

## Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

---

Tanya Seneviratne

---

Date

Evaluating the occupational and lifestyle risk factors among Chronic Kidney Disease of  
unknown etiology (CKDu) patients in the North Central Province of Sri Lanka

By

Tanya Seneviratne  
Master of Public Health

Global Environmental Health

---

N. Kyle Steenland, PhD  
Committee Chair

---

Paige Tolbert, PhD  
Committee Member

Evaluating the occupational and lifestyle risk factors among Chronic Kidney Disease of unknown etiology (CKDu) patients in the North Central Province of Sri Lanka

By

Tanya Seneviratne

B.S.

University of California, Irvine

2014

Thesis Committee Chair: N. Kyle Steenland, PhD

An abstract of

A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Public Health  
in Global Environmental Health  
2018

## Abstract

Evaluating the occupational and lifestyle risk factors among Chronic Kidney Disease of unknown etiology (CKDu) patients in the North Central Province of Sri Lanka

By Tanya Seneviratne

**Purpose** Chronic Kidney Disease of unknown etiology (CKDu) was first reported in the early 1990s in the North Central Province of Sri Lanka, and has now escalated to an epidemic level. This study attempts to further characterize the occupational exposures and daily practices of people in the disease endemic region that puts them at an increased risk of disease.

**Methods** A case-control study was conducted at the Medawachchiya renal clinic in the Anuradhapura district of the North Central Province. Cases (N=200) were selected from patients diagnosed with stage 3 or 4 CKDu from the clinic registry, and controls (N=200) were selected by asking each case to bring a spouse, relative, or neighbor who was confirmed to be disease-free. A survey questionnaire was administered to collect demographic information, disease history, and lifestyle habits regarding water consumption and occupational exposure to agrochemicals. Missing occupation data (31.5%) were imputed, and logistic regression modeling was performed to characterize disease risk using covariates sex, age, education, occupation, pesticide use, drinking from an abandoned well.

**Results** Among the diagnosed cases, the majority were male (N=117, 58.5%) and had up to a primary, or grade 5, education (N=107, 53.5%). The mean age for male and female cases was not statistically significant, but there was a 15.5 year age difference between cases and controls. Of the male CKDu patients in the non-imputed data, 74.36% worked as farmers, compared to 53.01% in female patients. Non-imputation or complete data modeling showed sex (OR=1.79, 95% CI 1.04-3.06), age (OR=1.12, CI 1.08-1.15), low education (OR=2.73, CI 1.44-5.15), pesticide use (OR=1.72, CI 1.01-2.92), and drinking water from an abandoned well (OR=2.82, CI 1.19-6.65) to be statistically significant predictors. Modeling with imputed data (c-statistic=0.855) showed sex (OR=2.01, CI 1.14-3.53), age (OR=1.11, CI 1.08-1.15), and drinking from an abandoned well (OR=2.69, CI 1.16-6.26) to increase risk of CKDu.

**Conclusion** This study supports previous findings that sex, age, and low education are important risk factors of disease. Both imputation and non-imputation modeling demonstrate that in this study population, being a farmer may not be the best predictor of CKDu, but rather the use of pesticides and drinking water from abandoned wells that may be contaminated with heavy metals and agrochemicals.

Evaluating the occupational and lifestyle risk factors among Chronic Kidney Disease of unknown etiology (CKDu) patients in the North Central Province of Sri Lanka

By

Tanya Seneviratne

B.S.  
University of California, Irvine  
2014

Thesis Committee Chair: N. Kyle Steenland, PhD

A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Public Health  
in Global Environmental Health  
2018

## **Acknowledgements**

I would first like to thank my thesis advisor, Dr. Kyle Steenland, who provided me with patient guidance throughout the entire process of writing my thesis. I learned a great deal during our many advising meetings, and I am honored to have worked with someone whom I admire so much.

I would also like to thank the faculty of the Environmental Health department at Rollins for teaching courses, organizing events, and providing opportunities that helped shape my career interests within the scope of environmental health, and that have led me to a path that I am happy to be on.

## Table of Contents

Introduction .....	1
Background and Significance .....	1
Additional Regions with CKDu Prevalence .....	3
Characteristics and History of the North Central Province .....	4
Major Hypothesized Etiologies .....	6
Methods .....	10
Study Participants .....	10
Ethical Consideration .....	11
Data Collection .....	11
Non-imputed (Complete) Data Analysis .....	11
Imputed Data Analysis .....	12
Results .....	14
Demographic Information .....	14
Non-imputed Logistic Modeling (Complete Data) .....	17
Imputation and Logistic Modeling .....	20
Discussion .....	25
Conclusions .....	29
References .....	30

## **Introduction**

### *Background and Significance*

Over the past several years, after a decades-long civil war, major public health successes, and improvements to infrastructure and education, Sri Lanka has transitioned from a low to middle income country, shown by a steady increase in GDP (Peiris, 2014). Overall improvements in all sectors have led to a shift in the prevalence of disease from infectious and maternal/childhood diseases to non-communicable diseases such as heart disease, diabetes, and chronic kidney disease (CKD). Sri Lanka's eradication of smallpox, polio, and even malaria has led to a transition from diseases normally prevalent in developing countries to diseases characteristic of developed nations. Margaret Chan, former Director-General of the World Health Organization, states that 85% of premature deaths resulting from non-communicable diseases worldwide occur in developing countries ("WHO highlights need for countries to scale up action on noncommunicable diseases. ," 2014).

Beginning in the early 1990s, an abnormal type of chronic kidney disease began to emerge in the North Central Province of Sri Lanka. Patients with this type of CKD lacked the traditional risk factors associated with the disease, including diabetes, hypertension, family history of kidney failure, or glomerulonephritis resulting from infectious disease. These patients had symptoms characteristic of CKD, which is defined as kidney damage or glomerular filtration rate (GFR) less than  $60\text{ml}/\text{min}/1.73\text{m}^2$  that persists for 3 months or more irrespective of cause ("Chapter 1: Definition and classification of CKD," 2013). GFR below  $60\text{ml}/\text{min}/1.73\text{m}^2$  falls into stages 3a-5 of CKD. This unusual type of CKD manifests as a tubulo-interstitial nephritis in which patients are seen to have small, irregularly contoured kidneys with glomerulosclerosis noted after renal biopsy (Jayasumana et al., 2017). This type



of CKD, with causative factors that have yet to be identified, was named Chronic Kidney Disease of unknown etiology, or CKDu.

Over the past two decades, studies have shown variations in the prevalence of CKDu in Sri Lanka, likely due to differing methods of disease assessment. The age-standardized prevalence of CKDu is estimated to be 15% (Jayatilake et al 2013). Chandrajith and colleagues conducted a study of CKDu patients in the endemic region and found a prevalence of 2-3% (2011). However, researchers in this study used an elevated albumin creatinine ratio (ACR), or proteinuria, as the primary indicator to diagnose CKDu. Proteinuria is a sensitive measure of screening CKD that is attributable to traditional risk factors and is a low-cost method often used population based studies. While this is a popular way to assess kidney damage, it is well documented that CKDu cases show minimal evidence of proteinuria in early stages of the disease (Redmon et al., 2014), and is thus an ineffective tool for screening for CKDu cases. As part of a collaborative effort between the World Health Organization and the Sri Lanka Ministry of Health between 2009 and 2012, the age-standardized population prevalence of CKDu was estimated to be 16.9% [95% confidence interval (CI) 15.5–18.3] in women and 12.9% (95% CI 11.5–14.4) in men (*International Expert Consultation on Chronic Kidney Disease of Unknown Etiology*, 2016). This study also used proteinuria as an indicator of disease. Using proteinuria as an indicator of disease likely results in the underreporting of CKDu cases since patients lacking proteinuria are excluded after initial screening. In order to more effectively assess prevalence, using a screening method such as estimated glomerular filtration rate (eGFR) measured with serum creatinine levels is ideal. Additionally, biomarkers for tubular kidney damage can be used, such as urinary kidney injury molecule (KIM-1) and urinary gelatinase-associated lipocalin (NGAL)- to creatinine ratio (De Silva, P. et al, 2016; Redmon et al., 2014) for an accurate assessment of CKDu. The Sri Lankan Ministry of Health estimates

that there are currently 60,000 cases of CKDu and 20,000 annual deaths due to CKDu (Ministry of Health). The high disease burden seen in Sri Lanka necessitates a nation-wide study using serum creatinine or specific urinary biomarkers such as KIM-1 or NGAL to assess CKDu population prevalence. Additionally, surveillance programs should be implemented to assess population prevalence of disease.

