InJAST
Volume 5, Number 2, October 2024
Pages: 65-83

P - ISSN: 2722 - 0133
E - ISSN: 2722 - 0141
DOI: 10.33751/injast.v5i2.10690

GIS-based analysis of water quality risk factors and CKDu prevalence in Northern Yobe State, Nigeria

MODIBBO BABAGANA-KYARI*, **NURA ALHAJI YARO, KABIRU M. YAKASAI**

Department of Environmental Health and Safety, School of Health Sciences, Maryam Abacha American University of Niger, Maradi

*Corresponding author: *modibbobgk@gmail.com*

Submitted 5 September 2024; Accepted 4 November 2024

ABSTRACT

Chronic Kidney Disease (CKDu) presents a major public health challenge in Northern Yobe State, Nigeria, particularly in the Bade community, where water quality is overly suspected to influence its prevalence. This study employs a Geographic Information Systems (GIS)-based framework to analyze the spatial distribution of CKD in relation to water quality parameters. Advanced spatial analysis techniques, including hexagonal tessellation and Moran's I for spatial autocorrelation, were utilized alongside community-based surveys conducted using Kobotoolbox and Qfield applications to map CKD hotspots. The Moran's I Index of 0.1046, with a z-score of 4.9546 and a p-value of 0.000001, indicates significant clustering of CKD cases rather than random distribution across the study area. Water samples from 30 water facilities, with 10 from each classified hotspot, were analyzed for nephrotoxic heavy metals, ionic concentrations, and hardness using an Atomic Absorption Spectrometer (Model 210V-GP). The spatial distribution of these parameters was modeled using Inverse Distance Weighting (IDW) interpolation in ArcGIS Pro 3.4. Descriptive statistics, hazard index calculations, and Water Quality Index (WQI) assessments were conducted, with box plots facilitating comparative analysis across High, Medium, and Low disease areas. ANOVA and Tukey's HSD post-hoc tests were performed to compare specific water parameters between the disease hotspots and water facilities. The results revealed elevated concentrations of nephrotoxic heavy metals in high-disease hotspots, with maximum values observed for Arsenic (0.21 mg/L), Cadmium (0.30 mg/L), Lead (0.23 mg/L), Chromium (0.50 mg/L), and Fluoride (55 mg/L). Additionally, Nitrite and Nitrate levels exhibited high Hazard Quotients, all surpassing WHO guidelines for safe drinking water. These findings underscore the potential health risks posed by these contaminants in affected areas. Results demonstrate a significant link between the prolonged use of handpumps water and high prevalence of chronic kidney disease incidence (CKD) among affected households. Additionally, the study identified strong spatial correlations between CKD incidence and high concentrations of nephrotoxic heavy metals in water from handpumps, providing critical insights for targeted public health interventions and guiding future research efforts.

ABSTRAK

Penyakit Ginjal Kronis (CKDu) menghadirkan tantangan kesehatan masyarakat yang besar di Negara Bagian Yobe Utara, Nigeria, khususnya di komunitas Bade, di mana kualitas air diduga sangat mempengaruhi prevalensinya. Penelitian ini menggunakan kerangka kerja berbasis Sistem Informasi Geografis (GIS) untuk menganalisis distribusi spasial CKD dalam kaitannya dengan parameter kualitas air. Teknik analisis spasial tingkat lanjut, termasuk tesselasi heksagonal dan Moran's I untuk autokorelasi spasial, digunakan bersamaan dengan survei berbasis komunitas yang dilakukan menggunakan aplikasi Kobotoolbox dan Qfield untuk memetakan hotspot CKD. Indeks Moran's I sebesar 0,1046, dengan skor z sebesar 4,9546 dan nilai p sebesar 0,000001, menunjukkan pengelompokan kasus CKD yang signifikan dibandingkan distribusi acak di seluruh wilayah penelitian. Sampel air dari 30 fasilitas air, dengan 10 dari masing-masing hotspot yang diklasifikasikan, dianalisis logam berat nefrotoksik, konsentrasi ionik, dan kekerasan menggunakan Spektrometer Serapan Atom (Model 210V-GP). Distribusi spasial parameter ini dimodelkan menggunakan interpolasi Inverse Distance Weighting (IDW) di ArcGIS Pro 3.4. Statistik deskriptif, penghitungan indeks bahaya, dan penilaian Indeks Kualitas Air (WQI) dilakukan, dengan plot kotak yang memfasilitasi analisis komparatif di wilayah dengan penyakit Tinggi, Sedang, dan Rendah. Uji post-hoc ANOVA dan HSD Tukey dilakukan untuk membandingkan parameter air spesifik antara titik panas penyakit dan fasilitas air. Hasilnya menunjukkan peningkatan konsentrasi logam berat nefrotoksik di titik rawan penyakit, dengan nilai maksimum yang diamati untuk Arsenik (0,21 mg/L), Kadmium (0,30 mg/L), Timbal (0,23 mg/L), Kromium (0,50 mg/L).), dan Fluorida (55 mg/L). Selain itu, tingkat Nitrit dan Nitrat menunjukkan Tingkat Bahaya yang tinggi, semuanya melampaui pedoman WHO untuk air minum yang aman. Temuan ini menggarisbawahi potensi risiko kesehatan yang ditimbulkan oleh kontaminan ini di daerah yang terkena dampak. Hasilnya menunjukkan adanya hubungan yang signifikan antara penggunaan air pompa tangan dalam waktu lama dan tingginya prevalensi kejadian penyakit ginjal kronis (CKD) di antara rumah tangga yang terkena dampak. Selain itu, penelitian ini mengidentifikasi korelasi spasial yang kuat antara kejadian CKD dan tingginya konsentrasi logam berat nefrotoksik dalam air dari pompa tangan, memberikan wawasan penting untuk intervensi kesehatan masyarakat yang ditargetkan dan memandu upaya penelitian di masa depan.

Keywords: *CKDu, CKD Hotspots, Disease mapping, Hazard Quotient, Voronoi tessellation, GIS, spatial Analysis, Nephrotoxicity*

INTRODUCTION

Chronic Kidney Disease of uncertain etiology (CKDu) has been a growing concern in various parts of the world, particularly in agricultural communities (Weaver, Fadrowski, & Jaar, 2015; Priyadarshani et al., 2023). The disease is characterized by its occurrence in individuals without the typical risk factors associated with Chronic Kidney Disease (CKD), such as diabetes or hypertension (Lunyera et al., 2016; Gifford et al., 2017). Despite this increase, the specific environmental and lifestyle-related risk factors contributing to CKDu remain poorly understood in many of the areas where the disease is so prevalent (Wesseling et al., 2013; Weaver et al., 2015). In regions like Northern Yobe, Nigeria, CKDu has become a significant public health issue, with a rising number of cases reported over the past decade based on hospital records observation (Sulaiman et al., 2019; Goni et al., 2024).

The geographical and environmental context of Northern Yobe State, characterized by its agricultural economy and reliance on groundwater, suggests a potential link between environmental exposures and CKDu. Preliminary studies in the CKD hotspots of Yobe region implicated groundwater and food resources grown in the area (Waziri et al., 2017; Gashua et al., 2018; Oyekanni et al., 2018; Ahmed et al., 2018) as heavy metals were found exceeded benchmark standards. However, the recent study of conducted by Aminu et al., (2022) have observed that heavy metals were found in biomedical samples of the patients diagnosed with CKD in the area. Similarly, Yuguda et al., (2022) study reported high concentrations of Cadmium, Lead and mercury exceeding permissible limits at all the sites where samples were collected.

Several studies have implicated water quality, particularly contamination by heavy metals, agrochemicals, and other pollutants, as a potential contributor to CKDu in similar settings (Wanigasuriya, 2012; Jayatilake et al., 2013). However, the link between these environmental risk factors, particularly water and the prevalence of CKDu in Northern Yobe State has not been systematically explored in spatial context despite the number of different study attempts.

Remarkably, Geographical Information Systems (GIS) provides a powerful tool for exploring the spatial dimensions of environmental health problems. GIS allows for the integration of various data sources (Saputro et al., 2023; Priatna & Monk, 2023), including environmental, health, and demographic data, to identify spatial patterns and relationships that might not be apparently seen through traditional epidemiological methods (Huang et al., 2021). In recent time, GIS has been very used to explore the spatial distribution of diseases and their potential associated environmental factors, providing deep insights that is not known (Oviasu, 2012; Sanati, 2015; Senanayake, 2016).

In the context of CKDu, GIS has the potential to uncover spatial correlations between disease prevalence and environmental factors such as water quality. Greatly mapping CKDu incidences and overlaying these with environmental data, one can identify potential hotspots and explore the geographic distribution of risk factors and their concentration as well. This approach has been successfully applied in other regions affected by CKDu, where it has helped to identify areas of high disease burden and potential environmental exposures (Rodriguez et al., 2013; Jayasumana et al., 2015; Vlahos et al., 2021).

The novelty of this study lies in its application of a community-centric approach, utilizing GIS to identify and map CKDu incidences at the household level in Northern Yobe State. Unlike traditional hospital-based studies, which may miss cases in remote or underserved areas, this approach involves direct engagement with communities to gather data on CKDu prevalence and potential environmental risk factors. Through integrating community surveys with GIS-based spatial analysis, the study seeks to offer a more thorough understanding of water risk factors and chronic kidney disease of unknown etiology (CKDu) prevalence in the region. This approach addresses the contentious hypothesis that has yet to be conclusively proven. This study is significant as it represents one of the first attempts to systematically investigate the relationship between water quality and CKDu in Northern Yobe State using a GIS-based approach.

METHODS

Study Area

The chosen study region encompasses the entirety of Northern Yobe State, with a specific focus on Bade Local Government Area. Bade Local Government Area is situated in the northeastern region of Nigeria, characterized by its geographical coordinates at latitude 12° 52' 25.12 " N and longitude 11° 2' 49.94" E respectively (Figure 1). This area shares its borders with Nguru Local Government Area to the north, as well as Bursari and Jakusko Local Government Areas to the east and west respectively. Notably, Bade LGA emerges as a significant focal point for this research due to its higher prevalence of Chronic Kidney Disease of Unknown etiology (CKDu). Due to the heightened prevalence of CKDu makes Bade LGA, particularly Gashua town, a key area for the research. Bade Local Government Area, holds a strategic position at the convergence of the Hadeija, and Jama'are rivers, which combine to form the Yobe River. This river system eventually joins the Kumadugu-Gana River at Damasak and flows into Lake Chad. This extensive water network serves as a vital water source for domestic and agricultural needs, supporting fisheries and sustenance for numerous local communities.

The region's soil fertility is enhanced by its loamy-clay and silty soils, fostering paddy rice cultivation that plays a pivotal role in the livelihoods of the local population. Traditionally, the area is primarily agrarian, with farming and fishing being the dominant occupations due to the abundance of accessible open and ground waters, often found at shallow depths in floodplains. A noteworthy aspect of this agricultural landscape is the practice of paddy rice cultivation, which has been a central occupation for the community for nearly four decades. This enduring commitment to paddy rice cultivation underscores its significance in the lives of the locals. The long-standing tradition of rice farming highlights the area's connection to its agricultural heritage and signifies the pivotal role that agriculture plays in the socio-economic fabric of the region. The area maintains an average elevation of approximately 370 meters above sea level. Its geological composition aligns with the broader geological features of the Lake Chad region. The geological makeup of the region can be broadly categorized into three main groups: a) Crystalline Basement Complex of Pre-Cambrian origin, b) sedimentary Chad Formation dating to the Tertiary and Quaternary periods, and c) Quaternary-age alluvium and aeolian sands, as reported by Alkali (1995).

