL/BSA, WDTH, VOL, VOL/BMI, LEN/CIRC, and WDTH/HGHT. Multivariate test for equality of mean vectors (permutation test) in the two groups showed a significant difference between healthy subjects and CKD\textsubscript{ALL} patients ($P < 0.001$). According to 95% bootstrapped confidence intervals (Bonferroni corrected), taken independently, all

### RESULTS

Anthropometric details of the two groups (CKD $n = 50$, healthy subject $n = 26$) are listed in Table 1. The mean ± standard deviations of all morphological features in this study are given in Table 2.

| Table 1: Demographic and anthropometric data of the subjects. |
|-----------------|------------------|------------------|
|                | CKD±SD ($\text{n}=50$) | Healthy subjects±SD ($\text{n}=26$) |
| Male (%)       | 56                | 54                |
| Age (years)    | 53±10             | 49±10             |
| Weight, (kg)   | 62±14             | 66±14             |
| Height (cm)    | 159±11            | 161±10            |
| Waist circumference (cm) | 91±13 | 90±12 |
| BMI (kg/m$^2$) | 24±4              | 25±5              |
| BSA (m$^2$)    | 1.6±0.2           | 1.6±0.2           |

CKD: Chronic kidney disease, BMI: Body mass index, BSA: Body surface area, SD: Standard deviation.

**Figure 1:** A 60-year-old male patient with chronic kidney disease. (a) Renal ultrasound image indicates the maximum longitudinal length (white) and maximum transverse width (yellow) measured in the coronal plane, (b) maximum thickness (THK) measured in the transverse plane, (c) cortical THK measurements in the ultrasound image from the outer border of the kidney cortex to the outer border of the medulla.
Table 2: Mean values of measured and calculated features.

<table>
<thead>
<tr>
<th>Feature</th>
<th>CKD_{KNOWN}</th>
<th>CKD_{UNKNOWN}</th>
<th>CKD_{all}</th>
<th>Healthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEN (cm)</td>
<td>8.2±0.2</td>
<td>9.2±0.9</td>
<td>8.7±0.1</td>
<td>10.0±0.8</td>
</tr>
<tr>
<td>WDTH (cm)</td>
<td>3.5±0.4</td>
<td>3.7±0.4</td>
<td>3.6±0.5</td>
<td>4.3±0.6</td>
</tr>
<tr>
<td>VOL (cm³)</td>
<td>53±16</td>
<td>64±15</td>
<td>58±17</td>
<td>87±25</td>
</tr>
<tr>
<td>ACT (cm)</td>
<td>1.0±0.2</td>
<td>1.0±0.1</td>
<td>1.0±0.2</td>
<td>1.1±0.1</td>
</tr>
<tr>
<td>Shape index (LTW)</td>
<td>0.44±0.06</td>
<td>0.4±0.1</td>
<td>0.42±0.06</td>
<td>0.43±0.05</td>
</tr>
<tr>
<td>LEN/CIRC</td>
<td>0.9±0.2</td>
<td>1.0±0.2</td>
<td>0.9±0.2</td>
<td>1.1±0.1</td>
</tr>
<tr>
<td>LEN/BSA (cm/m²)</td>
<td>52±8</td>
<td>57±6</td>
<td>54±8</td>
<td>61±7</td>
</tr>
<tr>
<td>LEN/BMI (cm²/kg)</td>
<td>3.5±0.7</td>
<td>3.8±0.8</td>
<td>3.6±0.7</td>
<td>4.0±0.7</td>
</tr>
<tr>
<td>LEN/HGHT (cm/m)</td>
<td>0.51±0.09</td>
<td>0.58±0.05</td>
<td>0.55±0.09</td>
<td>0.62±0.06</td>
</tr>
<tr>
<td>WDTH/CIRC</td>
<td>0.40±0.07</td>
<td>0.41±0.06</td>
<td>0.40±0.07</td>
<td>0.48±0.07</td>
</tr>
<tr>
<td>WDTH/BSA (cm/m²)</td>
<td>22±3</td>
<td>23±3</td>
<td>23±3</td>
<td>26±3</td>
</tr>
<tr>
<td>WDTH/BMI (cm²/kg)</td>
<td>1.5±0.3</td>
<td>1.5±0.4</td>
<td>1.5±0.3</td>
<td>1.7±0.4</td>
</tr>
<tr>
<td>WDTH/HGHT (cm/m)</td>
<td>0.22±0.03</td>
<td>0.23±0.02</td>
<td>0.22±0.03</td>
<td>0.27±0.03</td>
</tr>
<tr>
<td>VOL/CIRC (cm³)</td>
<td>0.6±0.2</td>
<td>0.7±0.2</td>
<td>0.6±0.2</td>
<td>0.9±0.2</td>
</tr>
<tr>
<td>VOL/BSA (cm³/kg)</td>
<td>33±9</td>
<td>39±8</td>
<td>36±9</td>
<td>52±13</td>
</tr>
<tr>
<td>VOL/BMI (cm³/m²)</td>
<td>2.3±0.7</td>
<td>2.6±0.7</td>
<td>2.4±0.7</td>
<td>4±1</td>
</tr>
<tr>
<td>VOL/HGHT (cm³/m)</td>
<td>0.3±0.1</td>
<td>0.40±0.09</td>
<td>0.4±0.1</td>
<td>0.5±0.1</td>
</tr>
</tbody>
</table>

