Tap water access and its relationship with cholera and other diarrhoeal diseases in an urban, cholera-endemic setting in the Democratic Republic of the Congo

AURELIE JEANDRON

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy of the University of London

JUNE 2020

Department of Infectious Diseases Epidemiology

Faculty of Epidemiology and Population Health

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by a grant from the French Development Agency and the Veolia Foundation to the London School of Hygiene and Tropical Medicine

Research group affiliation: Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases

Statement of own work

I, Aurelie Jeandron, confirmed that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

[redacted]

Aurelie Jeandron

Date: 4th of June 2020

Acknowledgements

I would first like to dedicate this thesis and this work to Dr. Jeroen Ensink, my supervisor, manager, mentor and friend, who was killed on the 29th of December 2015 outside his home. I would not have embarked on this research degree in January 2014 without his patient support and his optimism at project inception, despite numerous context and funding uncertainties. His implicit trust and contagious optimism were much needed, as were his engineering perspective and pragmatic advice on all aspects of the work. This project was carried on and this thesis written with his peaceful presence on my shoulder.

I would also like to express my deepest gratitude to Prof. Simon Cousens, for having taken over the supervision of this work in tragic circumstances and for having helped me focus and strive for excellence. Thank you for your patience, understanding and unlimited encouragements. I am grateful to Oliver Cumming and Prof. Sandy Cairncross as well for their support to the project, with a broader perspective and guidance.

Big thanks to all my colleagues in the Environmental Health Group, and especially those sharing the 411 office. Everyone's passion, tenacity and creativity have been an unwavering source of inspiration in the past 7 years, on good, bad and ugly days.

This project would never have happened without the funding from the French Development Agency and the Veolia Foundation – but it was about more than just funding research. Special thanks to Florent Bedecarrats, Martin Lemenager, Claire Zanuso and Marine Bepoix at AFD, Thierry Vandevelde, Thibaut Constant and Damien Machuel, without forgetting the early contributions of Frank Haaser and Clement Petit, from the Veolia Foundation. I am indebted to Charlotte Lauff, volunteer with the Veolia Foundation, who dazzlingly assisted with water samples analyses in Uvira over 3 weeks of extremely intense work in strenuous conditions in 2016.

Work in the Democratic Republic of the Congo was only made possible thanks to the OXFAM GB teams in Kinshasa, Goma, Bukavu and Uvira. Too many people helped one way or another to cite them all, but special thanks to Joanna Trevor, Georges Mugaruka, Lily Seguin, Geoffroy Sewa and Mazingi. I would also like to thank all those who assisted in data collection: Dieumerci, Felicien, Clarisse, Benjamin, John, Michel, Olivier, Balthazar, Josee, Oscar, Patrick and Sophie. Special mention for Jaime Mufitini Saidi and Baron Bashige Rumedeka, who made this work possible by collecting all CTC-related data with great professionalism. Laboratory work would not have materialised without the commitment, determination and generosity of the CRH team: Vercus Lumami Kapepula, Clement and Charlotte.

Many thanks as well to Dr. Peter Jensen, Prof. Anowara, Dr Marie-Laure Quilicci, Dr. Esmeralda Valiente, Dr Antonio Gasparrini, Prof. Ben Armstrong, Dr Lynn Grignard, Dr Sebastian Funk for their expert advice on specific aspects of the work.

Fieldwork in Uvira would have been unbearable at times without the presence of a very special ICRC team. Amelie, July, Karim and Renaud, you have contributed more to this work than you can imagine by providing a safe room when needed, rides around town on Sundays even under torrential rain, security intelligence, lots of fun around the babyfoot, incredible food, emotional comfort and much more… You had my back and I believe you saved my life several times. Thanks a million.

I can't thank enough my parents, broader family and in-laws, who provided an invaluable support with their infinite faith, patience and grounding when it was most needed.

Finally, I am immensely grateful to my husband Richard, who supported me in every way possible, close and remote, with patience, humour, attentions, and just… unconditional love. He shared the frustrations, the excitement, the satisfaction, the apprehension and everything else like the best companion I could ever dream of on a long and challenging journey. Team.

"Doubt is unpleasant,

but certainty is absurd."

Voltaire

Abstract

Globally 1.3 billion people are at risk of cholera in endemic countries, where nearly 3 million cases occur annually, of which 3 % are fatal. The burden is highest in South-Asia and Sub-Saharan Africa, where diarrhoeal diseases in general are also a leading cause of mortality and morbidity, especially for children under 5. Uvira, the second largest city of South-Kivu province in the eastern part of the Democratic Republic of the Congo, has been affected by cholera since it reached the region in the late 1970s. Suspected cholera cases have been reported nearly every week since 2004 by the Uvira cholera treatment centre (CTC).

This thesis first shows that about 40% of the patients admitted to the CTC test positive for cholera with rapid diagnostic tests, and that infections with other common enteric pathogens are highly prevalent. Two surveys of water-related practices in more than 500 households indicate that tap water is sometimes used by nearly 80% of households for drinking purposes but only systematically by less than 50%, whilst surface water is the main source of domestic water for nearly 40% of households. Tap water access is a predictor of the quality and quantity of domestic water used in households. Time-series regression reveals a 2.5-fold increase in CTC admissions within the 12 days following a 24h interruption in tap water supply. Finally, a multivariable time-series model highlights the influence of tap water supply variability in time and space on suspected cholera incidence, especially that attributed to epidemic transmission.