The epicenter of the disease in Sri Lanka is the North Central Province (NCP), which is comprised of the Anuradhapura and Polonnaruwa districts. The NCP is an area of predominantly low socioeconomic status, and the primary industry in this area is agriculture, with 45.7% of people in the NCP working as farmers (Economic Census 2013/14). Although renal clinics for CKD patients exist in this region, the immense crowding of hospitals and high cost of treatment prevents many CKDu patients from receiving care. A lack of availability of both dialysis and renal replacement therapy, along with the low SES of most patients, means that CKDu is often a terminal diagnosis, with prognosis usually 2 years from stage 3 CKDu diagnosis. The high mortality of disease is often devastating to those left behind- the loss of what is often the family's breadwinner means that children are often forced to work instead of attending school in order to provide income for their family. The growing CKDu epidemic could have national consequences as well- as the disease becomes more severe, Sri Lanka's economy may be impacted as the endemic region holds the major rice farming communities in the country.

#### *Additional Regions with CKDu Prevalence*

Although CKDu first appeared in the North Central Province of Sri Lanka, the disease is also seen in Central America, India, Egypt, and is emerging in other regions of Sri Lanka's dry zone. Mesoamerican nephropathy (MeN) is seen in the coastal lowlands of El Salvador,

Nicaragua, Guatemala, and Costa Rica. The characteristics of MeN have been shown to draw many parallels with that of CKDu in Sri Lanka- those primarily affected are outdoor laborers, and prevalence is particularly high in sugarcane workers in these countries. A similar type of CKDu is also seen in the Andra Pradesh region of India and the El-Minia Governorate in Egypt (Weaver, Fadrowski, & Jaar, 2015). In Egypt, 27% of end stage renal disease patients seeking dialysis treatment were seen to have CKD with no known etiology (El Minshawy 2011). Additionally, pesticide exposure and drinking from tube wells were seen to be associated with the disease in this region (Kamel & El Minshawy 2010). De Silva and colleagues have also found an emergent presence of CKDu cases in the Hambantota district, which is in the Southern Province of Sri Lanka but has similar dry climate and farming practices to the endemic region in the NCP.

#### *Characteristics and History of the North Central Province*

The North Central Province of Sri Lanka is the country's largest district, but not its most populous. The NCP was home to the ancient civilization of the Anuradhapura Kingdom, which is the oldest in Sri Lanka. Remnants of artifacts found in Anuradhapura excavations date to the Iron Age, in 900 BCE (Allchin, 1989). The NCP is located in the dry zone of Sri Lanka, characterized by a hot and arid climate with low elevation. This region of the country is mostly rural and has been used for agricultural purposes from the early stages of its civilization. Evidence of agricultural land use as early as 600 BCE have been found. A wide range of crops are farmed in this region, including lotus root, grains, grams, and many varieties of tropical fruits and vegetables, however rice paddy is the most abundant. The NCP contains 25% of Sri Lanka's total paddy farmland, the largest of any province (USAID). The NCP is

also known to have a higher yield of paddy compared to other provinces in the country, which is attributable to the extensive irrigation in the area.

In an effort to route water to farmland in early civilizations, a cascading irrigation system known as the Tank Cascade System (TCS) was constructed between 550 and 300 BCE. This reservoir based system functions by recycling water through a system of small and large reservoirs, or tanks. This irrigation system, still widely in use today, has evolved over the course of millennia as technology and land use changed. Since its beginning in ancient irrigation, the tank system in the region is now comprised of over 30,000 tanks of varying sizes on approximately 40,000km<sup>2</sup> of land in Sri Lanka's dry zone (Mendis 2003). The TCS in Sri Lanka is thought to be essential to water management systems in the country and has protected farmland from chronic and recurring droughts, seasonal flooding due to monsoons, and land degradation (Geekiyana & Pushpakumara 2013). This complex irrigation system has also been vital in providing food security to the country, allowing farmers to produce enough rice to sustain the country's population as well as meet export demands.

The 1960s Green Revolution that pushed forward new agricultural technologies in much of the developing world catalyzed a widespread and unregulated use of pesticides and other agrochemicals that peaked in Sri Lanka in the mid-1980s (Schreinemachers & Tipraqsa 2012). The scientific advancements that effectively managed pests and weeds and increased crop yields were welcomed in agricultural communities across the world, including in Southeast Asia. In the North Central Province of Sri Lanka, this widespread use of agrochemicals was later seen to have many detrimental effects on the environment and people. One such effect is the pollution of surface and groundwater by several agrochemicals (FAO 2015). The pollution of water by agrochemicals can occur through several mechanisms including surface runoff, erosion of contaminated soils, leaching through permeable soils, and

spray drift when chemicals are used near canals, drains, or water reservoirs. Additionally, the NCP's Tank Cascade System of irrigation has led to the accumulation of naturally occurring and chemical derived heavy metals in major reservoirs at lower elevations (J. M. Bandara et al., 2008). These reservoirs are often used for drinking, bathing, and washing of clothes by rural villagers, leaving them susceptible to ingestion, dermal, and inhalational exposures to a variety of agrochemicals in addition to heavy metals. Another factor unique to the NCP is the high prevalence of wells that have been abandoned due to unpalatability of the water. Villagers in the CKDu endemic region have reported that many shallow wells no longer provide water suitable for drinking and have now been abandoned due to high hardness and bad taste (Jayasumana 2015). Additionally, ground water in the region is known to be either hard or very hard (Calcium and Magnesium content of 121-180mg/L or greater than 180mg/L, respectively) (Paranagama 2013). Due to scarcity of adequate water sources in the agricultural fields, farmers in the NCP have reported having to drink from wells previously abandoned due to unpalatability, or travel far distances for potable water and risk dehydration and further exacerbation of renal damage.

#### *Major Hypothesized Etiologies*

The extent of published research on CKDu in Sri Lanka has proven that there is no singular cause for the disease, but rather that the etiology is multifactorial (Correa-Rotter, Wesseling, & Johnson, 2014; Jayasumana et al., 2017; Wanigasuriya, 2014). Many studies have attempted to identify the causative factors of the disease, but consistent results are yet to be found. A prominent hypothesis for the etiology of CKDu in the North Central Province was put forth in 2014 by a research team based out of Rajarata University of Sri Lanka and involves the combined effect of drinking water with high hardness and occupational exposure to

glyphosate. Consumption of water with high hardness has been linked to CKDu in multiple studies (Paranagama 2013, Jayasumana 2014, De Silva et al 2016, Valcke 2017). This hypothesis involves consumption of hard water along with exposure to another agent that could form stable complexes with hard water, bind with nephrotoxic metals, and act through multiple exposure routes. This secondary agent that aids in kidney toxicity when combined with hard water was proposed to be glyphosate. Glyphosate and its metabolite AMPA have been shown to have high chelating ability with many metals, including As, Ca, Mg, Zn, Cu, and others. The authors have suggested three possible mechanisms for formation of glyphosate-metal complexes: in drinking water, including wells, reservoirs, and surface water, in food farmed in the region, or formation of the complex within the blood circulation from ingested metals and inhaled or dermally absorbed glyphosate. Farmers often use hard water from wells to create the glyphosate solution used on crops, and the product can also runoff or be sprayed into waterways. It has been found that several crops farmed in the endemic region including rice, vegetables, and tobacco contain both cadmium and arsenic (Jayatilake 2013). The authors made several points to justify the implication of glyphosate in this hypothesis. First, it is a chemical that has been used in Sri Lanka since the late 1970s, and was the highest imported pesticide before being banned in 2017. In 2012, more glyphosate was imported into Sri Lanka than any other pesticide combined (Jayasumana 2014). Secondly, before being used as an herbicide, glyphosate was initially patented as a chelating agent to be used to clean calcium and minerals in hot water systems due to its efficiency in binding metals to make them water soluble. Glyphosate's chemical structure allows it to be an excellent binder of metals, and also gives it high water solubility. While its normal half-life in water and soil are 92 and 47 days, respectively, the absorption of chelating metals can increase glyphosate's soil half-life to between 7 and 22 years (Ebberbach 1999 and Nomura 1977). Jayasumana and

colleagues note that the timeline of glyphosate's introduction and widespread use lines up with the first documented cases of CKDu in the early 1990s, where clinically identifiable CKD could have taken between 12-15 years to appear. The implication of glyphosate led to a nationwide ban of the herbicide in 2015.