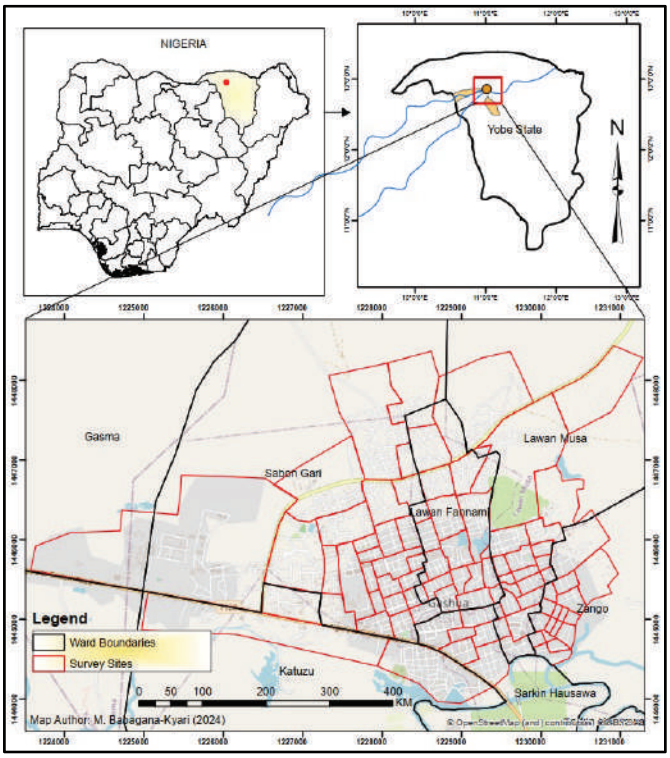


Figure 1. Study area map in northern Yobe State, Nigeria.

The local climate of the area adheres to the tropical pattern prevalent in northern Nigeria, characterized by distinct wet and dry seasons. Jajere et al., (2018) have identified three primary seasons within the study area: The hot-dry season spanning April to June, and the warm-moist season set in July to September, while the cold-dry season prevailing from December to February. As for Gashua's average temperature, it registers at 43 degrees Celsius.

Study Design

This study employs a descriptive, hotspot-based cross-sectional design to collect data on CKD prevalence directly from the community through household surveys of affected individuals. The primary aim is to identify CKD hotspots within Gashua main town, targeting areas with elevated incidences of the disease, encompassing both morbidity and mortality cases. Gashua town was purposively selected due to the high endemicity of CKD (Sulaiman et al., 2019). This approach allows for a comprehensive analysis of the relationship between groundwater quality, various confounding factors, and CKD occurrences (Figure 2).

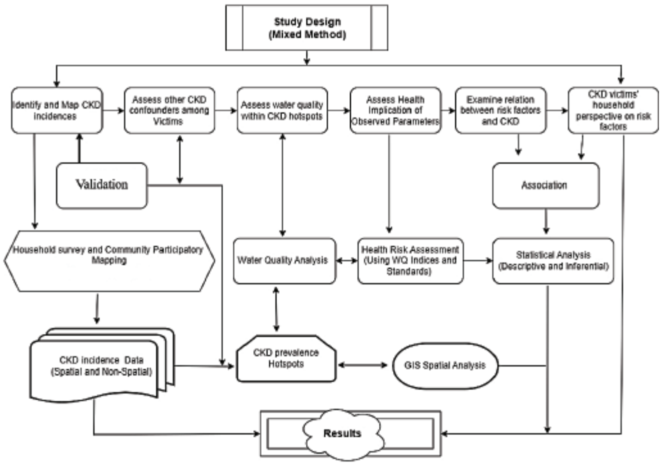


Figure 2. Methodology flow chart.

The choice of this design is particularly significant given the lack of geocoded hospital records, which are often crucial for accurate spatial analysis of disease prevalence (Oliver et al., 2005). While geocoded hospital data can offer valuable insights into disease distribution and hotspots, its absence in this context necessitates alternative methods for capturing localized disease incidences (Luby et al., 2015; Xie et al., 2017). Previous studies have demonstrated that community-based surveys and participatory mapping can effectively substitute for hospital records in areas where such data are unavailable (Fornace et al., 2018). Thus, through adopting a community-centric approach, this study underscores the importance of local factors and perspectives in understanding disease patterns in the region.

Data Collection

Incidence survey techniques

A community-based survey was employed using Kobotoolbox for survey design and deployment, and QField application for in-field data verification, GPS mapping. Community engagement included participatory and house-to-house surveys to identify households with confirmed chronic kidney disease (CKD) cases (mortality and morbidity) and collect detailed information on household water sources and usage patterns. Only cases confirmed through health facilities were included in the survey. During the survey, medical records (documents) of the individuals were requested for verification. Six field workers were recruited and trained specifically for the house-to-house survey, which was carried out for a period of 28 days in the study area after which the survey was halted.

Additionally, a reconnaissance was conducted to pretest the instrument in the Lawan Fannami ward. The electronic survey instrument captured both location and key sociodemographic variable of CKD patients/victims and household water usage pattern. Prior to survey, ethical approval was obtained first from Yobe State Ministry of Health's Research ethics committee. Moreover, informed consent was obtained from each study participant, with the consent process outlined at the beginning of the questionnaire which clearly explain the nature and purpose of the study. Participants were required to consent with either 'Yes' or 'No' before participating in the exercise.

A snowball sampling method was employed, with the assistance of neighborhood heads (Community Leaders), to identify households with CKD cases. The Enumeration Area of DLI 11.3 was utilized for the incidence survey to prevent duplication in data collection, as each fieldworker was assigned a specific spatial unit. Information was primarily provided by household heads and patients' attendants in households as majority of the victims are deceased while some are medically not fit to actively participate in the survey. The survey identified 441 medically confirmed cases across 430 households, accounting for instances where multiple cases were present in some households. To prevent duplication, field workers were assigned with coded Enumeration Areas (EAs) based on the 2022 Yobe State Geographic Information Service property enumeration area shapefile for Disbursement link indicator (DLI 11.3) project as depicted in Figure 1.

Water sample collection techniques

Water samples were systematically collected within the observed CKD hotspots and subjected to laboratory analysis for nephrotoxic heavy metals (Arsenic, Cadmium, Chromium, Mercury, Lead), key ions (Fluoride, Sodium, Nitrite, Nitrate, Phosphate), and water hardness indicators (Magnesium and Calcium). Additional parameters, including pH and turbidity, were also measured to provide a comprehensive assessment of water quality.

A hotspot-based stratified random sampling method was employed to ensure representative sampling across the three distinct CKD hotspot levels high, medium, and low incidence areas. From each hotspot category, 10 water facilities were randomly selected, resulting in a total of 30 water facilities being sampled. This approach ensured a balanced distribution of samples across different levels of disease prevalence, capturing variability in water quality across the identified areas. The sampling included a range of water facilities, such as hand pumps, deep boreholes, and tap stands, encompassing both public and private sources, except for newly constructed facilities.

To ensure the integrity and reliability of the samples, standard procedures were rigorously followed during collection. A total of 90 water samples were collected in triplicate from the 30 selected water points using clean plastic bottles. Samples were obtained during periods of active use, specifically after the facilities had been in continuous operation for over 5 hours. The samples were carefully tagged and promptly transported to Yobe State University Chemistry Research Laboratory, where they were analyzed under controlled conditions to maintain their integrity.

Geographic Information System (GIS) analysis techniques

The study employed a GIS-based methodology to explore the geographic distribution of CKD cases in relation to water quality parameters. To assess spatial autocorrelation, Moran's I was utilized, a statistical technique that identifies clustering or dispersion within spatial data, enabling the examination of disease prevalence patterns. To ensure uniform sampling, a hexagonal tessellation covering a total area of 20,000 square meters was applied across the study region. Each hexagon served as a sampling unit for surveying CKD prevalence, enabling a spatial join between CKD incidence data and the tessellated polygons. This method facilitated the quantification of disease counts per unit area.

Through the GIS analysis, three distinct hotspots were identified based on CKD prevalence and classified into high, medium, and low prevalence areas. A choropleth map was subsequently generated to visually represent the spatial distribution of CKD prevalence, using a red color gradient to differentiate the intensity of the disease across the study area. For the selected water points, Voronoi tessellation was employed to define the presumed service areas of the water facilities (see, Figure 11). This tessellation allowed for estimating the regions from which nearby residents are likely to collect water for domestic use.

Water quality analysis

Water samples collected from the 30 water points across the three identified CKD hotspots were analyzed to determine the concentration of nephrotoxic heavy metals, key ions, and water hardness indicators. An Atomic Absorption Spectrometer (Model 210V-GP) was employed for the analysis of these samples. The data obtained from these analyses were mapped and compared with CKD prevalence so as to identify potential environmental risk factors in water contributing to the disease.

Spatial modeling and data analysis

Spatial modeling and data analysis were conducted using ArcGIS Pro 3.4, where Inverse Distance Weighting (IDW) interpolation with a geometric interval classifier was used to model the spatial distribution of the water quality parameters. Descriptive statistics, including minimum, maximum, and mean values of the water samples for each hotspot, were calculated and compared against benchmark standards set by the World Health Organization (WHO) and the U.S. Environmental Protection Agency (EPA). ANOVA, Tukey's HSD post-hoc tests were conducted to identify specific differences between pairs of hotspots particularly high and low hotspots. For the health risk assessment, Hazard Quotient (HQ) was calculated to evaluate the potential risk posed by contaminants. The Hazard Quotient (HQ) for each heavy metal and ion was calculated using the following formula developed by United State Environmental Protection agency (EPA):

$$HQ = \frac{ED}{RfD}$$

Where:
ED = Exposure Dose (Contaminant Concentration values)
RfD = Reference Dose (the safe contaminant exposure limit provided by WHO or EPA (see, Appendix 1 & 2)

HQ<1 : The exposure is considered to be within the safe limit, and adverse health effects are unlikely
HQ≥1 : The exposure exceeds the safe limit, indicating a potential health risk. The higher the HQ value, the greater the risk of adverse health effects.

Using the formula, the Hazard Quotient (HQ) of each contaminant was computed. The average HQ for each contaminant was then calculated to obtain the overall average HQ for each of hotspot using the formula:

$$\text{Average conc. for Hotspot} = \frac{\sum \text{Concentration Values}}{\text{Number of Samples}}$$

These average concentrations of each contaminant within each hotspot were used to compute the HQ, facilitating a comparative analysis through chart depicted in Figure 15. Additionally, the Water Quality Index (WQI) was computed to provide an overall assessment of water quality across the hotspots. A comparative analysis of the observed water quality parameters was performed using box plots, while the water usage characteristics and duration of residency of affected households were examined through the use of heat map. This integrated analysis provided deeper insights into potential correlations between water quality and CKD prevalence in the studied community.

RESULTS

This section presents the study's findings, with a focus on the fundamental characteristics of households impacted by chronic kidney disease (CKD) and the individuals residing within them. Key variables highlighted include the geographical location of the affected households and the duration of stay of the CKD victims within these households.