By demonstrating the influence of coverage disparities and intermittency of the current tap water supply network on households' water-related practices and suspected cholera incidence in Uvira, this research establishes a solid base for a much-needed impact evaluation of on-going improvements on suspected cholera in an endemic area.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABU	Analysis base unit
AFD	Agence Française de Développement
AIC	Akaike's information criterion
aOR	Adjusted odds ratio
APW	Alkaline peptone water
AS	Aire de santé
BCZ	Bureau central de zone
BIC	Bayesian information criterion
CATI	Case area targeted intervention
CDF	Congolese Franc
CFR	Case fatality rate
CFU	Colony forming units
CHIRPS	Climate hazards group infrared precipitation with station data
CI	Confidence interval
CRH	Centre de recherche en hydrobiologie
CTC	Cholera treatment centre / Centre de traitement du choléra
DALY	Disability-adjusted life years
DNA	Deoxyribonucleic acid
DPD	N,N-diethyl-p-phenylenediamine
DPS	Division provinciale de la santé
DRC	Democratic Republic of the Congo
EBK	Empirical bayesian kriging
EHEC	Enterohaemorrhagic *Escherichia coli*
EHG	Environmental Health Group
ETEC	Enterotoxigenic *Escherichia coli*
EU	European Union
FC	Faecal coliforms
FIB	Faecal indicator bacteria
GIS	Geographical information system
GPS	Global positioning system
GTFCC	Global task force on cholera control
HDI	Human development index
HGR	Hôpital général de référence
IDP	Internally displaced people
IMERG	Integrated multi-satellite retrievals for global precipitation measurement
INGO	International non-governmental organisation
IPC	Infection prevention and control
IQR	Interquartile range
IRR	Incidence risk ratio
IV	Intravenous
lcd	litres per capita per day
LSHTM	London School of Hygiene and Tropical Medicine
MONUSCO	United Nations Organization Stabilization Mission in the Democratic Republic of the Congo
MPN	Most probable number

NGO	Non-governmental organisation
NTU	Nephelemetric turbidity unit
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
OCV	Oral cholera vaccination
ODK	Open Data Kit
OR	Odds ratio
ORS	Oral rehydration salt
OSM	Open Street Map
PCA	Principal component analysis
PCR	Polymerase chain reaction
PoUWT	Point of use water treatment
PPV	Positive predictive value
RDT	Rapid diagnostic test
RN	Route nationale
SD	Standard deviation
SDG	Sustainable Development Goals
SIR	Susceptible – infected - recovered
SIRS	Susceptible – infected – recovered - susceptible
SNEL	Société nationale d'électricité
STEC	Shiga-toxin producing *Escherichia coli*
swcRCT	Step-wedge clustered randomised controlled trial
TNTC	Too numerous to count
TTC	Thermotolerant coliforms
UN	United Nations
UNHCR	United Nations High Commissioner for Refugees
UNIKIN ESP	Université de Kinshasa – Ecole de Santé Publique
UV	Ultraviolet
VBNC	Viable but non-culturable
WASH	Water, sanitation and hygiene
WC	Whole cell
WGS	Whole genome sequencing
WHO	World Health Organization
ZS	Zone de santé

1

Introduction

Despite a substantial decrease in attributable mortality and a less important but nonetheless significant decrease in attributable morbidity in the past 25 years, diarrhoeal diseases, including cholera, remain a leading cause of death and ill-health globally, in particular in children under 5 years of age. It is estimated that 1.3 billion people are at risk of cholera, mostly in South Asia and Sub-Saharan Africa . A large part of the estimated annual 3 million cases also occur in those regions, and approximately 3 % of them are fatal **[1]**. Cholera is only one of the many pathogens causing diarrhoea, but it holds a particular place amongst diarrhoeal diseases for several reasons. Cholera is one of the diseases that was most feared during the 19^{th} century, when it hit major cities of Europe and America and contributed to the development of public health and epidemiology. It is also striking in its “explosiveness”, with epidemics having the potential to spread extremely fast to a very large number of people, even in our “modern era”, as was seen in 2010 in Haiti and in 2017 in Yemen. In addition, the most severe form of cholera can kill alarmingly quickly in a very distinctive clinical presentation. Like other diarrhoeal diseases, cholera is spread through faecal-oral transmission pathways, that can be addressed by an array of water, sanitation and hygiene (WASH) interventions. Specific characteristics of *Vibrio cholerae*, the organism responsible for cholera, and of cholera infection and disease, represent however opportunities and challenges for specific control strategies. These also need to be targeted to the poorest and most vulnerable populations, disproportionately affected by cholera in addition to other diarrhoeal diseases, often in a context of natural disaster or conflict.

1.1 Global epidemiology of cholera and diarrhoeal diseases

Cholera is believed to have originated in the Ganges delta region and was already prevalent in that region in ancient times. From there, six cholera “pandemics” spread to other parts of the world between 1817 and 1927, when it remained endemic in the Indian sub-continent but receded everywhere else. The current pandemic – the “Seventh pandemic” – originated in Indonesia in 1961, and spread to the Middle East, Africa, Europe and the Americas **[2]**.