Another major hypothesis on the etiology of CKDu is the chronic heat stress and dehydration hypothesis proposed by many studies on Mesoamerican Nephropathy (MeN) in Central America. Researchers suggest that the cause of the disease in sugarcane farmers is related to their occupational conditions in which intense manual labor is performed in hot conditions with little water consumption throughout the day. Several studies suggest that chronic exposure to dehydration and heat stress can contribute to acute kidney injury (AKI) and the eventual onset of CKDu (Glaser et al., 2016; Roncal Jimenez et al., 2014; Wesseling et al., 2016). While sugarcane workers on plantations in Central America are often exposed to long periods of heat without access to rest and shade for the long harvest season, paddy farmers in Sri Lanka are self-employed and can take breaks as needed. Additionally, they spend only 30-40 days per year in the field for harvesting. Additional points against the heat stress nephropathy etiology are detailed by Herath and colleagues (Forthcoming 2018). While occupational exposure to heat, along with the morphological and clinical characteristics of the disease (Wijkstrom et al., 2018) differ among CKDu patients in Sri Lanka and Central America, both regions have similar climates, and the heat stress hypothesis should be investigated further in Sri Lanka.

Over the past several decades, many additional hypotheses on the etiology of the disease have emerged in addition to the two major theories detailed previously. Arsenic from phosphate fertilizers is thought to contaminate food and water supplies and lead to nephrotoxicity (Jayasumana et al., 2015; Jayasumana, Gajanayake, & Siribaddana, 2014).

Exposure to naturally occurring fluoride in groundwater in endemic regions is thought to be a cause of the disease (Chandrajith, Dissanayake, Ariyaratna, Herath, & Padmasiri, 2011; Wasana et al., 2016), however this theory does not explain absence of CKDu in regions where fluoride in drinking water is high. Another proposed hypothesis is chronic low dose exposure to cadmium through both drinking water and food (J. M. Bandara, Senevirathna, D. M., Dasanayaka, D., Herath, V. Bandara, J. M., Abeysekara, T., & Rajapaksha, K.H., 2008; Jayatilake, Mendis, Maheepala, Mehta, & Team, 2013). These findings prompted an advisory by government officials to stop consuming lotus root and certain types of fish that were known to be high in cadmium content. In addition to hypotheses regarding occupational conditions, exposure to agrochemicals, and other heavy metals, researchers have also investigated the role of genetics in the onset of the disease (Nanayakkara et al., 2014). It is evident that there is no single etiology of CKDu, and that causation is likely multifactorial, possibly involving a combination of several aforementioned hypotheses.



## Methods

### *Study Participants*

The data analyzed at present were obtained from a case-control study conducted in Medawachchiya, a Divisional Secretariat of the Anuradhapura district in the in the North Central Province. This region has been previously identified as one of the highest CKDu prevalent regions in Sri Lanka.

At the time of study, there were 878 patients in stage 3 or 4 of CKDu as seen on the Medawachchiya Hospital Renal Clinic register. Patients were diagnosed with having CKD if they had kidney damage or glomerular filtration rate (GFR) less than  $60\text{ml}/\text{min}/1.73\text{m}^2$ , along with meeting the following criteria for the case definition of CKDu (Jayatilake et al 2013):

- No past history of, or current treatment for, diabetes mellitus or chronic/severe hypertension, history of snakebite, urological disease of known etiology, or glomerulonephritis
- Normal glycosylated hemoglobin ( $\text{HbA}_{1c} < 6.5$ )
- Blood pressure  $< 160/100$  mmHg if not on treatment for hypertension or  $< 140/90$  mmHg on up to two antihypertensive treatments

Recruitment of cases was initialized by sending letters to 300 of the 878 diagnosed CKDu patients on the clinic register at random. Those that responded to the first phase of recruitment were called in to ensure that all inclusion criteria were met. The first 200 cases that were confirmed in stage 3 or 4 of CKDu were selected as cases. No cases with end stage CKDu (stage 5) were selected for this study.

Controls were selected by asking each chosen case to bring a spouse or family member of the same household or a neighbor who was unaffected by CKDu. The cases and controls were matched by location. Non-CKDu status was verified by clinical examination, and 200 controls were included in the study. All cases and controls were compensated by financial

reimbursement for daily transportation fees, approximately 500 Rupees (approximately \$3.50 USD) per clinic visit.

### *Ethical Consideration*

Prior to data collection, approval from the Ethics Review Committee of the Faculty of Medicine at Rajarata University was obtained. Emory University IRB approval was obtained for secondary data analysis. Written informed consent was obtained from all study participants.

### *Data Collection*

A survey questionnaire was administered by resident physicians at the Medawachchiya Renal clinic to obtain information about participants' exposure history. Demographic data were collected, in addition to information about participants' family history of disease, agrochemical use, food and water consumption habits, and history of severe illness. Data from hospital records was obtained retrospectively for all cases. Serum creatinine measurements from all clinic visits in patient's history were noted for later GFR calculations.

### *Non-imputed (Complete) Data Analysis*

All data analyses were conducted using the statistical program SAS (Version 9.4, SAS Institute Inc., Cary, NC, USA 2013). Associations between outcome and exposure variables were investigated with frequency tables, and continuous variables were analyzed univariately by their means and standard deviations. Categorical variables were either dichotomized or coded into a small number of category levels for analysis. To simplify analysis and imputation, occupation was dichotomized as a binary variable with values Farmer vs Not Farmer. Age was

modeled as a continuous variable after noting monotonic trends in categorical analyses and better fit using a continuous variable compared to categorization. Education was coded as a 3 level numerical variable which the procedure treats as continuous. Sex, case-control status, pesticide use, and drinking from an abandoned well were binary variables. Seven logistic regression models were run with various combinations of covariates. Covariates were included based on whether they were significant predictors of the outcome, or were variables of a priori interest such as farmer/not farmer or pesticide use.

#### *Imputed Data Analysis*

Due to the high percentage (31.5%) of missing observations for the variable occupation, which is an important established risk factor of CKDu, multiple imputation method was used to impute data. Data was imputed for both occupation and drinking from abandoned well (8.25% missing), since this was shown to be a significant predictor in non-imputed modeling. This imputation method is a three step process involving filling in missing data using SAS PROC MI, analysis of multiple imputed datasets using logistic regression, and generation of valid statistical inferences using results from the logistic regression analyses with SAS PROC MIANALYZE. The multiple imputation analysis was run with occupation, drinking from abandoned well, sex, age, education level, pesticide use, and case-control status as predictors. Imputation generated 5 different imputed datasets, in which each predictor variable had unique variances. The imputed datasets were then treated as complete and modeled using logistic regression methods. Covariates included in imputed models were selected based on their importance in the analyses of complete data, and included age, sex, education, occupation, pesticide use, and drinking water from abandoned wells. In all logistic regression models, the concordance-statistic was noted to determine goodness of model fit

for the binary case-control outcome. Logistic analysis of five imputed datasets produced five sets of point and variance estimates for each predictor. The third step in the imputation process involved combining those results to generate accurate statistical inferences about each predictor. This was done via the SAS procedure, MIANALYZE, which derives an average of the parameters of interest across the five imputed data sets, as well as appropriate confidence intervals for those parameters, taking into account the additional variance due to imputation. Odds ratios (OR), 95% confidence intervals (CI), and p-values were calculated for all parameters in both non-imputed and imputed data. Results were considered significant at  $p < 0.05$ .