Victims' Household Characteristics

Figure 3 illustrates the distribution of responses to the consent question posed during the survey. The question was: "Consent for participation: If you are willing to participate in this survey, please respond 'Yes.' If you are not willing to participate, please respond 'No.' You are free to answer either way without any penalty".

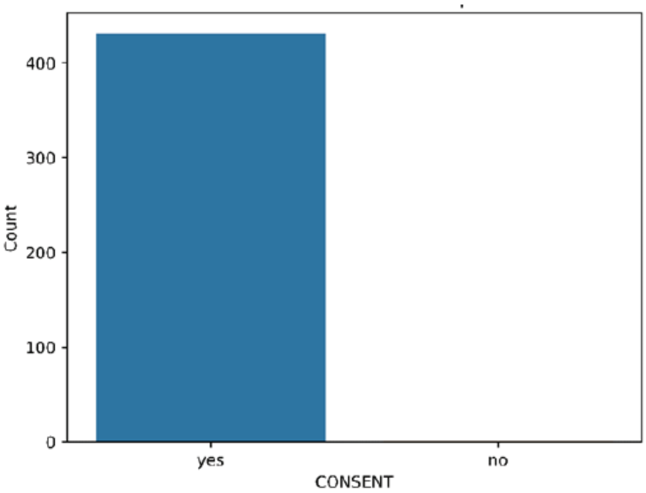


Figure 3. Distribution of study consent responses.

The 100% "Yes" responses indicate unanimous willingness among the respondents to participate in the CKD survey. Specifically, 431 out of 430 respondents expressed their consent by answering "Yes," with only one individual declining to participate. This demonstrates a high level of engagement and willingness to contribute to the study.

CKD Incidence Distribution in Household

The distribution of CKD incidence per household is illustrated to show how widespread the disease is within the study area. Table 1 presents the distribution of CKD incidences from the surveyed households. In the study area, 92.6% of households have one incidence of CKD, 2.46% have two incidences, and 4.7% have no incidences of CKD.

The results indicate that CKD is quite prevalent in the study area, with most households having at least one affected member. Specifically, 96.3% of households had only one CKD victim, while 2.46% had multiple cases. This suggests a need for targeted interventions for single-case households, while the presence of multiple cases may point to potential genetic or environmental factors.

Table 1. Number of CKD incidence in household in the study area.

Options	Frequency	Percentage (%)
1 incidence	414	92.6
2 incidences	16	2.46
None	11	4.7
Total	**441**	**100.0**

Source: Researchers' Fieldwork (2024)

Medical Confirmation of Household CKD Incidence

This section examines the responses to the questionnaire item regarding whether the household CKD incidence is medically confirmed. This information is critical for assessing the reliability of reported CKD cases and evaluating the potential for underreporting. Responses are categorized as "Yes" (medically confirmed) or "No" (not medically confirmed) to clarify the verification status of CKD cases within households (Table 2).

Table 2. Medical confirmation status of the incidence in household.

	Frequency	Percentage (%)
Yes	433	98.1
No	8	1.8
Total	**441**	**100**

Source: Researchers' Fieldwork (2024)

The result indicates that out of 433 CKD cases surveyed from 430 households, 98.1% (425 cases) were medically confirmed, while only 1.8% (8 cases) were not. This high percentage of confirmed cases indicates that most individuals had their CKD diagnosis validated by healthcare facilities, with physicians confirming the condition in most instances. The high level of medical verification is further supported by the fact that many households provided medical report documents to the research assistants during the survey

Respondent Status in Household

This section offers insight into the membership status of respondents within the household, with Figure 4 providing a visual breakdown of respondent statuses.

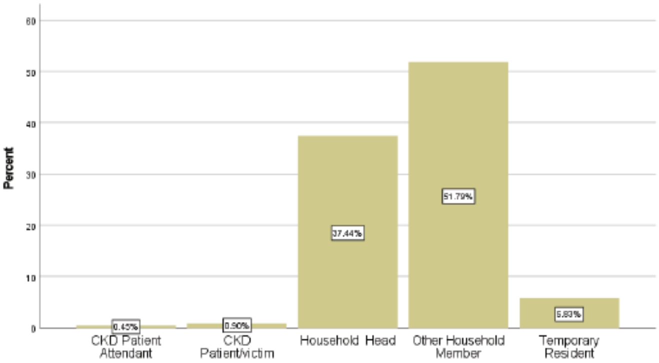


Figure 4. Household membership status of household respondents.

Notably, 37.44% of survey participants were household heads, while 51% were other household members. CKD patients comprised only 0.90% of respondents, suggesting a low representation likely due to mortality or morbidity. Nonetheless, household heads are considered reliable sources for information on household CKD incidence due to their comprehensive understanding of household affairs.

Age Group of the Respondent in Household

This section presents the age group distribution of respondents from households affected by the rampant incidence of chronic kidney disease (CKD). Figure 5 illustrates this distribution. It is noteworthy that the majority of the respondents fall within the 35-44 years age group, followed closely by those in the 25-34 years age group. The next largest group is the 45-54 years age group.

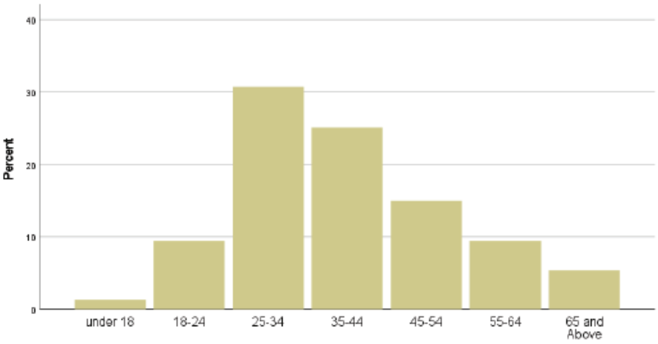


Figure 5. Age group distribution of the respondent.

Understanding the age distribution of respondents is crucial for assessing the quality of the information collected in the context of the rampant CKD incidence survey. Notably, respondents under the age of 18 years constituted less than 5% of the respondent, indicating a low representation of this age group.

Educational Level of Respondent

The part displays the various categories of the educational levels of respondents from surveyed household. Table 3 illustrate the educational profile of respondents who reported the incidence of household CKD. The data shows a variety of educational levels ranging from 'High Secondary School' to 'Postgraduate degree'. The majority of respondents have a high secondary school education (51.8%, 231). This is followed by those with informal education (16.8%, 72) and some college or bachelor's degree (15.7%, 68). Lesser percentages are noted in primary school education (6.5%, 28), Islamic education (4.4%, 19), and postgraduate degrees (2.8%, 12), highlighting a diverse educational background among the surveyed individuals. The total count of respondents is 430, accounting for 100% of the data presented in the table.

Table 3. Educational level of the respondent.

Category	Frequency	Percentage (%)
High Secondary School	231	51.8
Informal School	75	16.8
Some college / Bachelor's degree	70	15.7
Primary School	29	6.5
Islamic education	18	4.0
Postgraduate degree	7	1.6
Total	**430**	**100.0**

Source: Researchers' Fieldwork (2023)

The above insights indicate that a significant percentage of respondents were literate, suggesting they fully understood the questionnaire used. Relatively, with 51.8% having completed high secondary school and 15.7% having some college or a bachelor's degree, the majority have a solid educational foundation. Additionally, the presence of respondents with postgraduate degrees (2.8%) underscores their literacy and comprehension levels.

Gender of the Respondent

The gender of the respondent also provides valuable insights into the study, indicating who predominantly participated in the survey. The Table 4. illustrate the distribution of the gender of the respondents. From the data it can be seen that 81.6% were male while 14.8% were female. This because the majority of the persons participated in the study were house head (HH) who were Males.

Table 4. Gender distribution of the household respondents.

Category	Frequency	Percentage (%)
Male	364	81.6
Female	66	14.8
Total	**430**	**100.0**

Source: Researchers' Fieldwork (2023)

The gender distribution of respondents in the CKD incidence survey provides significant context for understanding the demographics of those reporting household CKD incidences. Figure 6 shows that a significant number of residents have lived in their current households for 6 to 10 years, as indicated by the highest percentage (38.57%).

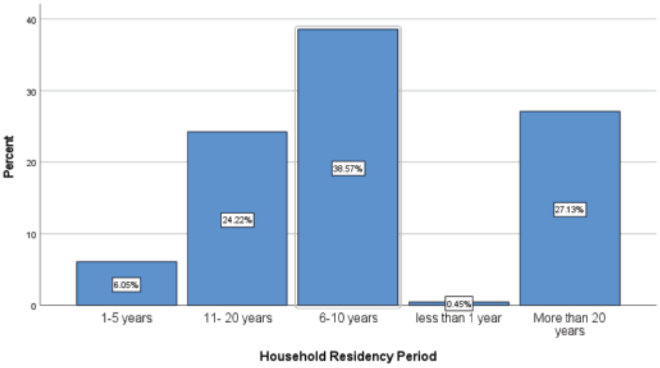


Figure 6. CKD victims' Household period of residency.

Additionally, a substantial percentage of households have been residing in the area for more than 20 years (27.13%). Conversely, households that have spent 1-5 years (6.05%) and less than 1 year (0.45%) in their current residences have lower percentages. This deep insight clearly highlights the importance of considering residency duration when assessing CKD risk. The relationship between the length of residency and CKD risk can be analyzed further to understand the impact of environmental factors on the development of the disease.

Spatial Prevalence of CKD

This section examines the spatial distribution of CKD incidences within the study area through the application of Geographic Information Systems (GIS). It includes a detailed map illustrating the hotspots of CKD occurrences in Northern Yobe State and evaluates the hypothesis concerning the spatial distribution patterns of CKD incidences using spatial statistical analysis. Figure 7 presents map illustrating the distribution pattern of CKD incidence within the study area. The point features on the map represent the locations of households where CKD victims reside. Global Moran's I index was employed to analyze the spatial patterns of CKD incidence within the study area. Moran's I is a measure of spatial autocorrelation, which determines whether the distribution pattern of spatial object is clustered, dispersed, or random.

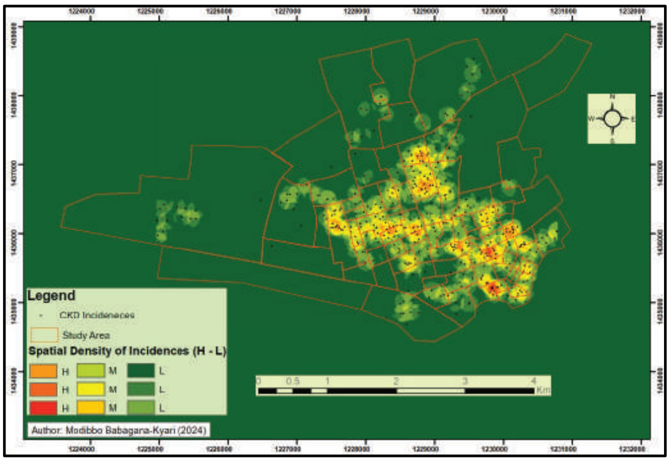
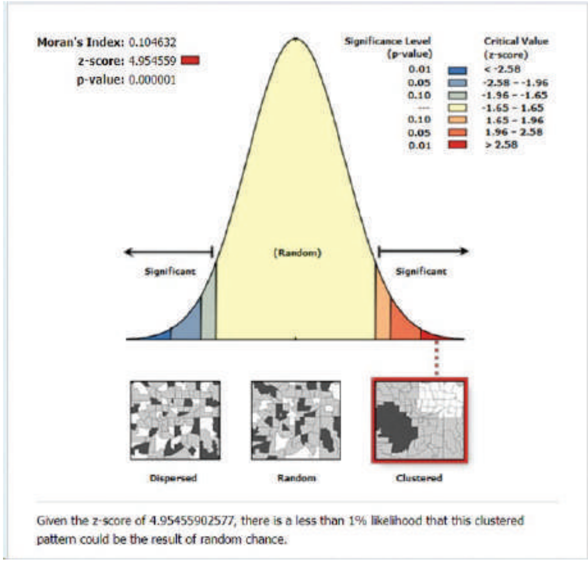


Figure 7. Map display of CKD hotspots in northern Yobe State.