Between 1970 and 2016, nearly 9.2 million cases of suspected cholera and nearly 900,000 deaths were reported to the World Health Organization (WHO) **[3]**. These numbers of cases and deaths are however believed to grossly understate the true numbers, due to the fear of economic and travel sanctions and challenges in adequately identifying, diagnosing and reporting cholera cases. Global estimates calculated using population-based incidence and spatial modelling suggest instead that every year between 2008 and 2012, nearly 2.9 million cases occurred in 69 endemic countries, with 91,000 deaths **[1]**.

Between 2010 and 2016, Lessler et al. estimated that a mean of more than 141,000 cholera cases per year occurred in Sub-Saharan Africa, highlighting however that this number varied substantially from one year to another and that only 4 % of all districts, representing less than 9 % of the total population, reported a high incidence rate above 1 case in 1,000 people annually (**Figure 1-1**)**[4]** .

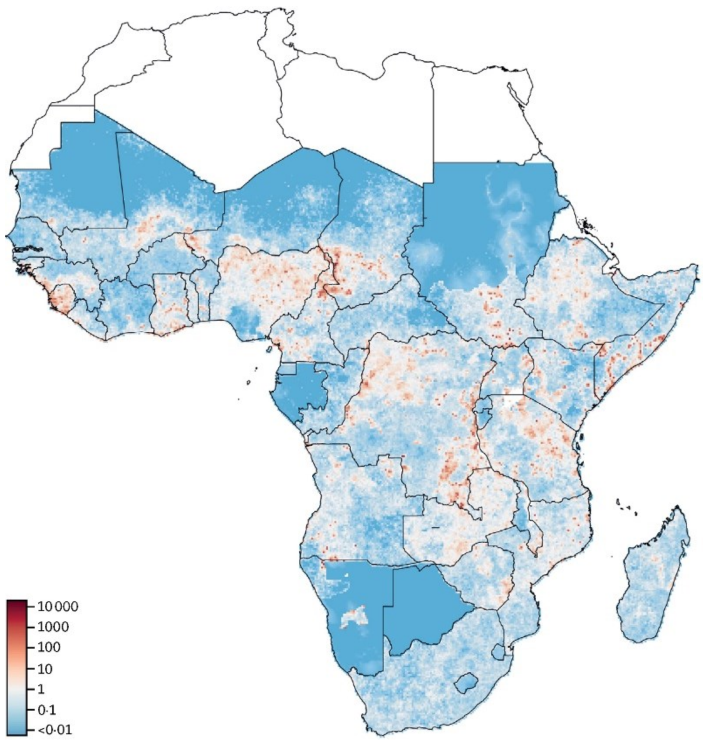


Figure 1-1 Mean annual cholera incidence per 100,000 people in sub-Saharan Africa between 2010 and 2016 [4]

Based on WHO figures, Haiti and the Democratic Republic of the Congo (DRC) reported the highest number of cases globally between 2010 and 2016, with approximately 796,000 and 166,000 cases respectively over 6 years. These figures do not include the largest cholera outbreak ever recorded that has been going on in Yemen since 2016, with well over a million cases reported the first year **[4]**.

In 2017, Sub-Saharan Africa was the region bearing most of the diarrhoeal diseases burden globally, with an estimated 38 million Disability-Adjusted Life Years (DALYs) lost of the 81 million worldwide. 9 % of this burden was attributed specifically to cholera, far behind rotavirus, shigella and cryptosporidium (33 %, 22 % and 21 % respectively) **[5]**. According to the same source, cholera is also the fourth leading cause of fatal diarrhoea in Sub-Saharan Africa (9 % of diarrhoea deaths), after the same three pathogens cited above (28 %, 23 % and 18 % respectively). Of note, morbidity attributed to cholera is low in relative terms in children under 5 years of age, who are particularly affected by other pathogens.

1.2 Diarrhoeal pathogens and cholera

1.2.1 Brief overview

A great number of different pathogens - viruses, bacteria or parasites - are responsible for infectious diarrhoea worldwide. Diarrhoea is the result of the disruption of two normal functions of the intestines: fluid and electrolyte secretion and re-absorption **[6]**. Clinical presentation of infectious diarrhoea can be classified into three major syndromes: acute watery diarrhoea, dysentery – diarrhoea with blood - and persistent diarrhoea – lasting more than 14 days. The same pathogen can lead to more than one of the above syndromes, and clinical presentation alone is unreliable to identify the causative agent, including for cholera **[7,8]**.

Most of the diarrhoeal diseases burden worldwide in all age groups is attributed to the following enteropathogens (by agent type): 1) viruses: adenovirus, norovirus, rotavirus; 2) bacteria: Aeromonas spp., Campylobacter spp., *Clostridium Difficile*, enteropathogenic *Escherichia coli*, enterotoxigenic *E. coli* (ETEC), non-typhoidal Salmonella spp., Shigella spp., *Vibrio cholerae*; 3) protozoa and parasitic agents: amoeba (mostly *Entamoeba histolytica*), Cryptosporidium spp. **[9]**.