## Results

### *Demographic Information*

Of the 400 study participants, 99.5% were from Anuradhapura, the district in which the Medawachchiya Renal clinic is located, and 97.5% were of Sinhalese ethnicity. Among the diagnosed cases of CKDu (N=200), 117 were male (58.5%) and 83 were female (41.5%). There was no significant difference in the mean age among male CKDu patients ( $63.32 \pm 10.38$  years) and females ( $64.49 \pm 10.25$  years). The difference in mean age of cases and controls was 15.5 years. Most cases (53.5%) were educated up to the primary level (grade 5), while most controls (55%) received a senior secondary education (grade 11) or higher. The study population was comprised of farmers, military personnel, masons, textile workers, businesspeople, among others. 55% of the total study population reported working as farmers, however, information on occupation was missing for 31.5% of the population, which will be discussed later. For ease of analysis, occupation categories were combined into a binary covariate for modeling. Of the male CKDu patients, 87 (74.36%) work as farmers, compared to 44 (53.01%) of female patients. The controls included 200 healthy individuals that were found to be free of CKDu at screening. This population included 33% males and 67% females. Among controls, 48.48% of males and 42.54% of females were farmers. 17% of cases reported drinking water from a well that had been previously abandoned due to unpalatability, compared to 6% of controls. Further information about the demographic characteristics of the study population can be found in Tables 1 and 2.

**Table 1. Characteristics of the Study Population (N=400)**

	<i>All Participants</i>		<i>Cases Only</i>		<i>Controls Only</i>	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
<b>Sex</b>						
Male	183	45.75	117	58.50	66	33.00
Female	217	54.25	83	41.50	134	67.00
<b>Highest Education Level</b>						
None-Primary	152	38.00	107	53.50	45	22.50
Junior Secondary	102	25.50	57	28.50	45	22.50
Senior Secondary +	146	36.50	36	18.00	110	55.00
<b>Occupation</b>						
Farmer	220	55.00	131	65.50	89	44.50
Military	7	1.75	2	1.00	5	2.50
Indoor Laborer	23	5.75	11	5.50	12	6.00
Outdoor laborer	8	2.00	4	2.00	4	2.00
Business owner	7	1.75	2	1.00	5	2.50
Professional	5	1.25	0	0.00	5	2.50
None	4	1.00	0	0.00	4	2.00
Missing	126	31.50	50	25.00	76	38.00
<b>Occupation Binary</b>						
Farmer	220	55.00	131	65.50	89	44.50
Not Farmer	54	13.50	19	9.50	35	17.50
Missing	126	31.50	50	25.00	76	38.00
<b>Pesticide Use</b>						
Yes	177	44.25	109	54.50	68	34.00
No	217	54.25	89	44.50	128	64.00
Missing	6	1.50	2	0.50	4	0.50
<b>Drink from Abandoned Well</b>						
Yes	46	11.50	34	17.00	12	6.00
No	321	80.25	150	75.00	171	85.50
Missing	33	8.25	16	8.00	17	8.50
<b>Home Water Source</b>						
Bottled	111	27.75	55	27.50	56	28.00
Well	226	56.50	114	57.00	112	56.00
Rainwater	19	4.75	11	5.50	8	4.00
Tap	29	7.25	13	6.50	16	8.00
River or Spring	5	1.25	3	1.50	2	1.00
Other	10	2.50	4	2.00	6	3.00

**Table 2. Univariate analysis of age in the study population (N=397)**

	<i>N</i>	<i>Mean</i>	<i>Std Dev</i>
Total study population*	397	56.07	13.98
Males	183	59.04	12.79
Females	214	53.53	14.47
All cases	199	63.8	10.32
Male Cases	117	63.32	10.38
Female Cases	82	64.49	10.25
All controls	198	48.3	12.83
Male Controls	66	51.44	13.19
Female Controls	132	46.73	12.39

\*3 Females with missing age

Tables 3 and 4 below show the relationships (frequencies and Spearman correlations) between pesticides, occupation (farmer/non-farmer), and drinking in abandoned well. 25 farmers (12.63%) who used pesticides also drank from abandoned wells. 10 farmers (5.05%) who did not use pesticides also drank from abandoned wells. As expected, pesticide use and occupation is significantly correlated, as is pesticide use and drinking water from an abandoned well.

**Table 3. Frequencies of Pesticide Use and Drinking from Abandoned Well by Occupation**

	<b>Farmer</b>				<b>Not Farmer</b>				
	<i>Yes AW</i>		<i>No AW</i>		<i>Yes AW</i>		<i>No AW</i>		
	<i>Freq</i>	<i>Percent</i>	<i>Freq</i>	<i>Percent</i>	<i>Freq</i>	<i>Percent</i>	<i>Freq</i>	<i>Percent</i>	
<b><i>Used pesticides</i></b>	25	12.63	113	57.07	<b><i>Used Pesticides</i></b>	2	4.08	13	26.53
<b><i>Did not use pesticides</i></b>	10	5.05	50	25.25	<b><i>Did not use pesticides</i></b>	4	8.16	30	61.22
	$\chi^2= 0.064, p=0.8059$				$\chi^2= 0.0238, p=0.8773$				

*Yes AW: Drank from abandoned well*

*No AW: Did not drink from abandoned well*

**Table 4. Spearman Correlation Coefficients for Occupation, Pesticide Use, and Drinking from Abandoned Well**

	<i>Occupation</i>	<i>Pesticide Use</i>	<i>Abandoned Well</i>
<i>Occupation</i>	1	0.334	0.058
	268	<.0001 268	0.362 247
<i>Pesticide Use</i>	0.334	1	0.124
	<.0001 268	391	0.019 362
<i>Abandoned Well</i>	0.058	0.124	1
	0.362 247	0.019 362	362

*Coefficient*  
*p-value*  
*N*

#### *Non-imputed Logistic Modeling (Complete Data)*

Using the complete data, 7 different models were run to examine the relationships between potential risk factors of CKDu. The covariates used in these analyses were sex, age, education, occupation, pesticide use, drinking from an abandoned well, and occpest. Occpest is a binary variable where 1 indicates a farmer who has used pesticides and 0 indicates that they are either not a farmer and/or did not use pesticides. The covariate pesticide use was coded as a binary variable for using pesticides or not. In this set of models, education was coded as a three level categorical variable with level 1 indicating none to a highest attainment of primary education, 2 indicating junior secondary school (grade 5) and 3 indicating senior secondary (grade 11) education or higher. When modeling age as a 3-level categorical variable, odds ratios showed a significant monotonic increase by category (data not shown). When age was modeled as a continuous variable, model fit was better, demonstrated by higher c-statistics, and therefore age was considered continuous in all models.