The map uses a color gradient to represent the density of CKD incidences, ranging from green (indicating low density) to red (indicating high density). The most intense hotspots, represented by red, are predominantly located in the central part of the study area, indicating a significantly higher concentration of CKD incidences. Surrounding these high-density zones are areas of medium density (yellow), suggesting a moderate level of CKD incidences, while the green regions, denoting low-density CKD incidence, are primarily situated on the periphery of the map

Spatial autocorrelation statistics

Figure 8 below illustrates the spatial statistics regarding the spatial autocorrelation Moran's I index, applied to understand the patterns of Chronic Kidney Disease (CKD) incidences in the area.



Given the z-score of 4.95455902577, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

Figure 8. Spatial authocorrealtion statistics.

The results, as depicted in Figure 8 reveal a Moran's Index of 0.104632, a z-score of 4.954559, and a p-value of 0.000001. With a z-score as high as 4.954559, which indicate there is a less than 1% likelihood that the observed clustered pattern of CKD incidences is due to random chance. This significant statistical evidence supports the presence of non-random, spatially dependent factors influencing CKD distribution in the study area.

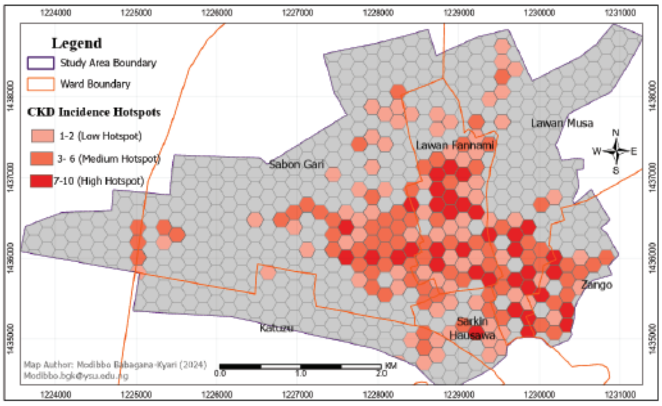


Figure 9. The classified three hotspots by incidence counts density.

The map in Figure 9 illustrates the incidence of Chronic Kidney Disease (CKD) across a surveyed area using hexagonal tessellation. It employs color coding to highlight three distinct clusters of CKD hotspots. Apparently, areas with the highest CKD incidence are marked by darker red hexagons, indicating high hotspots with values ranging from 7-10 cases. Medium hotspots, where CKD incidence is moderate (values of 3-6), are shown in lighter red. Regions with low CKD incidence, termed low spots, are represented by blue hexagons (values of 1-2). The survey area is delineated in grey. This visual representation provides a clear spatial distribution of CKD incidence, allowing for the identification of areas with varying disease prevalence in the study area as depicted in Figure 9.

Water Quality Exploration and CKD Incidence Using GIS

This section of the study focuses on analyzing groundwater quality in CKD hotspots within Northern Yobe State to identify potential environmental risk factors contributing to the prevalence of the disease. Figure 10 provides a visual representation of the geographical distribution of the water sampled locations. The blue dots indicate the locations of the 30 water facilities selected for analysis, which were strategically chosen across the identified disease hotspots as shown in Figure 10.

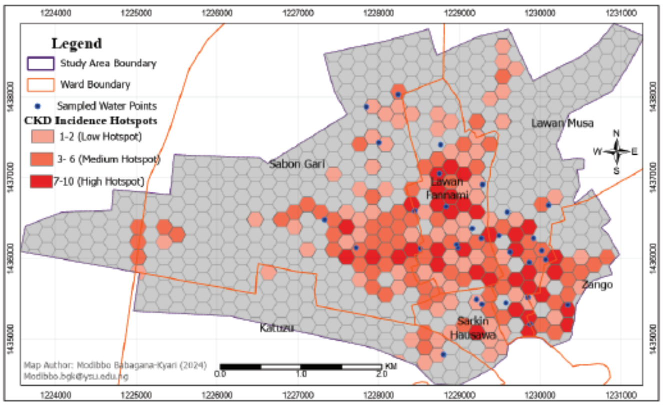


Figure 10. The locations of water points across the study area.

The map above clearly shows that the selected water points align with the hotspot polygons. Overall, the map effectively demonstrates the hotspot-based approach used for selecting the water samples in the area.

Water Quality Parameters in the Disease Hotspots

The analysis of water samples from disease hotspots examined the following parameters, (Arsenic, Cadmium, Lead, Mercury, Chromium, Fluoride, Nitrite, Nitrate, Sodium and phosphate as well as water hardnesses (Magnesium and Calcium) each essential for assessing water quality and its potential health impacts. These parameters and their observation are presented in Table 5.

Table 5. Descriptive statistics of water parameters across the three disease hotspots.

Serial No.	Parameters	High Hotspot (Min, Max, Mean)	Medium Hotspot (Min, Max, Mean)	Low Hotspot (Min, Max, Mean)	Benchmark Standard (WHO)
1	As(mg/L)	0.01-0.129, 0.0456	0.005-0.089, 0.0364	0.003-0.079, 0.0188	0.01 mg/L
2	Cd (mg/L)	0.021-0.317, 0.1081	0.009-0.173, 0.0604	0.000-0.020, 0.005889	0.003 mg/L
3	Pb (mg/L)	0.01-0.238, 0.0624	0.004-0.090, 0.0359	0.001-0.026, 0.008667	0.01 mg/L
4	Hg(mg/L)	0.002-0.02, 0.0058	0.000-0.005, 0.002	0.000-0.001, 0.000444	0.001 mg/L
5	Cr (mg/L)	0.048-0.552, 0.2441	0.014-1.67, 16.7963	0.005-0.080, 0.025111	0.05 mg/L
6	F (mg/L)	6.253-55.48, 25.385	1.78-34.48, 12.9546	0.9733-9.43, 3.278889	1.5 mg/L
7	Na (mg/L)	26.44-226.86, 111.12	7.43-137.42, 50.13	2.02-25.56, 9.88	200 mg/L
8	NO2 (mg/L)	0.2-16.84, 3.39	0.06-11.43, 2.39	0.03-2.95, 0.94	0.3 mg/L
9	NO3 (mg/L)	0.13-197.75, 56.86	4.96-92.92, 38.15	1.96-10.71, 5.39	50 mg/L
10	PO4 (µg/L)	194.72-209.97, 204.29	88.16-213.09, 192.76	186.68-671.51, 254.17	500 µg/L
11	Hardness-Ca (mg/L)	31.96-35.02, 33.89	14.7-35.48, 31.94	31.14-111.03, 42.13	200 mg/L (EPA)
12	Hardness-Mg (mg/L)	20.32-22.42, 21.36	16.26-22.28, 21.10	20.33-44.37, 23.98	150 mg/L (EPA)

The results in Table 5 present the descriptive statistics of water quality parameters tested across three CKD hotspots categorized as high, medium, and low within the study area. Several parameters, including arsenic (As), cadmium (Cd), lead (Pb), and mercury (Hg), were found to exceed the World Health Organization (WHO) benchmark standards, particularly in the high hotspot areas. For instance, fluoride (F) levels in the high hotspot areas showed a significant increased, with a mean concentration of 25.385 mg/L, far exceeding the WHO limit of 1.5 mg/L. Similarly, other contaminants like chromium (Cr), sodium (Na), and phosphates (PO4) also displayed elevated levels, particularly in the high hotspot, indicating significant environmental contamination. The data suggests a correlation between higher contaminant levels in groundwater and the increased incidence of CKD in these hotspots, underscoring the importance of addressing water quality as a potential risk factor for CKD in the region.

Spatial Relationship between Quality Parameters and CKD incidence

The Figure 11. depict the four maps showing the spatial distribution of Chronic Kidney Disease (CKD) incidences in relation to the concentrations of arsenic (As), cadmium (Cd), lead (Pb), and mercury (Hg), with blue dots representing sampled water points at the center of the study area. The analysis reveals that higher concentrations of these heavy metals often coincide with clusters of CKD incidences (denoted by red dots).

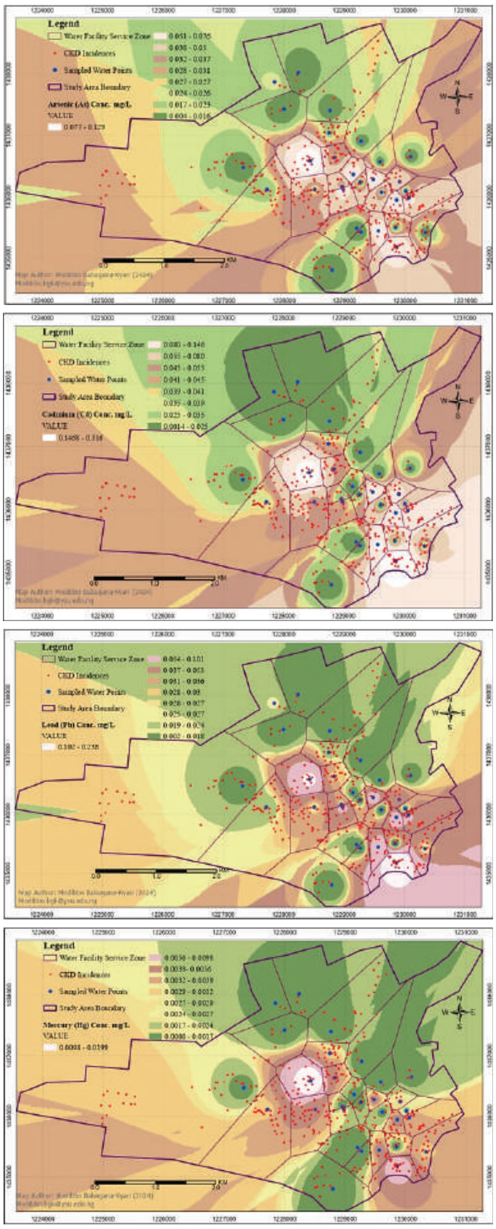


Figure 11. Maps for Arsenic, Cadmium, Lead and Mercury concentration level.

Notably, the maps illustrate that regions with elevated levels of arsenic, cadmium, lead, and mercury are associated with a higher density of CKD cases, suggesting a potential link between contaminated water sources and the prevalence of CKD. The sampled water points (blue dots) in the central area indicate critical locations for assessing water quality in relation to the surrounding CKD incidence, reinforcing the importance of monitoring these contaminants to understand their health impacts.