Aetiology distribution between age groups and regions can vary widely, and most studies on diarrhoeal aetiologies have focussed on children under 5, for whom diarrhoea mortality rates are the highest.

Infectious doses vary widely by pathogen: 1 to 100 viral particles or cysts/oocysts for parasitic agents, a few hundred colony forming units (CFU) for some bacteria (for example Shigella spp., Campylobacter spp) and 10^5 to 10^8 for others (ETEC and *V. cholerae* for instance) **[10]**. Incubation time can vary from a few hours to 14 days, and duration of illness from 2 days to several months **[11]**. Pathogens are usually shed in large numbers in the faeces of infected individuals, especially during the diarrhoeal illness episode, but shedding can last longer that the symptoms **[12]**.

Replacing fluids and electrolytes lost is the key treatment for infectious diarrhoea, and oral rehydration is often sufficient unless diarrhoea is associated with vomiting or the dehydration is severe. Antimicrobial therapy may be beneficial for specific bacterial pathogens and can reduce symptoms severity and illness duration but is not systematically recommended except for shigellosis in children **[7,13]**.

Most enteropathogens induce a degree of immunity but many uncertainties remain on the level of protection and its duration **[14]**. Naturally acquired immunity against viruses is generally stronger than against bacteria and protozoa, although prior symptomatic infection with Cryptosporidium spp. and Shigella spp. was noticeably protective in young children in the MAL-ED cohort study **[15]**. Better understanding of naturally acquired immunity against enteric pathogens has led to the development of vaccines ; vaccines are licensed for human use for rotavirus, *Salmonella Typhi* and *V. cholerae* (see below) **[16,17]** and progress is being made on vaccines against Shigella spp. and ETEC **[18]**.

1.2.2 Cholera infection and disease

Vibrio cholerae

Vibrio cholerae is a gram-negative bacterium with a flagellum and a pili (**Figure 1-2**). More than 200 serogroups of *V. cholerae* exist, with O_1 and O_{139} being the only ones responsible for epidemic cholera. *V. cholerae* can persist in aquatic environments, especially brackish or estuarine waters, often in association with zooplankton, shellfish, and aquatic plants. *V. cholerae* can enter a viable but non-culturable (VBNC) state within biofilms – a structured aggregate of micro-organisms in a protective coating -, in response to unfavourable environmental conditions such as nutrient deficiencies or salinity and temperature fluctuations.

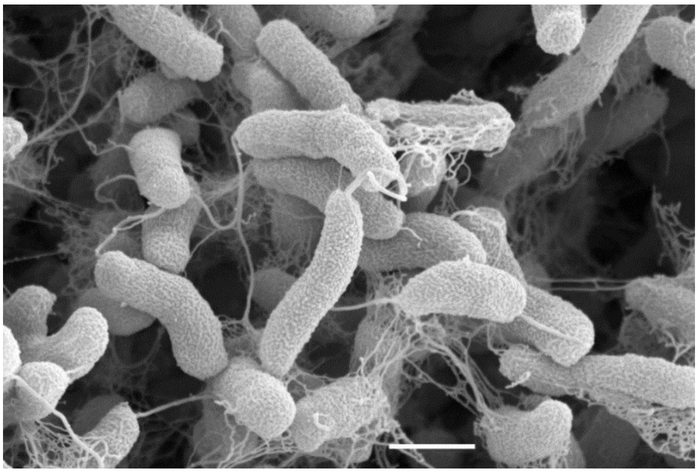


Figure 1-2 *Vibrio cholerae*. Scanning electron microscopy
Bar = 1 μm [19]

Pathogenesis

The capacity of *V. cholerae* strains to cause disease depends on a combination of virulence factors that allow the organism to colonize the small intestine and to produce cholera toxin. The toxin provokes electrolyte and water loss in the small intestine's epithelial cells and reduces water re-absorption in the colon, leading to watery diarrhoea.

To reach the small intestine, *V. cholerae* must resist the normal defense mechanisms of the upper gastro-intestinal tract, one of which is gastric acidity. *V. cholerae* is sensitive to low pH and only a large inoculum size will allow the organism to withstand gastric acidity.

An infectious dose of 10^8 to 10^{11} bacteria was required to obtain consistent colonization in healthy human volunteers, but this infective dose decreased to 10^4 to 10^8 when inoculated with a bicarbonate buffer that reduced gastric acidity **[20-22]**. Food acts as a gastric acid buffer, which means that the cholera infectious dose is lower when ingested with food than when ingested with water. In natural settings, it is believed that doses as low as 10^2 to 10^3 organisms are enough to cause disease, although possibly with a lower attack rate and less severe cases **[21,23,24]**.

It has been shown in animal models that, the infectious dose required for colonization was 10 times lower for a short period of time (several hours) after *V. cholerae* was shed in rice-water stool **[25,26]**. Although the mechanisms are not yet fully understood, the decreasing infectiousness of *V. cholerae* as time since it was shed from a human host increases has been suggested as an explanation for the high transmission rates within households and for the explosive nature of outbreaks **[27]**.