Results are summarized in Table 5 below for non-imputed (complete data) models. In Model 1, with occupation the only key exposure variable, we find the odds ratio for farmers vs. non-farmers to be 1.72, which did not attain statistical significance. Model 2 adds pesticides



to the model and shows that both farmers and the pesticide-exposed have an increased risk of CKDu but neither is statistically significant, and the inclusion of pesticides decreases the OR for farmers. The variable pesticides has a stronger effect than farming, and pesticides and farming are, as expected, correlated (see Table 4). Models 1 and 2 motivated Model 3 where we look at farmers who are exposed to pesticides ( $ocpest=1$ ), and find that the combination has a 59% increased risk of CKDu which is close to significance ( $p=0.13$ ). Model 4 with pesticide as the only exposure variable finds that those exposed to pesticides have a 75% increased risk of CKDu, which attains statistical significance ( $p=.04$ ). This model has significantly more observations ( $N=391$ ) than the other models ( $N=268$ ), because it does not include occupation which is missing for 32% of the population. In Models 5-7 we see that drinking from an abandoned well is an important predictor of disease, but only modestly affects the prior estimates of occupation,  $ocpest$ , and pesticides. In particular we find that that in Model 7, with both pesticides and drinking from an abandoned well (which may involve exposure to additional pesticides as well as hard water), is our best model in the non-missing (complete) data, with a c-statistic of 0.855. This model also has the highest number of subjects ( $N=391$ ), since it did not use the occupation variable which is missing large amounts of data.

**Table 5. Covariate Models with Non-Imputed (Complete) Data**

	<b>Model 1 (c-statistic=0.817) N=268</b>			<b>Model 2 (c-statistic=0.821) N=268</b>		
<i>Source</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>
<i>Sex</i>	<b>2.44</b>	<b>[1.33, 4.45]</b>	<b>0.004</b>	<b>2.22</b>	<b>[1.20, 4.12]</b>	<b>0.011</b>
<i>Age</i>	<b>1.11</b>	<b>[1.07, 1.14]</b>	<b>&lt;.0001</b>	<b>1.11</b>	<b>[1.07, 1.14]</b>	<b>&lt;.0001</b>
<i>Education level 1 vs 3</i>	1.75	[0.82, 3.72]	0.145	1.80	[0.85, 3.83]	0.128
<i>Education level 2 vs 3</i>	1.41	[0.66, 3.03]	0.378	1.45	[0.67, 3.12]	0.350
<i>Occupation</i>	1.72	[0.77, 3.84]	0.187	1.40	[0.60, 3.28]	0.438
<i>Pesticide use</i>				1.62	[0.85, 3.08]	0.142
<i>Drink from abandoned well</i>						
<i>OccPest</i>						
	<b>Model 3 (c-statistic=0.818) N=268</b>			<b>Model 4 (c-statistic=0.849) N=268</b>		
<i>Source</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>
<i>Sex</i>	<b>2.20</b>	<b>[1.22, 3.98]</b>	<b>0.009</b>	<b>2.03</b>	<b>[1.20, 3.43]</b>	<b>0.008</b>
<i>Age</i>	<b>1.11</b>	<b>[1.07, 1.14]</b>	<b>&lt;.0001</b>	<b>1.11</b>	<b>[1.08, 1.15]</b>	<b>&lt;.0001</b>
<i>Education level 1 vs 3</i>	1.88	[0.91, 3.90]	0.089	<b>2.53</b>	<b>[1.35, 4.73]</b>	<b>0.004</b>
<i>Education level 2 vs 3</i>	1.46	[0.68, 3.14]	0.331	1.77	[0.91, 3.45]	0.092
<i>Occupation</i>						
<i>Pesticide use</i>				<b>1.75</b>	<b>[1.04, 2.96]</b>	<b>0.036</b>
<i>Drink from abandoned well</i>						
<i>OccPest</i>	1.59	[0.87, 2.88]	0.1297			
	<b>Model 5 (c-statistic=0.834) N=268</b>			<b>Model 6 (c-statistic=0.833) N=268</b>		
<i>Source</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>
<i>Sex</i>	<b>1.93</b>	<b>[1.03, 3.65]</b>	<b>0.044</b>	<b>1.92</b>	<b>[1.04, 3.54]</b>	<b>0.036</b>
<i>Age</i>	<b>1.11</b>	<b>[1.07, 1.15]</b>	<b>&lt;.0001</b>	<b>1.11</b>	<b>[1.08, 1.15]</b>	<b>&lt;.0001</b>
<i>Education level 1 vs 3</i>	1.93	[0.89, 4.16]	0.094	1.99	[0.95, 4.17]	0.069
<i>Education level 2 vs 3</i>	1.55	[0.71, 3.41]	0.274	1.56	[0.72, 3.42]	0.263
<i>Occupation</i>	1.31	[0.55, 3.12]	0.542			
<i>Pesticide use</i>	1.67	[0.86, 3.23]	0.128			
<i>Drink from abandoned well</i>	<b>3.18</b>	<b>[1.26, 8.05]</b>	<b>0.015</b>	<b>3.17</b>	<b>[1.26, 7.97]</b>	<b>0.014</b>
<i>OccPest</i>				1.595	[0.87, 2.93]	0.131
	<b>Model 7 (c-statistic=0.855) N=391</b>					
<i>Source</i>	<i>OR</i>	<i>95% CI OR</i>	<i>p</i>			
<i>Sex</i>	<b>1.79</b>	<b>[1.04, 3.06]</b>	<b>0.035</b>			
<i>Age</i>	<b>1.12</b>	<b>[1.08, 1.15]</b>	<b>&lt;.0001</b>			
<i>Education level 1 vs 3</i>	<b>2.73</b>	<b>[1.44, 5.15]</b>	<b>0.002</b>			
<i>Education level 2 vs 3</i>	1.91	[0.97, 3.76]	0.060			
<i>Occupation</i>						
<i>Pesticide use</i>	<b>1.72</b>	<b>[1.01, 2.92]</b>	<b>0.045</b>			
<i>Drink from abandoned well</i>	<b>2.82</b>	<b>[1.19, 6.65]</b>	<b>0.018</b>			
<i>OccPest</i>						

Education level 1: None-Primary  
Education level 2: Junior Secondary  
Education level 3: Senior Secondary +  
Occupation referent: Non-farmer  
OccPest: Not farmer and/or does not use pesticides

### *Imputation and Logistic Modeling*

As mentioned previously, the occupation variable, shown in previous studies to be a significant risk factor of CKDu, had a large proportion of missing observations ( $N_{\text{miss}}=126$ , 31.5%). In order to perform analyses on a full dataset without missing occupation observations, data were imputed using a binary occupation variable- farmer or not farmer. Missing observations for abandoned well were also imputed simultaneously. Observations with missing data for the variable pesticide ( $N_{\text{miss}}=6$ , 1.50%) were deleted to simplify imputation. Table 6 below describes the distribution of missing binary occupation data by sex and case control status. It is evident that there is a much higher amount of missing occupation information in females (47.47%) compared to in males (12.57%). Additionally, a higher percentage of controls have missing occupation information when compared to cases- 38% and 25%, respectively. Values for missing occupation observations were generated using the multiple imputation method in SAS which produced five imputed datasets.

**Table 6. Distribution of missing occupation data by sex and case control status**

	<i>Farmer</i>		<i>Not Farmer</i>		<i>Missing</i>		<b>TOTAL</b>
	<b>Frequency</b>	<b>Percent</b>	<b>Frequency</b>	<b>Percent</b>	<b>Frequency</b>	<b>Percent</b>	
<b>Cases</b>	131	65.50	19	9.50	50	25.00	200
<b>Controls</b>	89	44.50	35	17.50	76	38.00	200
<b>Males</b>	119	65.03	41	22.40	23	12.57	183
<b>Females</b>	101	46.54	13	5.99	103	47.47	217

**Table 7a. Selected demographics from participants with occupation information in non-imputed data (N=391)**

	<b>Occupation</b>				
	<i>N</i>	<i>Farmer</i>		<i>Not Farmer</i>	
		Frequency	Percent	Frequency	Percent
<b>Cases</b>	148	129	87.16	19	12.84
<b>Controls</b>	124	85	70.83	35	29.17
<b>Males</b>	160	119	74.38	41	25.63
<b>Females</b>	108	95	87.96	13	12.04

**Table 7b. Selected demographics from 5 imputed datasets N=1955**

	<b>Occupation</b>				
	<i>N</i>	<i>Farmer</i>		<i>Not Farmer</i>	
		Frequency	Percent	Frequency	Percent
<b>Cases</b>	985	872	88.53	113	11.47
<b>Controls</b>	970	677	69.69	293	30.21
<b>Males</b>	915	687	75.08	228	24.92
<b>Females</b>	1040	862	82.22	178	17.12

Tables 7a and 7b above show the distribution of occupation from both non-imputed and imputed data summarized by case-control status and sex. When comparing the ratios of stratified populations between imputed and non-imputed data, it can be seen that the distribution of imputed occupation was similar to that in the data without missing occupation. The model fit for our imputation model to predict missing occupation was good, with a c-statistic of 0.903. The same model was run with drinking from an abandoned well as the outcome, giving a c-statistic of 0.726, which was not as high as the model to predict occupation but was still reasonably good.