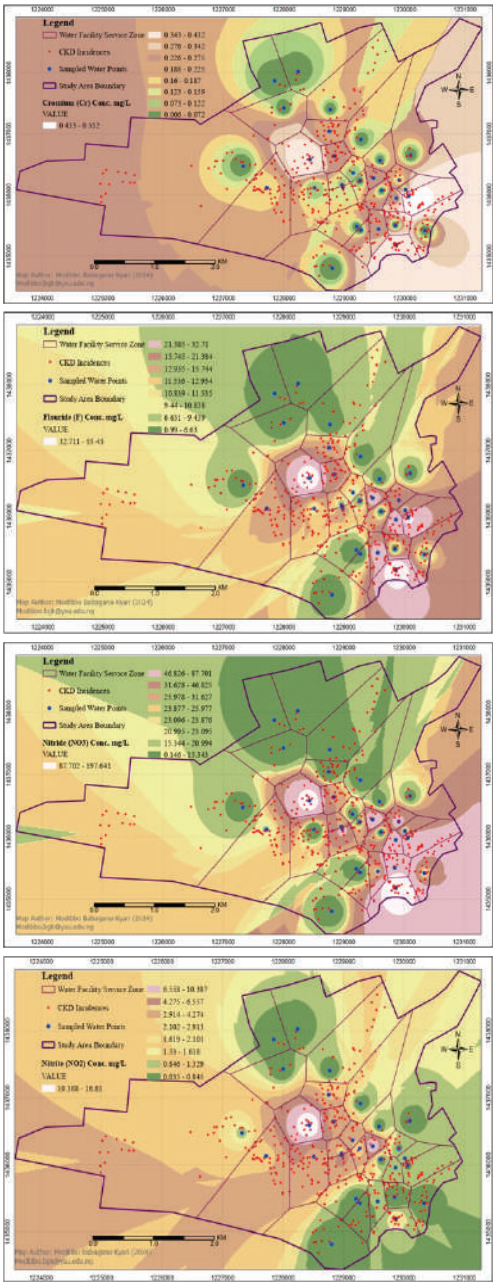


Figure 12. Maps for Chromium, Fluoride, Nitride, and Nitrite.

The four maps in Figure 12 depict the spatial relationship between CKD incidences and the concentrations of chromium (Cr), fluoride (F), nitrate (NO_3), and nitrite (NO_2) in the study area, showing a significant overlap between higher contaminant concentrations and areas with increased CKD cases. The darker shaded areas represent higher levels of these chemicals, which correspond to clusters of CKD incidences, particularly in the central and northwestern parts of the study area. This spatial pattern suggests that these contaminants may be linked to CKD prevalence, highlighting the need for an in-depth exploration.

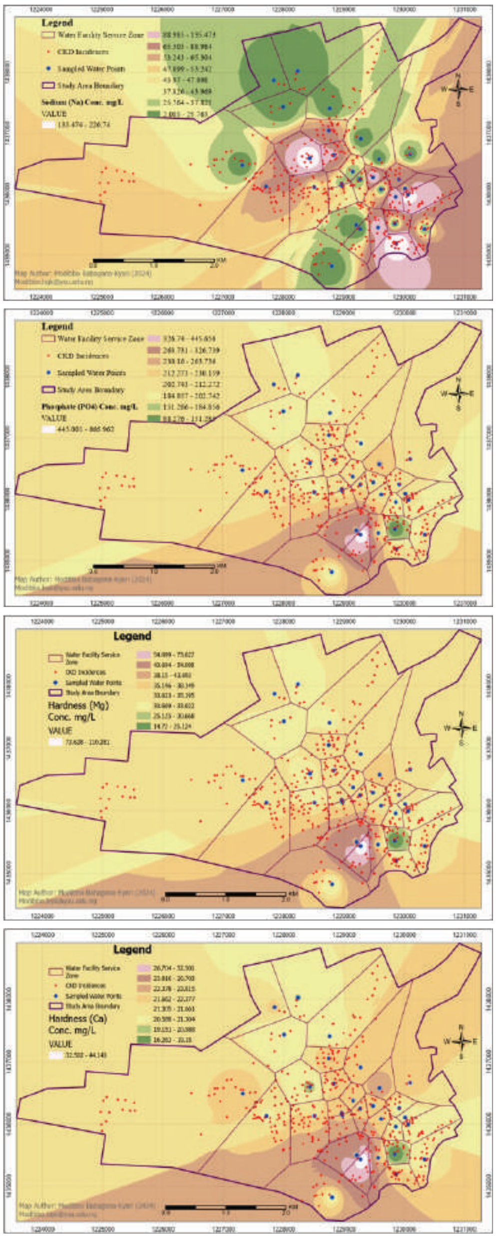


Figure 13. Maps for Sodium, Phosphate, water hardness (Mg), and (Ca).

Similarly, Figure 13 presents the spatial distribution of CKD incidences relative to sodium (Na), phosphate (PO_4), magnesium hardness (Mg), and calcium hardness (Ca) concentrations in the study area. A noticeable overlap is also observed between regions with higher concentrations of these elements and clusters of CKD cases. Specifically, elevated levels of sodium and phosphate in the central and southeastern parts of the study area align with significant CKD incidence pockets.

The maps for magnesium and calcium hardness further support this trend, as areas with higher hardness levels also correspond to CKD hotspots. Overall, in all the parameters particularly heavy metals, there are spatial correlation between the concentration of the disease incidence and water contaminants concentration.

Comparative Analysis of Contaminants Across CKD Hotspots

This section compares the water quality parameters across hotspots. The Box plot in Figure 14 depict the comparison of the observed parameters. Box plots are particularly useful for comparing these parameters against benchmark standards. This section details the use of box plots for comparative analysis, presenting the variations in water quality across the three CKD hotspots areas and their potential health implications.

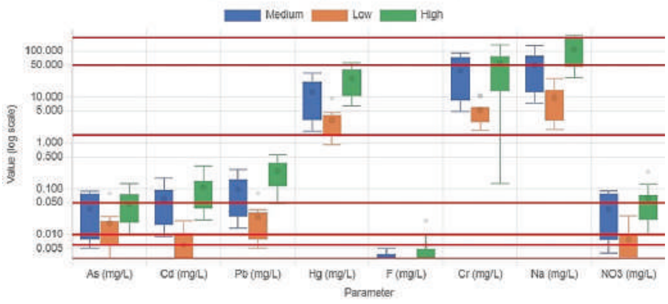


Figure 14. Comparison of water quality elements across CKD hotspots.

Figure 14 presents a comparative analysis of various water quality elements across three CKD hotspot levels Medium, Low, and High using a Box Plot on a logarithmic scale. The analysis shows that contaminant concentrations, such as arsenic (As), cadmium (Cd), lead (Pb), and mercury (Hg), are generally highest in the medium hotspot, followed by the High and Low hotspots, with most values exceeding recommended safety limits. The Medium hotspot also exhibits the greatest variability in contamination levels. Sodium (Na) and fluoride (F) levels peak in the High hotspot, while nitrate (NO_3) concentrations are highest in the medium hotspot. Overall, the Medium and High hotspots display significant contamination, with the Low hotspot having comparatively lower levels.

Water Quality Risk Assessment

This section present health risk assessment of the observed parameters using models such as Hazard Quotient (HQ) and Water Quality Index (WQI) to evaluate the potential health impacts of groundwater consumption in the study area.

Contaminant Hazard Quotients (HQ)

Contaminant Hazard Quotient (HQ) is a measure used to assess the health risks associated with exposure to specific contaminants in groundwater. Through calculating HQ values, the potential risk levels for different contaminants can be determined using their reference dose (RfD) standards (see, Goumenou et al., 2019). Figure 15 present the hazard Quotient (HQ) of each water parameters analyzed in the study. Find the reference dose of the parameters used for model in the link.

The Figure 15 is a line graph depicting the Hazard Quotients for various contaminants across three CKD hotspot (High, Medium and Low incidence areas). Each line represents a different contaminant (As, Hg, Cd, F, Pb, Cr, Na, PO4). The x-axis displays the Hotspot Levels, while the y-axis shows the Hazard Quotient values. The values were calculated using the HQ model. The colored legend identifies the average contaminants represented by each line. Arsenic, Mercury, Cadmium and Fluoride appeared to have high hazard quotient in the high hotspot. While in the low hotspot, virtually all the parameters except arsenic tend to have low HQ values (Figure 15).

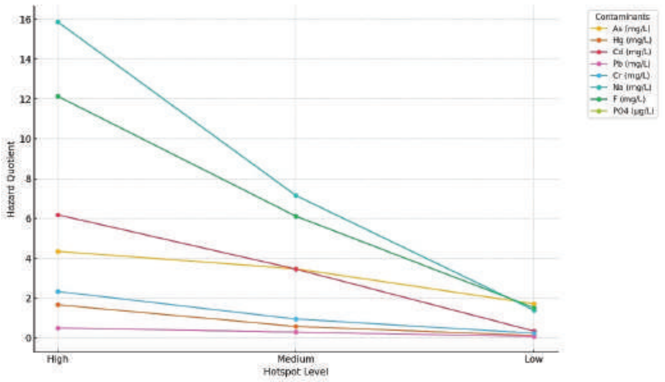


Figure 15. Hazard quotient of contaminants across the identified CKD hotspots.

Water Quality Index

Water Quality Index (WQI) is a comprehensive metric that summarizes the overall quality of water based on various parameters. It provides an easily interpretable indication of water safety for consumption. The below in Figure 16, present the calculated WQI values across the three disease hotspots comparing them to benchmark standards of WHO so as to identify regions with suboptimal water quality.

Figure 16 depict graphically the Water Quality Index (WQI) for selected parameters across three CKD hotspot levels: Medium, Low, and High. The box plot depicts the WQI distribution for each hotspot level, with WQI values on the y-axis and hotspot levels on the x-axis. The colored lines represent different water quality classifications: Excellent ($WQI < 50$), Good ($50 \leq WQI$

< 100), Poor (100 ≤ WQI < 200), Very Poor (200 ≤ WQI < 300), and Unsuitable (WQI ≥ 300) respectively.

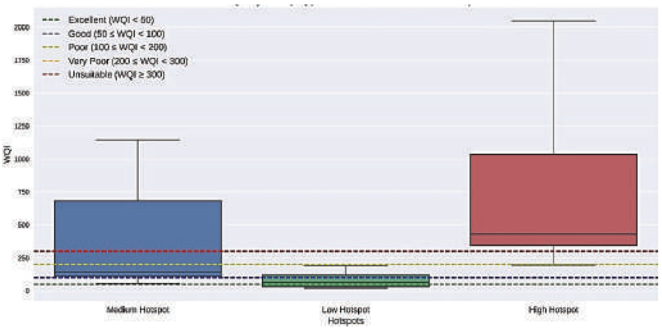


Figure 16. Water Quality Index (WQI) for selected parameters across hotspots.