In addition to the infectious dose ingested and pH levels in the stomach, the risk of infection and/or clinical presentation in individuals is affected by several genetic and nutritional risk factors. In particular, retinol deficiency – an indicator of malnutrition – and zinc deficiency are associated with an increased risk of infection and clinical disease **[28]**. People with blood group O are more likely to develop severe disease if infected **[29]**. Concomitant infection with other enteropathogens or intestinal parasites may also interfere with the immune response to *V. cholerae* O_1 **[30,31]**.

The immune response in people with cholera is mainly directed at bacterial surface antigens and at the cholera toxin, and involves distinct types of intestinal and serum immunoglobulins, as well as memory B and T-cells **[32]**. These are crucial for long-term protection against subsequent cholera infections induced by either vaccination or natural exposure. Several studies suggest that an initial symptomatic infection with cholera leads to a protection against the same cholera serogroup infections for at least 3 years **[33,34]**. Asymptomatic or mild infections may not confer as much protection or any protection at all **[35]**. The combined roles of antibacterial and antitoxin immune responses have guided the development of oral cholera vaccines (OCV) described in section 1.3.3 below.

Clinical presentation and treatment

The incubation period between infection and first symptoms ranges between 8 h and 5 days **[36]**. Most people infected with cholera will show no symptoms at all or mild diarrhoea indistinguishable from other aetiologies **[8]**. The proportion of asymptomatic cases among infected people ranges between 33 % and 99 % **[34]**.

The classic clinical presentation with abundant watery diarrhoea ("rice water" diarrhoea) and concurrent vomiting represents the most severe form of the disease, *cholera gravis*. It is estimated that only 2 to 5 % of people infected with cholera will actually suffer from *cholera gravis* . In that case, rapid water and electrolyte loss can lead to fatal dehydration in less than 6 h if not properly replaced **[37-39]**. As in other non-inflammatory diarrhoea, fever is typically absent, except in the presence of concurrent infections, such as malaria.

Cholera is easily treatable, and the cornerstone of treatment consists of appropriate oral or intravenous rehydration. Antibiotic administration may be useful to reduce symptoms duration, although antibiotics sensitivity of the circulating cholera strain should ideally be checked as antimicrobial resistance is common **[40-44]**. With timely and adequate treatment, the case fatality rate (CFR) should be below 1 %, but the death rate in untreated patients with severe cholera can reach 70 % **[45]**.

Rice water stools can contain up to 10^{11} - 10^{12} *V. cholerae* organisms per liter, and severe cases can purge up to 1 l/h in the early stages **[46,47]**. Vibrio shedding in stools commonly lasts less than a week in both symptomatic and asymptomatic cases, although prolonged shedding - up to several months – has been be observed in association with HIV or malnutrition **[46-49]**.

1.3 Diarrhoeal diseases transmission and control strategies

1.3.1 Fecal-oral transmission and WASH interventions

Cholera has been closely linked with contaminated water ever since John Snow suggested removing the pump handle from the Broad Street pump in 1854, during a cholera outbreak in London. Water is considered, along with contaminated food, to be the dominant vehicle of transmission of *V. cholerae*, although like other diarrhoeal pathogens transmitted via the fecal-oral route, cholera can be transmitted along the pathways summarized in the F-diagram (**Figure 1-3**) **[50]**.

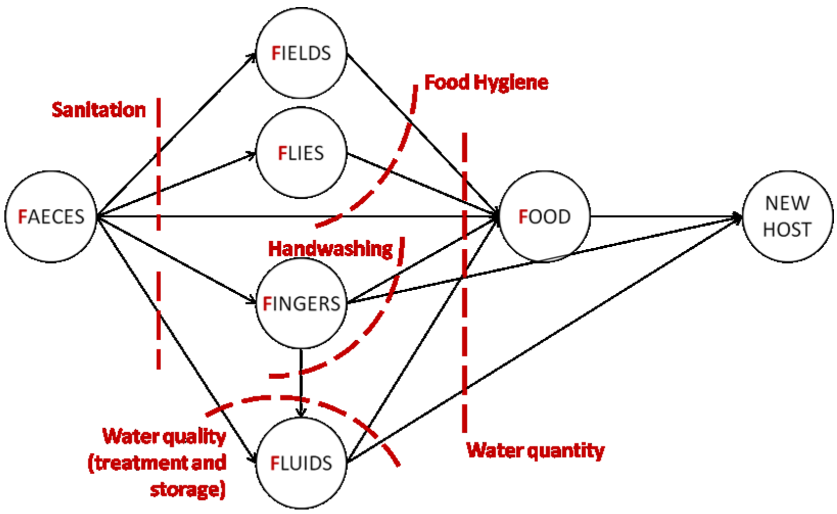


Figure 1-3 F-diagram representing fecal-oral transmission routes and potential WASH prevention mechanisms (adapted from Wagner and Lanoix [50])

WASH interventions that effectively halt faecal-oral transmission are therefore a key prevention approach for diarrhoeal diseases and are briefly summarised below. Several reviews also describe in great detail WASH prevention mechanisms (*see section 1.4.1 WASH interventions and diarrhoeal diseases below*)

Consistent use of appropriate sanitation is the primary barrier to prevent human faeces and associated pathogens from contaminating the environment. Open defaecation, unsafe disposal of children's faeces and inadequate sanitation technologies can all lead to environmental contamination either directly or by intermediate vectors such as insects, pests or domestic animals. Wastewater and "night soil" are also rich sources of nutrients that may be used as fertilizer to improve soil composition and crop yield. In the absence of precautions to remove or reduce the number of pathogens present, the crops may however become contaminated by faecal pathogens, either within or on their surface.