**Table 8. Spearman Correlation Coefficients for Occupation, Pesticide Use, and Drinking from Abandoned Well using 5 Imputed Datasets (N=1955)**

	<i>Occupation</i>	<i>Pesticide Use</i>	<i>Abandoned Well</i>
<i>Occupation</i>	1	0.249 <.0001	0.048 0.034
<i>Pesticide Use</i>	0.249 <.0001	1	0.103 <.0001
<i>Abandoned Well</i>	0.048 0.034	0.103 <.0001	1

*Coefficient*  
*p-value*

Table 8 above shows the Spearman correlation coefficients using the imputed data, which is comprised of 5 datasets with a total N of 1955. When compared to the coefficients in the non-imputed data (Table 4), we can see that occupation, is significantly correlated with both pesticide use and drinking from an abandoned well. Pesticide use and drinking from an abandoned well are also significantly correlated.

Using the imputed data, five separate logistic regression models were run to examine the effects of multiple key covariates on case-control status. The variables included in these models were sex, age, education, drinking from an abandoned well, occupation, pesticide use, and occpest. Age was modeled as a continuous numerical variable. The variable education was coded into a 3-level numerical categorical variable, but was treated as a continuous variable for ease of use in later imputation analyses. There is no model with imputed data for “pesticides only”, similar to Model 4 in the non-imputed data, because there are no missing pesticide data in the imputed data and therefore no need for the imputation.

Model 1 contains just the variables sex, age, education, and drinking from an abandoned well. From the non-imputation modeling, these covariates were shown to be of importance and are included in all imputation models. In this model, those who drink from

an abandoned well are seen to have a 2.78 times higher risk of CKDu. Model 2, which adds occupation, shows that farmers have a twofold increased risk of CKDu, however this did not attain statistical significance. Model 3 adds pesticide use, and gives the best c-statistic of all 4 models (0.855). Drinking from an abandoned well has a 2.69 times increased risk of CKDu, and pesticide use shows a 47% increased risk, which is close to significance ( $p=0.11$ ). As per Table 4 in the non-imputed data and Table 8 in the imputed data, occupation, pesticides, and drinking from an abandoned well are correlated. Model 4 includes occpest without occupation or pesticide, which shows a 59% increased risk of CKDu for farmers who use pesticides, which is close to significance ( $p=0.09$ ). Lastly, Model 5, which is comparable to the best fitting model (Model 7) in the non-imputed data, shows a 2.75 times increased risk in drinking from an abandoned well and a 74% increased risk for those that use pesticides. While the c-statistic (0.854) is not higher than imputed Model 3, Model 5 has 4 significant covariates with education close to significant ( $p=0.065$ ) for a 89% increased risk of CKDu in less educated groups. Table 9 below summarizes beta coefficients, standard errors, p-values, ORs and 95% OR confidence intervals for all five logistic regression models with imputed data.

**Table 9. Logistic Modeling with 5 Imputed Datasets from PROC MIANALYZE**

<i>Covariates</i>	<b>Model 1 (c-statistic=0.852)</b>					<b>Model 2 (c-statistic=0.853)</b>				
	<i>B</i>	<i>SE B</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>	<i>B</i>	<i>SE B</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>
<i>Sex</i>	<b>0.74</b>	<b>0.26</b>	<b>0.004</b>	<b>2.11</b>	<b>[1.26, 3.52]</b>	<b>0.81</b>	<b>0.27</b>	<b>0.003</b>	<b>2.32</b>	<b>[1.37, 3.93]</b>
<i>Age</i>	<b>0.11</b>	<b>0.01</b>	<b>&lt;.0001</b>	<b>1.12</b>	<b>[1.08, 1.15]</b>	<b>0.11</b>	<b>0.01</b>	<b>&lt;.0001</b>	<b>1.11</b>	<b>[1.08, 1.15]</b>
<i>Education</i>	0.62	0.34	0.067	1.86	[0.96, 3.62]	0.61	0.35	0.078	1.81	[0.92, 3.57]
<i>Drink from ab. well</i>	<b>1.02</b>	<b>0.42</b>	<b>0.016</b>	<b>2.78</b>	<b>[1.21, 6.35]</b>	<b>1.01</b>	<b>0.43</b>	<b>0.020</b>	<b>2.71</b>	<b>[1.17, 6.29]</b>
<i>Occupation</i>						0.49	0.44	0.278	2.09	[0.88, 4.95]
<i>Pesticide use</i>										
<i>OccPest</i>										
<i>Covariates</i>	<b>Model 3 (c-statistic=0.855)</b>					<b>Model 4 (c-statistic=0.854)</b>				
	<i>B</i>	<i>SE B</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>	<i>B</i>	<i>SE B</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>
<i>Sex</i>	<b>0.65</b>	<b>0.29</b>	<b>0.026</b>	<b>2.01</b>	<b>[1.14, 3.53]</b>	<b>0.65</b>	<b>0.27</b>	<b>0.016</b>	<b>1.91</b>	<b>[1.13, 3.23]</b>
<i>Age</i>	<b>0.11</b>	<b>0.01</b>	<b>&lt;.0001</b>	<b>1.11</b>	<b>[1.08, 1.15]</b>	<b>0.11</b>	<b>0.01</b>	<b>&lt;.0001</b>	<b>1.11</b>	<b>[1.08, 1.15]</b>
<i>Education</i>	0.62	0.35	0.072	1.84	[0.93, 3.64]	0.64	0.34	0.064	1.89	[0.96, 3.69]
<i>Drink from ab. well</i>	<b>0.99</b>	<b>0.44</b>	<b>0.023</b>	<b>2.69</b>	<b>[1.16, 6.26]</b>	<b>1.00</b>	<b>0.42</b>	<b>0.018</b>	<b>2.73</b>	<b>[1.19, 6.27]</b>
<i>Occupation</i>	0.31	0.47	0.519	1.79	[0.72, 4.45]					
<i>Pesticide use</i>	0.47	0.29	0.109	1.53	[0.87, 2.69]					
<i>OccPest</i>						0.46	0.27	0.086	1.59	[0.94, 2.69]
<i>Covariates</i>	<b>Model 5 (c-statistic=0.854)</b>									
	<i>B</i>	<i>SE B</i>	<i>p</i>	<i>OR</i>	<i>95% CI OR</i>					
<i>Sex</i>	<b>0.59</b>	<b>0.27</b>	<b>0.033</b>	<b>1.80</b>	<b>[1.05, 3.07]</b>					
<i>Age</i>	<b>0.11</b>	<b>0.01</b>	<b>&lt;.0001</b>	<b>1.12</b>	<b>[1.08, 1.15]</b>					
<i>Education</i>	0.63	0.34	0.065	1.89	[0.96, 3.70]					
<i>Drink from ab. well</i>	<b>1.01</b>	<b>0.43</b>	<b>0.018</b>	<b>2.75</b>	<b>[1.19, 6.34]</b>					
<i>Occupation</i>										
<i>Pesticide use</i>	<b>0.55</b>	<b>0.27</b>	<b>0.041</b>	<b>1.74</b>	<b>[1.02, 2.95]</b>					
<i>OccPest</i>										