Additionally, the WQI values for the Medium Hotspot range roughly, with most values falling into the Poor to Very Poor categories. The median WQI value is around 500, illustrating a significant contamination level. Conversely, the WQI values for the Low Hotspot are significantly lower than those of the Medium and High Hotspots. Most values fall within the Excellent to Good categories, with the median WQI value well below 50, indicating relatively clean water quality. The High Hotspot exhibits the highest WQI values, with a median value around 1000 and this clearly suggests that the water quality in this hotspot is predominantly Unsuitable, with significant contamination levels.

CKD and Household Water Usage Patterns

Analyzing household water usage patterns provides valuable context regarding community exposure to groundwater contaminants. The frequency and methods of groundwater usage were examined, shedding more light on behaviours influencing CKD risk factors. Specifically, the data for household water usage includes information on household water source, frequency of usage, duration of water usage and victims' period of residency.

Household primary source of water

The histogram in Figure 17 illustrates the primary water sources used by the CKD victim's household in the study area.

It can be seen from the chart a substantial majority, amounting to 66.89% of the population, rely on public handpumps for their water needs. Private boreholes follow at a significant 18.39%, while pure water/table water supplies 10.03% of the population. The remaining water sources, including harvested rainwater (0.33%), "I don't know" (0.50%), traditional hand-dug wells (0.84%), paddy shallow wells (1.34%), and open river water (1.67%) are used by a relatively small proportion of the population. These distributions signify a heavy dependence on public handpumps, highlighting their critical role in the community's water supply.

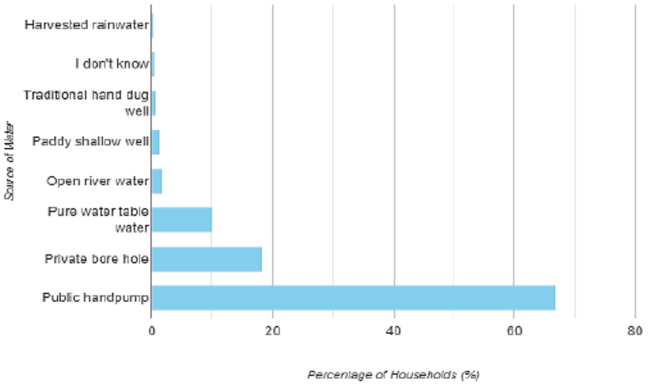


Figure 17. Primary source of water for victim's households.

Relationship between CKD rates and water quality in the study

The section presents the results on the relationship between water sources, CKD cases, and residency periods, portraying significant patterns crucial for understanding the impact of water quality on health over varying durations of residence. Figure 18 presents a heatmap illustrating the distribution of CKD cases across different water sources and residency periods. The x-axis categorizes water sources, while the y-axis represents residency periods in years. The color intensity of each cell corresponds to the number of CKD cases within that specific combination of water source and residency period as can be seen in the chart.

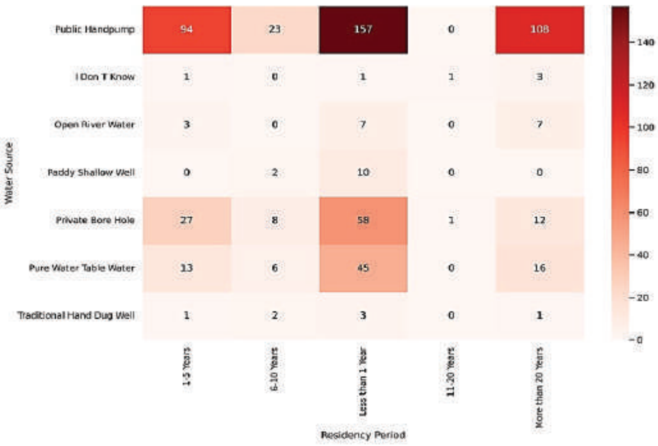


Figure 18. Distribution of CKD cases by water Source and period of residency.

The numbers in the heatmap boxes represent the count of Chronic Kidney Disease (CKD) cases for each combination of water source and victim's residency period in household. Each cell's color intensity varies according to the count, with darker shades indicating higher numbers of CKD cases. This visualization helps identify patterns or trends in CKD occurrences relative to different water sources and the length of residency of victims in households. The heatmap reveals a strong association between public handpumps and CKD cases, particularly among residents with a residency period of 11-20 years. A significant number of CKD cases were also linked to private boreholes across all residency periods. Conversely, water sources such as open river water, paddy shallow wells, and traditional hand-dug wells demonstrated lower frequencies of CKD cases.

Water quality variation across hotspots and water point types

The ANOVA conducted to analyze water quality parameters across three CKD hotspots revealed diverse levels of significance among different contaminants and ions. The F-statistic and corresponding p-values provided insights into the disparities in contaminant levels are presented in Table 6 below.

Table 6. ANOVA results for parameters across the three disease hotspots.

Parameter	F-statistic	P-value	Interpretation
Arsenic (As mg/L)	1.73	0.196	No significant difference
Lead (Pb mg/L)	3.39	0.049	Significant at the 0.05 level
Cadmium (Cd mg/L)	5.22	0.012	Significant difference
Mercury (Hg mg/L)	6.74	0.004	Highly significant difference
Chromium (Cr mg/L)	10.01	0.001	Highly significant difference
Fluoride (F mg/L)	7.2	0.003	Highly significant difference
Sodium (Na mg/L)	8.89	0.001	Highly significant difference
Nitrite (NO2 mg/L)	1.78	0.188	No significant difference
Nitrate (NO3 mg/L)	3.58	0.042	Significant at the 0.05 level
Phosphate (PO4 µg/L)	1.28	0.294	No significant difference
Hardness (Ca mg/L)	1.26	0.299	No significant difference
Hardness (Mg mg/L)	3.62	0.041	Significant at the 0.05 level

Source: Researchers' Analysis (2024)

It worth observing from the above Table 6, that Arsenic (As mg/L) parameter exhibited an F-statistic of 1.7299, associated with a p-value of 0.1963, suggesting that differences in arsenic levels across the hotspots are not statistically significant. Cadmium (Cd mg/L), on the other hand, showed more substantial variation, with an F-statistic of 5.2242 and a p-value of 0.0120, indicating significant differences in cadmium concentrations among the disease hotspots. This pattern of significant variation continued with Lead (Pb mg/L), which maintained an F-statistic of 3.3894 and a p-value of 0.0486, signifying marginal significance.

Similarly, it can be seen mercury (Hg mg/L) and Chromium (Cr mg/L) both demonstrated highly significant differences, with F-statistics of 6.7441 and 10.006, and p-values of 0.0042 and 0.0005, respectively. Notably, Fluoride (F mg/L) and Sodium (Na mg/L) also showed significant disparities, with F-statistics of 7.1983 and 8.8897, and p-values of 0.0031 and 0.0010, respectively, indicating noticeable distinction in their levels across the three disease hotspots.

In contrast, Nitrite (NO2 mg/L) and Phosphate (PO4 µg/L) indicate no significant differences with F-statistics of 1.7811 and 1.2809, and p-values of 0.1876 and 0.2941, respectively. Nitrate (NO3 mg/L), with an F-statistic of 3.5840 and a p-value of 0.0416, indicated significant differences, marking it as a parameter with notable variation.

These statistical results emphasize the spatial variation in water quality across the studied hotspots, with significant differences in several key parameters, implicating environmental or anthropogenic impacts influencing the variation.

Parameters variation across water points

Following the ANOVA, Tukey's HSD post-hoc tests were conducted to identify specific differences between pairs of hotspots particularly high and low hotspots. The Tukey's Honest Significant Difference (HSD) post-hoc test results presented compare water quality parameters between Deep Bore Hole and Hand Pump water points across different hotspots. The table delineates F-statistics and corresponding p-values for each parameter, highlighting statistical differences between the two water point types.

Relatively, for the Hand Pump water point type, the results indicated that the 'High hotspot' consistently exhibited higher levels of contamination compared to the 'Low hotspot' respectively. For instance, Cadmium (Cd) levels were significantly higher between the High and low hotspots with a mean difference of (0.1089) (p-value (0.001). Equally, Chromium (Cr) showed significant differences with a mean difference of (0.1160) (p-value (0.026)), and Sodium (Na) also demonstrated significant differences with a mean difference of (112.1762) (p-value (0.001) as clearly shown in Table 6.

Table 7. Tukey's HSD Post-hoc test for Deep Bore Hole and Hand Pump water points.

Parameter	Deep Bore Hole	Deep Bore Hole	Hand Pump	Hand Pump
	F-statistic	p-value	F-statistic	p-value
As (mg/L)	0.0699	0.9329	2.6246	0.1264
Cd (mg/L)	1.0462	0.3948	5.4256	0.0284
Pb (mg/L)	0.9400	0.4298	2.4539	0.1410
Hg (mg/L)	3.0053	0.1062	3.7681	0.0647
Cr (mg/L)	1.2070	0.3482	19.524	0.0005
F (mg/L)	0.9319	0.4326	11.419	0.0033
Na (mg/L)	1.1761	0.3566	17.023	0.0008
NO2 (mg/L)	0.9433	0.4286	2.0206	0.1884
NO3 (mg/L)	0.7901	0.4862	3.3697	0.0808
PO4 (Âµg/L)	0.0494	0.9520	1.1915	0.3474
Hardness (Ca) (mg/L)	0.0320	0.9685	1.1793	0.3508
Hardness (Mg) (mg/L)	0.3130	0.73980	1.3815	0.2997

The results Table 7, highlight the significant variations in water quality parameters across the hotspots for the Hand Pump water point type. The High hotspot consistently shows higher levels of contaminants such as fluoride, cadmium, sodium, chromium emphasizing the need for targeted interventions and further investigations. Overall, Tukey's HSD provides pairwise comparisons and is particularly useful in identifying which specific groups (hotspots) have significant differences.

DISCUSSION

Survey Characteristics Overview

This section provides a comprehensive overview of the key characteristics and findings of the survey conducted on chronic kidney disease (CKD). It includes details on consent rates, CKD incidence distribution, medical confirmation, household respondent demographics, educational background, and residency duration. These aspects collectively paint a picture of the survey's scope and the data's reliability, providing insights into CKD's prevalence and the community's engagement with the study.

Firstly, the survey achieved an exceptionally high consent rate, with 431 out of 430 respondents agreeing to participate. Even though some refusals were encountered in the course of the survey. This large consent rate indicates a strong willingness among the community to contribute to the study. Consequently, this indicates successful recruitment strategies. High participation rates are often associated with increased reliability and validity of the survey findings, as engaged participants are more likely to provide accurate and comprehensive data. Similarly, it also indicates that the local community are in serious need for uncovering root of the problem in the area.

Moreover, the distribution of CKD cases among surveyed households reveals that 92.6% of households reported at least one incidence of CKD, 2.46% reported two incidences, and 4.7% reported no CKD cases. This distribution highlights a high prevalence of CKD in the study area, suggesting that CKD is a significant health issue within the community. Therefore, such findings are consistent with research indicating that certain regions experience higher CKD prevalence due to local environmental or genetic factors forming hotspots (Friedman, 2019).

Additionally, out of 433 CKD cases surveyed, 98.1% were medically confirmed, while only 1.8% were not. This indicates the quality of the data reported as many of the disease victim's household presented medical records evidence to the field research assistants during the survey. Overall, the high rate of medical confirmation supports the accuracy and reliability of the reported CKD prevalence. Importantly, medical verification is crucial for ensuring the validity of health survey data, as it provides objective evidence of disease status. However, Self-reported disease incidence data through surveys can be a valid source of information in some instances, especially for hypertension, diabetes, and cancer in some settings (see, for example, Jeong et al., 2024).