Food preparation practices may not be adequate to remove or inactivate these pathogens on food items eaten raw.

The risk of ingesting harmful pathogens, including cholera, is also decreased by improving the microbiological quality of the water used for drinking and cooking, either by protecting water sources from faecal contamination, by improving water collection and storage practices or by promoting drinking water treatment. Reliable and more convenient access to water may lead to better drinking water microbiological quality, as water is stored for shorter times and in better conditions.

Additionnally, increasing the amount of domestic water used is likely to translate into better hand, food and domestic hygiene practices that all require water. Practicing hand washing with soap at key times – after using a toilet or defecation, before eating or feeding a child, after caring for a child, before preparing food – prevents pathogens from contaminating hands and hands from contaminating food or water. Domestic hygiene, for example cleaning the floor of the dwelling and the kitchen, or washing clothes soiled by faeces or vomit from a sick individual, is a water-consuming activity as well. Water is also needed to wash and clean produce eaten raw and cooking utensils or surfaces, possibly faecally contaminated by hands or flies.

Food hygiene includes adequate cooking and re-heating, to inactivate pathogens present on or in the food, and better food storage practices – in covered containers to avoid contamination or in a refrigerator to reduce bacterial growth.

Although the WASH interventions described above can all contribute to preventing any disease caused by pathogens transmitted along the fecal-oral routes, their respective effectiveness in reducing cholera transmission is influenced by specific survival characteristics of *V. cholerae* in the environment.

1.3.2 Survival and inactivation of enteric pathogens in the environment

All enteric pathogens are sensitive to high temperatures in liquids, and most will be inactivated at temperatures between 60 and 80 °C in less than 5 min **[51,52]**. Similarly, chlorine inactivates most pathogens in water : a 2-log reduction is achieved at a concentration of 1 mg/l in 2 to 30 min for viruses, less than a few seconds for bacteria and 25 min to 4 h for protozoa **[53]**. A notable exception is for Cryptosporidium oocysts, against which chlorine is not effective. The disinfection performance of chlorination is affected by pH and the presence of particles or biofilms, that allow organisms to evade chlorine effects **[54]**. Inactivation by ultraviolet (UV) treatment of water is variably effective for bacteria and viruses, with efficacy generally higher for protozoa **[53,54]**.

Viral and parasitic agents can generally survive longer in the environment than bacterial pathogens, but all three types can be recovered from water, wastewater, soil and crops after 10 days to 3 months **[11]**. Transmission through fomites has been shown to be particularly relevant for viruses **[55]** but many bacterial enteric pathogens can also be found on fomites, especially when humidity and temperature conditions are favourable **[56,57]**. Some bacterial enteric pathogens – Salmonella spp and vibrios for instance - have the ability to adapt to stressful environments – lack of nutrients or moisture for example – by entering into a VBNC state, that increases their survival. They can then revert to a culturable and infective state when conditions are more favourable, especially when passing through a host's intestines **[58-60]**.

Animal faeces can also be a source of environmental contamination for specific human enteric pathogens **[61]**. Domestic husbandry has also been shown to increase the risk of diarrhoea caused by some protozoa (e.g. Cryptosporidium spp and Giardia spp), Campylobacter spp. and enterohemorrhagic *E. coli* (EHEC) / shiga-like toxin-secreting *E. coli* (STEC) **[62]**. A large number of diarrhoea-causing pathogens, in particular bacteria, have also been isolated from the surface and guts of the common housefly *Musca Domestica* **[63]**; houseflies preferences for food and faecal matter make them a meaningful vector for diarrhoeal pathogens.

V. cholerae

V. cholerae is part of the autochtonous bacterial flora of natural waters and non-pathogenic strains are commonly isolated from the environment **[64]**. This persistence in the environment is linked to *V. cholerae*'s ability to assume "survival" forms in association with zooplankton, phytoplankton and aquatic life **[65]**. It is accepted that *V. cholerae* in the environment is mostly found in VBNC state, and in aggregates or biofilms. These forms allow vibrios to resist unfavourable conditions (low temperature, low osmolarity, low pH, lack of nutrients) and evade predation, especially from lytic bacteriophages (viruses infecting and replicating in bacteria). VBNC cholera retains the capacity to colonize the human intestine and be shed as a culturable form. VBNC cholera is therefore considered infectious, although its infectivity is lower than vibrios freshly shed by an infected individual **[66]**. Recent research has however uncovered two separate mechanisms that suggest hyperinfectivity of *V. cholerae* in biofilms or when released by protozoa **[67,68]**.

Environmental modifications in endemic areas, such as seasonal plankton blooms or El-Nino related changes in water temperature, are associated with an increased likelihood of cholera epidemics **[69-71]**. It is suggested that these events disrupt the balance found in environmental waters between *V. cholerae* and lytic bacteriophages, resulting in an increase in numbers of *V. cholerae* in these waters.