## Discussion

This study attempted to uncover further details about the demographic and lifestyle factors that can lead to the onset of Chronic Kidney Disease of unknown etiology (CKDu). Since this disease affects an estimated tens of thousands of people in Sri Lanka, it is more important than ever to discover what may predispose one to being diagnosed with CKDu. While not an initial aim, this study used analytical techniques to attempt to estimate missing patterns of occupation among the study population in order to test that variable in predicting disease outcome. Although the Medawachchiya renal clinic is near the border of Anuradhapura and Vavuniya districts, almost all study participants were from Anuradhapura district. This may be because while Anuradhapura is a large district primarily used for agriculture, Vavuniya has only 10.8% rice paddy agricultural land use (ESDAC). A possible explanation for the high proportion of females in the control population is that during study recruitment, cases were asked to bring a spouse, family member, or neighbor to be used as a control in the study. The high numbers of females could be representative of the wives of the male cases in the study. Although the entire study population had a greater amount of females, the greater proportion of males to females among cases is consistent with published literature regarding the population most affected by CKDu. The sex covariate was a significant predictor of disease outcome across all models in both imputed and non-imputed data. The best fitting model for imputed data (Model 3) showed that the odds of having CKDu for males were 2.01 times that of females when controlling for other variables. CKDu patients had a higher mean age than healthy individuals, and no significant difference was found between mean age of male and female CKDu patients, supporting findings that indicate CKDu becomes clinically diagnosable in the 5<sup>th</sup>-6<sup>th</sup> decades of life. Across all models with both imputed and non-imputed data, age was a significant covariate in predicting case-control status, and all models



showed an OR of 1.11. The odds of being a CKDu patient are 2.84 times higher for a person who is 10 years older than another, when all other variables are held constant.

When examining the distribution of missingness for the occupation variable in the study population, it can be seen that the majority of missing occupation observations was found in females. Additionally, controls had a substantially higher percentage of missing occupation data when compared to cases. As can be seen in Table 7a and 7b, imputing the data led to an approximately 6% decrease in the proportion of female farmers compared to the non-imputed data. This is attributed to the fact that most participants with missing occupation were female, and imputation predicted the missing values based on values for females in the complete data, who were mostly not farmers. Before imputing, occupation had no effect on case control status in the models. After imputing the missing third of occupation observations with a good concordance statistic in all models, there is little difference in odds ratios for occupation between imputed and non-imputed data, and it remained a non-significant covariate in predicting disease. We can conclude that the missing occupation data is not telling a lot about the study population, even though when imputed it predicts occupation well based on the non-missing data. Instead, other covariates were shown to have greater significance in predicting disease.

Over half of cases were uneducated or had attained a maximum of grade 5 education, while the majority of controls had received a grade 11 education or higher. This finding is consistent with prior literature, and education level can be a reliable predictor of occupation. Most uneducated people in this region are employed in manual labor jobs, including farming. Conversely, those with at least a high school education are qualified for skilled jobs including business, military, and other professional careers. In non-imputed data modeling, low education proved to be a significant predictor of case status in two models (4 and 7), indicating

that the odds of being a CKDu patient were 2.53-2.73 times higher for someone with no or primary education compared with those with senior secondary education and above.

Pesticide use was shown to be a driver of disease in two non-imputation models, demonstrating 72-75% increased risk of CKDu. This finding suggests that pesticide use, and not necessarily being a farmer, can contribute to CKDu risk. An increased risk of CKDu for those who use pesticides, as opposed to the variable farmer/non-farmer, supports the argument against the heat stress hypothesis discussed by Herath et al (Forthcoming 2018) under the assumption that farmers have high exposure to heat. However, we did not have data on heat exposure, and were unable to directly evaluate this hypothesis. The combined variable occpest was modeled to show if being a farmer who used pesticides was a significant predictor of disease. This variable was close to significance in imputation Model 4, which showed a 59% increased risk for farmers who used pesticides.

Lastly, drinking water from a previously abandoned well was shown to be a significant predictor in two models with non-imputed data. There was a 3.18 times increased risk of CKDu for those who reported drinking from an abandoned well. This covariate was significant in every non-imputation model it was included in, and therefore was included in all imputation models, where it remained significant in all models. This finding is important, as drinking from a previously abandoned well can lead to consumption of water with extremely high hardness that may be contaminated with agrochemicals and other heavy metals, as shown by Jayasumana et al (2015). Drinking from abandoned well can contribute to the pesticide exposure that those with CKDu may already be experiencing. Additionally, as more wells in the disease endemic region become abandoned due to unpalatability of water, residents must travel further to collect water to be used for drinking and cooking purposes. This scarcity of

water can contribute to dehydration in those with decreased renal function, and can possibly exacerbate CKD.

The large sample size (N=400) is a strength of the study, as associations between populations could be analyzed using traditional methods, particularly after imputing missing data. The location of the study population, Medawachchiya, is at the epicenter of the CKDu epidemic, and results of the study can help inform future studies for populations at risk. An additional strength of the present study is that it has a population with considerable exposure to hypothesized risk factors. One limitation of this study is the lack of complete data on the duration and types of specific agrochemicals that the population was exposed to. Additionally, there was no matching by age, sex, or education, resulting in large disparities between cases and controls on these variables. This led to difficulty in proper control of these covariates during analysis. Lastly, there was no data on heat stress or dehydration in CKDu cases and control participants, and therefore this hypothesis could not be tested.

## **Conclusions**

This study strongly supports existing literature that show CKDu primarily affects males at a later stage in life. Pesticide use and drinking from an abandoned well were seen to be the primary drivers of disease prediction in both complete and imputed data. These factors should be further investigated in future studies to characterize the risk of using specific types of pesticides and to determine the prevalence and use of abandoned wells in the endemic region. Spatial mapping of abandoned and serving wells should be conducted to determine if a colocation exists between CKDu patients and wells that have been abandoned. Additionally, further chemical analysis of both serving and abandoned wells in the endemic region should be conducted to investigate the types of contaminants CKDu patients are exposed to. Future studies should examine biomarkers of pesticide exposure and kidney damage in CKDu patients to quantify outcome in conjunction with exposure data to further test hypotheses regarding agrochemical exposure and disease.

## References

- Allchin, F. R. (1989). City and State Formation in Early Historic South Asia. *South Asian Studies*, 5(1), 1-16.
- Bandara, J. M., Senevirathna, D. M., Dasanayake, D. M., Herath, V., Bandara, J. M., Abeysekera, T., & Rajapaksha, K. H. (2008). Chronic renal failure among farm families in cascade irrigation systems in Sri Lanka associated with elevated dietary cadmium levels in rice and freshwater fish (Tilapia). *Environ Geochem Health*, 30(5), 465-478. doi:10.1007/s10653-007-9129-6
- Bandara, J. M., Senevirathna, D. M., Dasanayaka, D., Herath, V. Bandara, J. M., Abeysekera, T., & Rajapaksha, K.H. (2008). Chronic renal failure among farm families in cascade irrigation systems in Sri Lanka associated with elevated dietary cadmium levels in rice and freshwater fish (Tilapia). *Environ Geochem Health*, 30(5), 465-478. doi:10.1007/s10653-007-9129-6
- Chandrajith, R., Dissanayake, C. B., Ariyaratna, T., Herath, H. M., & Padmasiri, J. P. (2011). Dose-dependent Na and Ca in fluoride-rich drinking water--another major cause of chronic renal failure in tropical arid regions. *Sci Total Environ*, 409(4), 671-675. doi:10.1016/j.scitotenv.2010.10.046
- Chandrajith, R., Nanayakkara, S., Itai, K., Aturaliya, T. N., Dissanayake, C. B., Abeysekera, T., . . . Koizumi, A. (2011). Chronic kidney diseases of uncertain etiology (CKDue) in Sri Lanka: geographic distribution and environmental implications. *Environ Geochem Health*, 33(3), 267-278. doi:10.1007/s10653-010-9339-1
- Chapter 1: Definition and classification of CKD. (2013). *Kidney Int Suppl* (2011), 3(1), 19-62. doi:10.1038/kisup.2012.64