Furthermore, the survey respondents included 37.44% household heads and 51% other household members, with CKD patients representing only 0.90% of respondents. The predominance of household heads among respondents suggests that the data collected reflects a comprehensive understanding of household health issues. Nevertheless, the low representation of CKD patients highlights a potential gap, likely due to the morbidity severity and mortality associated with the disease in the community as the disease victims eventually die when the disease manifest.

In addition, the majority of respondents (51.8%) had completed high secondary school, with 16.8% having informal education and 15.7% holding some college or a bachelor's degree. This educational profile indicates a generally high level of literacy and comprehension among respondents, which is beneficial for the accuracy of survey responses. Consequently, higher education levels are associated with better understanding of health survey questions and more reliable data (Van et al., 2013).

Finally, a significant portion of households (27.13%) had been residing in the area for more than 20 years, while only 6.05% had lived in the area for 1-5 years and 0.45% for less than a year. The long-term residency of many participants highlights the importance of considering environmental factors when assessing CKD risk particularly the water risk factors. Thus, prolonged exposure to local conditions may contribute to CKD prevalence, highlighting the need to advance research into environmental and lifestyle factors contributing to CKD (Floris et al., 2021).

Overall, this survey overview provides valuable insights into the community's engagement with the study, the prevalence of CKD, and the reliability of the data collected. Consequently, understanding these aspects is key for interpreting the survey findings and informing public health strategies and future research on chronic kidney disease.

Disease Spatial Prevalence Patterns

The significant clustering of Chronic Kidney Disease (CKD) cases in specific geographic hotspots, particularly around the sampled water facilities directly aligns with patterns observed globally in regions facing similar environmental health issues. Specifically, the spatial analysis in this study, utilizing Moran's I and hexagonal tessellation, clearly reveals that CKD cases are not randomly distributed but are concentrated in areas with poor water quality. Consequently, the result strongly supports the hypothesis that environmental factors, especially ground water risk factors, play a role in the etiology of CKD in the study region.

Interestingly, the observed geographic clustering of CKD cases in this study is consistent with findings from research conducted in Central America, South Asia, and other regions where CKD of unknown etiology (CKDu) is prevalent. Similarly, these studies have identified spatial clusters of CKD cases in agricultural and rural settings where water sources are often compromised by environmental contaminants and poor healthcare access (Orantes-Navarro et al., 2017; Jayasumana et al., 2015). Therefore, the spatial concentration of CKD in these hotspots effectively highlights the importance of localized environmental risk factors assessment, particularly those related to water quality.

Moreover, the spatial analysis using Moran's I and hexagonal tessellation in this study provides strong evidence of a potential environmental trigger for CKD, closely linked to water contamination. This pattern is further supported by a growing body of literature associating CKD with exposure to nephrotoxic substances in water, such as heavy metals and agrochemicals, particularly in regions reliant on groundwater affected by agricultural runoff (Wanigasuriya, 2012; Wesseling et al., 2013). Equally, similar spatial patterns have been documented in Sri Lanka and El Salvador, where higher CKD prevalence correlates with areas exhibiting elevated levels of these contaminants (Jayasumana et al., 2015; Crowe et al., 2019).

Consequently, the application of spatial analysis techniques, such as Moran's I and hexagonal tessellation, offers a robust framework for identifying and understanding the geographic patterns of CKD prevalence. Notably, Moran’s I have been widely used to detect non-random spatial patterns, providing critical insights into the underlying environmental or socio-economic determinants of disease (Elliott & Wartenberg, 2004). Furthermore, hexagonal tessellation, with its ability to enhance the precision of spatial representation, corroborates the link between poor water quality and CKD incidence, thereby allowing for more targeted public health interventions (Mennis, 2006).

Ultimately, the spatial clustering of CKD cases in specific geographic hotspots, particularly those with poor water quality, underscores the critical role of environmental factors in the spatial epidemiology of CKD. These findings compellingly emphasize the need to address environmental determinants of health, particularly in regions with limited healthcare access, to effectively combat the growing burden of CKD in vulnerable communities. Therefore, future research should continue to explore these spatial patterns, carefully integrating environmental and socio-economic variables to develop targeted interventions aimed at mitigating CKD in affected regions.

Water Quality Implications

The analysis of water quality in the study area revealed concerning levels of contamination, particularly in the high hotspot regions where several parameters exceeded the World Health Organization (WHO) benchmark standards. The most striking finding was the significant elevation of fluoride (F) levels, with a mean concentration of 25.385 mg/L, which far exceeds the WHO limit of 1.5 mg/L. This is particularly alarming given that high fluoride levels are known to cause various health issues, including dental and skeletal fluorosis, and may exacerbate CKD conditions (Gupta et al., 2020). Similarly, the elevated levels of other contaminants such as arsenic (As), cadmium (Cd), lead (Pb), and mercury (Hg) further underscore the potential health risks faced by communities in these areas. The presence of these contaminants, especially in high concentrations, suggests significant environmental contamination, likely emanated through agrochemical residue runoff since the local engages in intensive paddy rice cultivation.

Spatial Analysis of Water Quality and CKD Incidence

The spatial analysis clearly illustrated a strong correlation between areas of high contaminant concentration and the density of CKD cases. Regions with elevated levels of arsenic, cadmium, lead, and mercury were particularly associated with higher incidences of CKD, supporting the hypothesis that water contamination is a significant environmental risk factor for CKD in the study area. This finding is consistent with global studies that have identified similar patterns, such as in regions of Central America and South Asia, where CKD prevalence is high in areas with contaminated water sources (Jayasumana et al., 2015). The spatial relationship depicted in Figure 12 between CKD incidences and the concentrations of chromium (Cr), fluoride (F), nitrite (NO_2), and nitrate (NO_3) emphasizes the need for quick interventions to remedy water contamination.

Comparison of Water Quality Across Hotspots

The comparative analysis of water quality across different CKD hotspots using Box Plots on a logarithmic scale revealed significant variability in contaminant concentrations. Interestingly, the medium hotspot exhibited the highest variability in contamination levels, with concentrations of key contaminants such as arsenic, cadmium, lead, and mercury generally highest in this hotspot, followed by the High and Low hotspots. This pattern suggests that the medium hotspot may be experiencing fluctuating environmental conditions or intermittent sources of pollution, which could be contributing to the variability (Figure 13). The high levels of sodium (Na) and fluoride (F) in the High hotspot further implicate these areas as critical zones for an intervention due to adverse nephrotoxic effects of these contaminants. This aligns with the findings of Waziri et al. (2017; Gashua et al. 2018; Yuguda et al., 2022), who previously identified heavy metal toxicity (Cadmium, lead and Arsenic, mercury) as a suspected environmental risk factors for the disease. More recently, Goni et al. (2024) also viewed poor groundwater quality as a contributing factor in the aetiology of the disease.

The hazard quotient (HQ) analysis revealed that arsenic, mercury, cadmium, and fluoride have high HQ values in the high disease incidence zone, indicating a significant health risk. Conversely, in the Low hotspot, most parameters tends to have low HQ values, suggesting relatively safer water conditions in these areas. This variation in hazard quotient aligns with the Water Quality Index (WQI) result, where the High hotspot displayed WQI values indicative of unsuitable water quality, with a median value around 1000. This stark contrast between the High and Low hotspots emphasizes the urgent need for environmental remediation in areas with poor water quality.

Mover, the reliance of 66.89% of the population on public handpumps highlights the critical role these water sources play in the community. The strong association between CKD cases and public handpumps, particularly among residents with 11-20 years of residency, suggests a chronic exposure to contaminated water sources. The heatmap analysis further supports this, showing that CKD cases are notably linked to public handpump across high residency periods, whereas open river water, paddy shallow wells, and traditional hand-dug wells exhibited lower frequencies of CKD cases. These findings suggest that interventions should prioritize the improvement of water quality in public handpumps and private boreholes, which are the primary sources of water for these communities. This is consistent with findings from Bihar, India, where Bhatia et al. (2014) reported high arsenic contamination in hand pump drinking water, with 57% of samples exceeding 200 ppb, posing significant health risks to residents. Similarly, Kumar et al. (2016) identified arsenic contamination in Buxar district, Bihar, noting correlations between contamination levels, well depth, and proximity to the Ganga River.

Furthermore, the ANOVA and Tukey's HSD tests confirmed significant spatial variation in water quality across the studied hotspots. Specifically, cadmium, lead, mercury, chromium, fluoride, and sodium levels showed statistically significant differences across the hotspots, with the High and Medium hotspots consistently exhibiting higher contamination levels compared to the Low hotspot. These findings are crucial as they indicate that certain water facilities, particularly hand pumps in high hotspots, are more prone to contamination. This may be due to factors such as proximity to pollution sources, geological conditions, or inadequate water treatment practices as rightly reported by (Gupta & Gupta, 2020). The significant differences identified through these statistical analyses emphasize the need for serious water quality monitoring and remediation efforts in the most affected areas such as reverse osmosis (RO).

Overall, the findings from this study highlight the critical role of water quality in the prevalence of CKD in the study area. The significant contamination levels observed, particularly in the High and Medium hotspots, suggest that addressing environmental and water management issues is essential to reducing the incidence of CKD.

Methodological Limitation and Future Study

This study encountered several limitations that should be considered when using the study findings. The cross-sectional design, while effective in identifying correlations, does not allow for the establishment of causal relationships between water quality and CKD prevalence. Furthermore, reliance on self-reported data for certain sociodemographic variables may introduce potential biases or inaccuracies. Although the study's spatial focus provides detailed insights into the specific area, it may limit the generalizability of the findings to other regions or populations.

Future research should investigate additional factors that may contribute to CKD risk, such as dietary patterns, occupational exposures, and genetic predispositions. Longitudinal studies would be valuable in providing more robust evidence of causal relationships between environmental contaminants and CKD. Expanding the research to encompass a broader geographic area could help determine whether the observed patterns are consistent across different contexts. Moreover, studies on the effectiveness of public health interventions aimed at improving water quality and healthcare access could offer critical insights into reducing CKD prevalence in affected regions.

CONCLUSION AND RECOMMENDATIONS

This study revealed a significant correlation between poor water quality and the prevalence of CKD in the identified hotspots. High concentrations of hazardous contaminants, including arsenic, cadmium, lead, and fluoride, were particularly prevalent in areas with higher incidences of CKD. The spatial analysis revealed a significant concentration of chronic kidney disease (CKD) cases in regions with the highest levels of water contamination, underscoring the substantial environmental influence on the disease's prevalence. This study reinforces the widely believe hypothesis that water quality risk factors may be a significant etiological factor contributing to CKD in Northern Yobe. The spatial correlations observed, alongside the consistency with previous studies, further validate this link, suggesting that environmental determinants like water contamination could play a critical role in the prevalence of CKD in the region.

The findings of this study emphasize the critical role of environmental factors, especially water quality, in the development and spread of CKD within the study area. The clustering of CKD cases in hotspots with poor water quality suggests that these environmental risks must be urgently addressed to halt the disease occurrence. The study findings highlight the need for a multifaceted approach to CKD prevention, focusing on environmental health, early detection thorough biomedical screening.