It is therefore not surprising that *V. cholerae* were demonstrated to survive up to several months when experimentally spiked into different types of water. Low temperatures (4 °C) do not seem to reduce significantly the survival duration of *V. cholerae* in water, but temperatures above 60 °C for 10 min or more inactivated all vibrios **[54,72,73]**. Exposure of water to sunlight is also detrimental for *V. cholerae* survival in water **[73]**.

Cholera organisms in water are generally sensitive to chlorine. "Smooth" colonies of *V. cholerae*, ie those not expressing a gene coding for the production of a particular exopolysaccharide favouring aggregation into biofilms, are inactivated in less than 20 s at a concentration of 1 mg/l free chlorine **[54]**. However, some strains have the ability to shift to a "rugose" form, particularly under environmental stress, which remains infective and virulent but is more resistant to chlorine. Viable organisms persist after 30 min in 2 mg/l free chlorine **[74,75]**. After 1 h exposure to 3 mg/l of free chlorine, 10 times more rugose colonies of *V. cholerae* strains isolated in Haiti survived in comparison with the smooth colonies of the same strains **[76]**. Large aggregates of *V. cholerae*, like those potentially found in water supply distribution systems, also show an increased resistance to chlorine, attributed to both the rugose type and the clumping **[77,78]**. Reduction in the number of *V. cholerae* cells after exposure to chlorine was also lower when they were attached to shrimps, even at chlorine concentrations up to 10 mg/l for 5 min **[79]**.

Figure 1-4 Colony morphology of smooth (S) and rugose (R) strains after 48h incubation on L-agar [75]

V. cholerae survival in various food items fluctuates widely between a few hours and several days, with evidence available for meat, fish, dairy products, vegetables, fruits and drinks. Refrigeration generally increases survival duration, while acidic food or alcoholic beverages reduce it to a couple of hours . *V. cholerae* not only survives, but also grows on most cooked foods at temperatures ranging between 20 and 35 °C. In cooked chicken, the number of *V. cholerae* cells increased by 2 to 5 logs in 8 h at 22 and 37 °C, and growth was generally faster in more alkaline foods, such as eggs or shellfish **[80]**.

V. cholerae is particularly sensitive to dessication and survives a few hours only on non-absorbent materials (glass, metal). It can however survive several days on absorbent materials such as linen **[81,82]**. Efficacy of 5 disinfectant types (pH-adjusted bleach, 1 % citric acid, 70 % ethanol, quaternary ammonia, or Pine-Sol®), on 5 different types of materials (aluminum, carpet, concrete, glass, and wood) on *V. cholerae* reduction showed a pooled mean recovery of 41 % across the 5 materials and the 5 disinfectants **[83]**.

V. cholerae was recovered from houseflies body surface and guts in laboratory experiments and in cholera-affected areas – although the number of organisms was relatively small on their legs, larger quantities – about 10^6 CFU per fly - were isolated from their guts and regurgitated saliva 3-4 days after they were infected with *V. cholerae* **[84-87]**.

1.3.3 Oral cholera vaccines

As soon as *V. cholerae* was extracted and purified in 1883 by Koch, parenteral vaccines were prepared with cultures and tested for nearly a hundred years **[23]**. The immunity they conferred only lasted a few months and caused unpleasant side effects. With the development of microbiology, other vaccine development strategies have been followed since 1980 and in 2018 three oral cholera vaccines are WHO pre-qualified for vaccination campaigns and vaccination of travellers **[32]**. All three are killed whole cell (WC) vaccines, and one of them includes cholera toxin B sub-unit (WC-rBS) as well. The primary regimen is based on two doses administered two to six weeks apart, and they are considered safe for use during pregnancy. Dukoral® (Valneva) can be used from 2 years of age, whereas Sanchol™ and Euvichol® (Shantha Biotech and Eubiologics respectively) can be used from 1 year of age.

A recent meta-analysis of the protection conferred by these whole-cell vaccines concluded that they provide a moderate to high level of protection for at least 3 years, possibly longer **[88]**. It also concluded that the efficacy was significantly lower in children younger than 5 years than in older age groups. In Bangladesh between 1985 and 1990, an efficacy trial concluded that two doses of Dukoral® provided 85 % protection for the first 6 months after vaccination, and approximately 60 % after 2 years, but the efficacy dropped much faster in children younger than 5 years of age **[89]**. In India since 2006, another trial found an efficacy of the Sanchol™ vaccine of 67 % at 2 years, across all age groups, with no evidence of a reduction in efficacy in the second year **[90]**. Recent re-analysis of the same trial data estimated effectiveness at 38 % for those vaccinated from ages 1 to under 5 years, 85 % for those 5 to under 15 years, and 69% for those vaccinated at ages 15 years and older **[91]**. Effectiveness estimates overall are more heterogeneous with varied approaches to vaccination targeting and coverage, as reported by a recent systematic review of cost effectiveness analyses of OCV **[92]**. While OCV was generally cost-effective in the absence of other interventions, even when no indirect effect of vaccination was assumed (herd effect), no economic evaluation found

OCV cost-effective in comparison with WASH interventions. Cost-effectiveness of OCV is highest for populations at high-risk of cholera and when access to health facilities is limited **[92]**.