- Correa-Rotter, R., Wesseling, C., & Johnson, R. J. (2014). CKD of unknown origin in Central America: the case for a Mesoamerican nephropathy. *Am J Kidney Dis*, *63*(3), 506-520. doi:10.1053/j.ajkd.2013.10.062
- De Silva, P., Abdul, K. S., Eakanayake, E. M., Jayasinghe, S. S., Jayasumana, C., Asanthi, H. B., . . . Siribaddana, S. H. (2016). Urinary Biomarkers KIM-1 and NGAL for Detection of Chronic Kidney Disease of Uncertain Etiology (CKDu) among Agricultural Communities in Sri Lanka. *PLOS Neglected Tropical Diseases*, *10*(9). doi:10.1371/journal.pntd.0004979
- El Minshawy, O. (2011). End Stage Renal Disease in El-Minia Governorate, Egypt: Data of the Year 2007. *Nephro-Urology Monthly*, *3*(2), 118-121. Retrieved from [http://numonthly.portal.tools/?page=article&article\\_id=1064](http://numonthly.portal.tools/?page=article&article_id=1064)
- FAO International Code of Conduct on the Distribution and Use of Pesticides Sri Lanka Report. , 2005, International Code of Conduct on the Distribution and Use of Pesticides Sri Lanka Report, [www.fao.org/docrep/008/af340e/af340e0k.htm#bm20](http://www.fao.org/docrep/008/af340e/af340e0k.htm#bm20).ESDAC. (n.d.). Vavuniya District. Land Use. Retrieved from <https://esdac.jrc.ec.europa.eu/content/vavuniya-district-land-use>
- Glaser, J., Lemery, J., Rajagopalan, B., Diaz, H. F., Garcia-Trabanino, R., Taduri, G., . . . Johnson, R. J. (2016). Climate Change and the Emergent Epidemic of CKD from Heat Stress in Rural Communities: The Case for Heat Stress Nephropathy. *Clin J Am Soc Nephrol*, *11*(8), 1472-1483. doi:10.2215/CJN.13841215
- Herath, C., Jayasumana, C., De Silva, P.M.C.S., De Silva, P.H.C., Siribaddana, S., De Broe, M.E. (Forthcoming 2018). Kidney Diseases in Agricultural Communities: A Case Against Heat-Stress Nephropathy. *Kidney International Reports*.

- International Expert Consultation on Chronic Kidney Disease of Unknown Etiology*. (2016). Retrieved from WHO South-East Asia Regional Office:  
[http://www.searo.who.int/srilanka/documents/report\\_international\\_expert\\_consultation\\_on\\_ckdu.pdf](http://www.searo.who.int/srilanka/documents/report_international_expert_consultation_on_ckdu.pdf)
- Jayasumana, C., Fonseka, S., Fernando, A., Jayalath, K., Amarasinghe, M., Siribaddana, S., . . . Paranagama, P. (2015). Phosphate fertilizer is a main source of arsenic in areas affected with chronic kidney disease of unknown etiology in Sri Lanka. *Springerplus*, *4*, 90. doi:10.1186/s40064-015-0868-z
- Jayasumana, C., Gajanayake, R., & Siribaddana, S. (2014). Importance of Arsenic and pesticides in epidemic chronic kidney disease in Sri Lanka. *BMC Nephrol*, *15*, 124. doi:10.1186/1471-2369-15-124
- Jayasumana, C., Orantes, C., Herrera, R., Almaguer, M., Lopez, L., Silva, L. C., . . . De Broe, M. E. (2017). Chronic interstitial nephritis in agricultural communities: a worldwide epidemic with social, occupational and environmental determinants. *Nephrol Dial Transplant*, *32*(2), 234-241. doi:10.1093/ndt/gfw346
- Jayatilake, N., Mendis, S., Maheepala, P., Mehta, F. R., & Team, C. K. N. R. P. (2013). Chronic kidney disease of uncertain aetiology: prevalence and causative factors in a developing country. *BMC Nephrol*, *14*, 180. doi:10.1186/1471-2369-14-180
- Kamel, E. G., & El Minshawy, O. (2010). Environmental Factors Incriminated in the Development of End Stage Renal Disease in El-Minia Governorate, Upper Egypt. *International Nephrology and Urology*, *2*(3), 431-437.
- Ministry of Health. Data presented at the presidential task force for prevention of kidney diseases. Colombo, Sri Lanka. Presidents house. 6 June 2014

- Nanayakkara, S., Senevirathna, S. T., Abeyssekera, T., Chandrajith, R., Ratnatunga, N., Gunarathne, E. D., . . . Koizumi, A. (2014). An integrative study of the genetic, social and environmental determinants of chronic kidney disease characterized by tubulointerstitial damages in the North Central Region of Sri Lanka. *J Occup Health*, *56*(1), 28-38.
- Peiris, D. (2014). Sri Lanka: Transforming into a Middle Income Country. Retrieved from <http://www.worldbank.org/en/results/2014/04/16/sri-lanka-transforming-middle-income-country>
- Redmon, J. H., Elledge, M. F., Womack, D. S., Wickremashinghe, R., Wanigasuriya, K. P., Peiris-John, R. J., . . . Levine, K. E. (2014). Additional perspectives on chronic kidney disease of unknown aetiology (CKDu) in Sri Lanka--lessons learned from the WHO CKDu population prevalence study. *BMC Nephrol*, *15*, 125. doi:10.1186/1471-2369-15-125
- Roncal Jimenez, C. A., Ishimoto, T., Lanasa, M. A., Rivard, C. J., Nakagawa, T., Ejaz, A. A., . . . Johnson, R. J. (2014). Fructokinase activity mediates dehydration-induced renal injury. *Kidney Int*, *86*(2), 294-302. doi:10.1038/ki.2013.492
- USAID. (2009, November). *Assessment of Horticulture in Eastern, Uva, and North Central Provinces of Sri Lanka* (Rep.). Retrieved [https://pdf.usaid.gov/pdf\\_docs/Pnadt118.pdf](https://pdf.usaid.gov/pdf_docs/Pnadt118.pdf)
- Wanigasuriya, K. (2014). Update on uncertain etiology of chronic kidney disease in Sri Lanka's north-central dry zone. *MEDICC Rev*, *16*(2), 61-65.
- Wasana, H. M., Aluthpatabendi, D., Kularatne, W. M., Wijekoon, P., Weerasooriya, R., & Bandara, J. (2016). Drinking water quality and chronic kidney disease of unknown etiology (CKDu): synergic effects of fluoride, cadmium and hardness of water. *Environ Geochem Health*, *38*(1), 157-168. doi:10.1007/s10653-015-9699-7



- Weaver, V. M., Fadrowski, J. J., & Jaar, B. G. (2015). Global dimensions of chronic kidney disease of unknown etiology (CKDu): a modern era environmental and/or occupational nephropathy? *BMC Nephrol*, *16*, 145. doi:10.1186/s12882-015-0105-6
- Wesseling, C., Aragon, A., Gonzalez, M., Weiss, I., Glaser, J., Rivard, C. J., . . . Johnson, R. J. (2016). Heat stress, hydration and uric acid: a cross-sectional study in workers of three occupations in a hotspot of Mesoamerican nephropathy in Nicaragua. *BMJ Open*, *6*(12), e011034. doi:10.1136/bmjopen-2016-011034
- WHO highlights need for countries to scale up action on noncommunicable diseases. . (2014). Retrieved from <http://www.who.int/mediacentre/news/notes/2014/action-on-ncds/en/>
- Wijkstrom, J., Jayasumana, C., Dassanayake, R., Priyawardane, N., Godakanda, N., Siribaddana, S., . . . Wernerson, A. (2018). Morphological and clinical findings in Sri Lankan patients with chronic kidney disease of unknown cause (CKDu): Similarities and differences with Mesoamerican Nephropathy. *PLoS One*, *13*(3), e0193056. doi:10.1371/journal.pone.0193056