To address these issues, it is recommended that healthcare infrastructure and accessibility be significantly improved in the community so as to facilitate early diagnosis and effective treatment. Additionally, stringent water quality monitoring and regulation should be enforced in areas with identified poor water quality, including regular testing, remediation, and provision of alternative safe water sources as well as imposition of standards for drilling water point in the area. Finally, public awareness campaigns should be implemented to educate communities about CKD risk factors, particularly those associated with water quality, encouraging preventive behaviors and timely medical consultation to reduce the disease's impact.

REFERENCES

Baffa, A., Abdulkarim, I., & Yaro, N. (2021). Assessment of the spatial distribution of chronic kidney disease in Hadejia emirate, Jigawa state. *Bima Journal of Science and Technology, 5(1), 213-224.*

Babagana-Kyari, M. (2023). Exploratory GIS-based mapping and analysis of chronic kidney disease incidences in Gashua Town, Yobe State, Nigeria. *International Journal of Advances in Nephrology Research, 6(1), 31-45.*

Bhatia, S., Balamurugan, G., & Baranwal, A. (2014). High arsenic contamination in drinking water hand-pumps in Khap Tola, West Champaran, Bihar, India. *Frontiers in Environmental Science, 2, 49.*

Elliott, P., & Wartenberg, D. (2004). Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives, 112(9), 998-1006.*

Floris, M., Lepori, N., Angioi, A., Cabiddu, G., Piras, D., Loi, V., ... & Pani, A. (2021). Chronic kidney disease of undetermined etiology around the world. *Kidney and Blood Pressure Research, 46(2), 142-151.*

Fornace, K. M., Surendra, H., Abidin, T. R., Reyes, R., Macalinao, M. L., Stresman, G., ... & Cook, J. (2018). Use of mobile technology-based participatory mapping approaches to geolocate health facility attendees for disease surveillance in low resource settings. *International Journal of Health Geographics, 17, 1-10.*

Friedman, D. J. (2019). Genes and environment in chronic kidney disease hotspots. *Current Opinion in Nephrology and Hypertension, 28(1), 87-96.*

Gifford, F. J., Gifford, R. M., Eddleston, M., & Dhaun, N. (2017). Endemic nephropathy around the world. *Kidney International Reports, 2(2), 282-292.*

Goni, B. W., Kwairanga, H. S., Abdu, A., Ummate, I., Abdu, A., Ba'aba, A. I., ... & Caplin, B. (2024). High burden of chronic kidney disease of unknown cause among patients receiving renal replacement therapy in Northeast Nigeria: A cross-sectional survey of haemodialysis units. *medRxiv, 2024-06.*

Gupta, A., Kumar, B., Kumar, P., & Gupta, R. (2020). Prevalence of chronic kidney disease and its association with risk factors in Eastern Uttar Pradesh, India. *J. Clin. Exp. Nephrol., 5(4), 90.*

Goumenou, M., & Tsatsakis, A. (2019). Proposing new approaches for the risk characterisation of single chemicals and chemical mixtures: The source related Hazard Quotient (HQS) and Hazard Index (HIS) and the adversity specific Hazard Index (HIA). *Toxicology reports, 6, 632-636.*

Gupta, S. K., & Gupta, I. C. (2020). *Drinking water quality assessment and management*. Scientific Publishers.

Huang, Y., Wu, J., Li, X., & Chen, X. (2021). The application of GIS in environmental health research: Recent advances and future prospects. *Environmental Science & Technology, 55(3), 1234-1250.*

Jayasinghe, S., de Silva, K., & Senanayake, S. (2019). Community participatory approaches in the investigation of CKDu: An innovative model for disease mapping. *BMC Public Health, 19(1), 234.*

Jayasumana, C., Gunatilake, S., & Senanayake, P. (2015). Glyphosate, hard water and nephrotoxic metals: Are they the culprits behind the epidemic of chronic kidney disease of unknown etiology in Sri Lanka? *International Journal of Environmental Research and Public Health, 11(2), 2125-2147.*

Jayasumana, C., Paranagama, P., Agampodi, S., Wijewardane, C., Gunatilake, S., & Siribaddana, S. (2015). Drinking well water and occupational exposure to herbicides is associated with chronic kidney disease, in Padavi-Sripura, Sri Lanka. *Environmental Health, 14, 1-10.*

Jayatilake, N., Mendis, S., Maheepala, P., & Mehta, F. R. (2013). Chronic kidney disease of uncertain aetiology: Prevalence and causative factors in a developing country. *BMC Nephrology, 14(1), 180.*

Jeong, A., Tran, T. X. M., Moon, S. G., Kim, M. K., Kim, I., Kim, Y. M., & Park, B. (2024). Validation of self-reported morbidities of the Korean Atomic Bomb Survivor Cohort. *Epidemiology and Health*, e2024058.

Kumar, A., Rahman, M. S., Iqubal, M. A., Ali, M., Niraj, P. K., Anand, G., ... & Ghosh, A. K. (2016). Groundwater arsenic contamination: A local survey in India. *International Journal of Preventive Medicine, 7(1), 100*.

Luby, S. P., Saha, S., & Andrews, J. R. (2015). Towards sustainable public health surveillance for enteric fever. *Vaccine, 33, C3-C7*.

Lunyera, J., Mohottige, D., Von Isenburg, M., Jeuland, M., Patel, U. D., & Stanifer, J. W. (2016). CKD of uncertain etiology: A systematic review. *Clinical Journal of the American Society of Nephrology, 11(3), 379-385*.

Oliver, M. N., Matthews, K. A., Siadaty, M., Hauck, F. R., & Pickle, L. W. (2005). Geographic bias related to geocoding in epidemiologic studies. *International Journal of Health Geographics, 4, 1-9*.

Orantes-Navarro, C. M., Herrera-Valdés, R., Almaguer-López, M., López-Marín, L., Vela-Parada, X. F., Hernandez-Cuchillas, M., & Barba, L. M. (2017). Toward a comprehensive hypothesis of chronic interstitial nephritis in agricultural communities. *Advances in Chronic Kidney Disease, 24(2), 101-106*.

Priatna, D., & Monk, K.A. (2023) Progress, Challenges, and the Nexus of Research and Impact - the importance of technology in biodiversity conservation in Indonesia. *Indonesian Journal of applied Environmental Studies, 4(1): 3-9*. DOI: 10.33751/injast.v4i1.8778

Priyadarshani, W. V. D., de Namor, A. F. D., & Silva, S. R. P. (2023). Rising of a global silent killer: critical analysis of chronic kidney disease of uncertain aetiology (CKDu) worldwide and mitigation steps. *Environmental Geochemistry and Health, 45(6), 2647-2662*.

Rodriguez, R. A., Hotchkiss, J. R., & O'Hare, A. M. (2013). Geographic information systems and chronic kidney disease: Racial disparities, rural residence and forecasting. *Journal of Nephrology, 26(1), 3*.

Sanati, A. (2015). Chronic kidney disease mortality in Costa Rica; geographical distribution, spatial analysis and non-traditional risk factors. *LUMA-GIS Thesis*.

Saputro, B., Priatna, D., Rosadi, R., & Miyazawa, N. (2023). Spatial analysis of paddy field conversion in Purwakarta Regency, West Java, Indonesia. *Indonesian Journal of applied Environmental Studies, 4(1): 22-29*. DOI: 10.33751/injast.v4i1.6644

Sengupta, M., & Dalwani, R. (Eds.). (2008). *Proceedings of Taal 2007: The 12th World Lake Conference (Vol. 1)*. Ministry of Environment & Forests.

Siddique, K., Rafiq, M., Awan, A., Rehman, A., & Shaheen, M. (2013). Assessment of fluoride levels in drinking water and its health impacts in districts Peshawar and Charsadda, Pakistan. *Journal of Public Health, 21(1), 49-56*.

Smith, A. H., Lingas, E. O., & Rahman, M. (2010). Contamination of drinking-water by arsenic in Bangladesh: A public health emergency. *Bulletin of the World Health Organization, 78(9), 1093-1103*.

Van Der Heide, I., Wang, J., Droomers, M., Spreeuwenberg, P., Rademakers, J., & Uiters, E. (2013). The relationship between health, education, and health literacy: Results from the Dutch Adult Literacy and Life Skills Survey. *Journal of Health Communication, 18(sup1), 172-184*.

Vlahos, P., Schensul, S. L., Anand, S., Shipley, E., Diyabalanage, S., Hu, C., ... & Nanayakkara, N. (2021). Water sources and kidney function: Investigating chronic kidney disease of unknown etiology in a prospective study. *NPJ Clean Water, 4(1), 50*.

Wanigasuriya, K. P. (2012). Update on uncertain etiology of chronic kidney disease in Sri Lanka's north-central dry zone. *MEDICC Review, 14(2), 61-65*.

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 Nephrology, 16(1), 145*.

Wesseling, C., Crowe, J., Hogstedt, C., Jakobsson, K., Lucas, R., & Wegman, D. H. (2013). Mesoamerican nephropathy: Report from the first international research workshop on MeN. *American Journal of Kidney Diseases, 63(3), 502-508*.

Xie, S., Greenblatt, R., Levy, M. Z., & Himes, B. E. (2017). Enhancing electronic health record data with geospatial information. *AMIA Summits on Translational Science Proceedings, 2017, 123*.

Appendix 1. Reference Doses for Contaminants

The following table summarizes the reference doses (RfD) for each contaminant, along with their sources and hyperlinks.

Contaminant	Reference Dose (RfD)	Source	Hyperlink
Arsenic (As)	0.0003 mg/kg/day	U.S. EPA IRIS	https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0278_summary.pdf
Mercury (Hg)	0.0001 mg/kg/day	U.S. EPA IRIS	https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0370_summary.pdf
Cadmium (Cd)	0.0005 mg/kg/day	U.S. EPA IRIS	https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0141_summary.pdf
Fluoride (F)	0.06 mg/kg/day	U.S. EPA IRIS	https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0053_summary.pdf
Lead (Pb)	0.0004 mg/kg/day	U.S. EPA IRIS	https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0277_summary.pdf
Chromium (Cr)	0.003 mg/kg/day	U.S. EPA IRIS	https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0144_summary.pdf
Sodium (Na)	0.2 mg/kg/day	WHO Guidelines for Drinking-water Quality	https://www.who.int/water_sanitation_health/dwq/guidelines/en/

Appendix 2. Reference Doses (RfD) for Contaminants

Contaminant	Reference Dose (RfD)	Source	Hyperlink
Arsenic (As)	0.0003	EPA	https://www.epa.gov/iris
Mercury (Hg)	0.0001	EPA	https://www.epa.gov/iris
Cadmium (Cd)	0.0005	EPA	https://www.epa.gov/iris
Fluoride (F)	0.06	EPA	https://www.epa.gov/iris
Lead (Pb)	0.0004	EPA	https://www.epa.gov/iris
Chromium (Cr)	0.003	EPA	https://www.epa.gov/iris
Nitrite (NO2)	0.1	EPA	https://www.epa.gov/iris
Nitrate (NO3)	1.6	EPA	https://www.epa.gov/iris
Hardness (Ca)	0.2	EPA	https://www.epa.gov/iris
Hardness (Mg)	0.2	EPA	https://www.epa.gov/iris
Phosphate (PO4)	0.03	EPA	https://www.epa.gov/iris