CKD: Chronic kidney disease, LEN: Length, WDTH: Width, VOL: Volume, ACT: Average cortical thickness, LTW: Length to kidney width, LEN/BMI: Kidney length normalized to body mass index, LEN/BSA: Kidney length normalized to body surface area, LEN/CIRC: Kidney length normalized to body circumference, LEN/HGHT: Kidney length normalized to body height, VOL/BMI: Kidney volume normalized to body mass index, VOL/BSA: Kidney volume normalized to body surface area, VOL/CIRC: Kidney volume normalized to body circumference, VOL/BSA: Kidney volume normalized to body surface area, VOL/CIRC: Kidney volume normalized to body circumference, VOL/HGHT: Kidney volume normalized to body height, WDTH/BMI: Kidney width normalized to body mass index, WDTH/BSA: Kidney width normalized to body surface area, WDTH/CIRC: Kidney width normalized to body circumference, WDTH/HGHT: Kidney width normalized to body height.

Figure 2: Mean decrease accuracy of the random forest classifier for chronic kidney disease (CKD)ALL patients and healthy subjects (a) and CKDKNOWN and CKDUNKNOWN groups (b).
patients ($P = 0.01$). According to 95% Bonferroni confidence intervals, taken independently, except the kidney LEN normalized to CIRC (LEN/CIRC), other six features were significantly low in CKD\textsubscript{unknown} patients compared to that of the healthy subjects. The 95% confidence intervals of the features are given in Table 3.

### CKD\textsubscript{known} versus CKD\textsubscript{unknown}

RF classification of CKD\textsubscript{known} and CKD\textsubscript{unknown} groups found the highest accuracy index for LEN/HGHT, followed by LEN/BSA, VOL/BSA, LEN/CIRC, and VOL/HGHT [Figure 2b]. A multivariate test for equality of mean vectors in the two groups did not find a significant difference between CKD\textsubscript{known} patients and CKD\textsubscript{unknown} patients ($P = 0.062$). However, according to the 95% bootstrapped confidence intervals, taken independently, LEN normalized to HGHT and VOL normalized to HGHT were significantly different between the groups [Table 3]. On the other hand, the eGFR and age of CKD\textsubscript{known} (eGFR = 43 ± 20, age = 57 ± 3 years) patients were significantly different ($p=0.01$) to that of CKD\textsubscript{unknown} (eGFR = 57 ± 19, $P = 0.02$ age = 47 ± 7 years).

### DISCUSSION

Our study demonstrates that US-based kidney morphological features in different combinations can distinguish healthy subjects and CKD patients. According to the RF model, kidney VOL normalized to waist CIRC, BSA, BMI, and kidney WDTH are the most important features in the classification. This is in line with the prior studies which report correlations between VOL and functional ability of kidney in different disease profiles.\cite{38,39} The reduction in kidney VOL in CKD may be associated with known pathological alternations such as glomerular sclerosis and tubular atrophy.\cite{40}

Serum creatinine level of healthy subjects was not measured in this study, which could be considered a limitation, as their kidney health could not be verified biochemically.

Our results show that the normal kidney LEN (10 ± 1 cm) of Sri Lankan population is similar to that of other Asian countries; Northwest Indians = 10.0 cm and Pakistanis = 10.4 cm. The kidney VOL (87 cm\textsuperscript{3}) of Sri Lankans as reported in this study appears to be smaller than that of Danish (140 cm\textsuperscript{3}), Korean (158 cm\textsuperscript{3}), and Caucasian (130 cm\textsuperscript{3}) ethnicities.\cite{41}

In this study, we identified two groups of CKD patients on their clinical presentation at the time of the study. One group of patients presented with classical risk factors (CKD\textsubscript{known}), and the etiology of the other CKD group was unknown (CKD\textsubscript{unknown}). The etiology of the CKD\textsubscript{unknown} group cannot be determined from the available data but could be hypothesized based on the literature as CINAC (or CKDu) patients. The present study population was drawn from a region of high prevalence of CINAC\textsuperscript{10} and CKD\textsubscript{unknown} patients.
patients did not have any of the classical risk factors. Therefore, it is very likely that CKD\text{unknown} patients may be clinically defined as CINAC patients. However, it was not able to verify this retrospectively and can be regarded as a limitation of this study. It should also be pointed out that since all CKD patients in this study were from CINAC endemic areas, some of the CKD\text{unknown} patients could have complex etiologies.

We found that the kidney LEN normalized to HGHT and VOL normalized to HGHT were significantly different between CKD\text{unknown} and CKD\text{known}. However, CKD\text{unknown} patients were younger, had higher eGFR, and did not have any of the classical risk factors.

CONCLUSION

It is likely that the progression and the severity of CKD\text{unknown} (possibly CINAC) are different to that of the CKD\text{known}. Ultrasound-based kidney morphological measurements in combination with RF classification could play a key role in future studies of disease progression in this group.

**Declaration of patient consent**

The authors certify that they have obtained all appropriate patient consent.

**Financial support and sponsorship**

National Research Council, 120/07, Vidya Mawatha, Colombo 07, Sri Lanka (Grant Number: NRC 16-044).

**Conflicts of interest**

There are no conflicts of interest.

**REFERENCES**

23. Levey AS, Bosch PJ, Lewis JB, Greene T, Rogers N, Roth D. A
27. R Core Team. R: A Language and Environment for Statistical Computing; 2013.