Until recently, the use of these vaccines had been modest, due to their limited availability and concerns about their cost **[93]**, acceptability **[94]**, and the logistics of administering two doses in a short period of time **[95]**. In 2011 however, the WHO revived its Global Task Force on Cholera Control (GTFCC) and established in 2013 a global stockpile of Sanchol™ vaccine for use in both emergency and non-emergency settings **[96]**. Between 2013 and 2018, more than 36 million doses from this stockpile were shipped for more than 104 mass vaccination campaigns in 22 countries **[97]**.

1.4 WASH interventions against diarrhoeal diseases and cholera

1.4.1 WASH interventions and diarrhoeal diseases

Evidence of WASH interventions' impact on diarrhoeal diseases has been summarised by numerous systematic reviews and meta-analyses since the 1980s, with pooled estimates of impact in each intervention domain. WASH interventions are all separately expected to have a positive impact on diarrhoeal diseases, with risk reductions ranging from 3 % to 61 % (**Figure 1-5**). Evidence is lacking however on how these interventions interact with each other. For example, increasing water availability may also promote more hygienic practices in household, or improve the quality of water at point-of-use although this may only be observed in particular contexts or interventions.

Pooled estimates do not reflect the variability of interventions grouped under the same category. For example, Wolf et al. highlighted in their reviews that each step of improvement along the water source ladder, from unimproved source to continuous tap water supply on premises through community improved source and basic piped water, had a positive impact on diarrhoeal diseases **[98,99]**.

Measures of diarrhoeal disease incidence used in many of the impact studies included in the systematic reviews summarised above relied largely on diarrhoea episodes reported weekly or bi-weekly by community members or primary carers for children. They are therefore particularly vulnerable to courtesy bias from participants unless blinding on intervention or control status is implemented **[109-111]**. This is naturally impossible for many WASH interventions, such as providing a piped water connection on the premises or building an improved latrine.

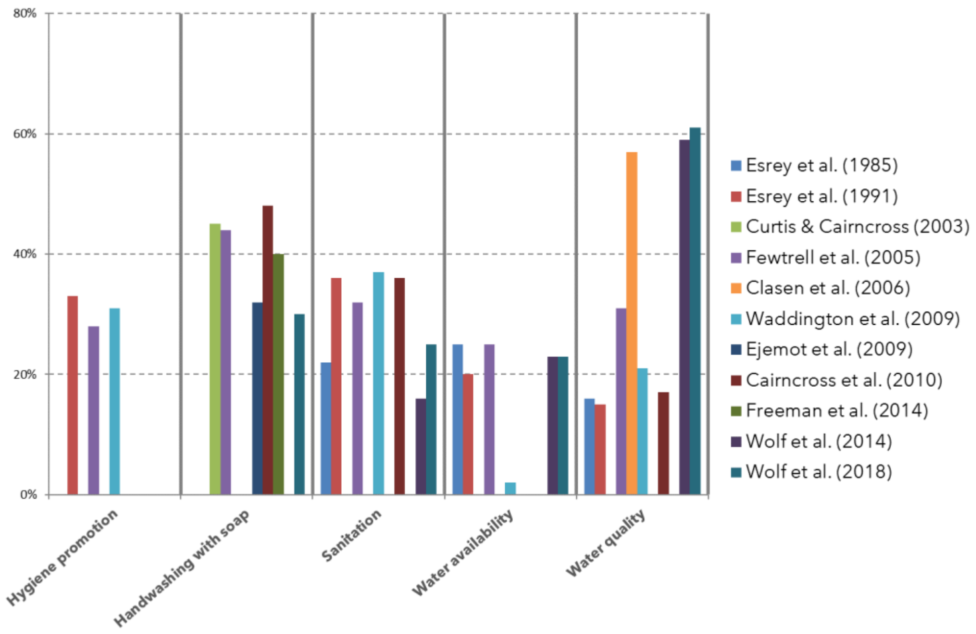


Figure 1-5 Summary of pooled estimates of diarrhoeal diseases risk reduction for different WASH interventions [93 - 103]

Many recent studies on the impact of WASH interventions have added to a reported diarrhoeal episodes outcome, a measure of enteric pathogen shedding in participants' faeces, or an assessment of environmental faecal contamination **[112]**.

Results of three large trials on the impact of WASH interventions on child health, that included, amongst other primary outcomes, diarrhoeal episodes, enteric pathogen infection and faecal contamination in the household environment, highlighted that the WASH interventions delivered had a generally positive impact on children's exposure to enteropathogens, although this impact was small in comparison with the considerable levels of faecal contamination encountered in households of all three study sites. This may explain why diarrhoea incidence did not decrease in two study sites out of the three **[113]**.

1.4.2 WASH interventions and cholera

Although it is highly plausible that WASH interventions that are effective against diarrhoea in general also decrease the risk of cholera infection, evidence of impact for the same interventions against cholera specifically is scarce. In 2015, Taylor and colleagues conducted a systematic review of the literature pertaining to the impact of WASH interventions during cholera epidemics **[114]**. 18 studies were identified and only five of them included a direct measure of the impact on cholera disease, while the others assessed the impact by proxy indicators of function or use. Of the five studies with a health impact measure, four assessed a water quality intervention (filtration or chlorination at point-of-use, safe storage) and one evaluated a community water supply and sanitation intervention.

Three of the four water quality studies reported a significant positive impact on cholera incidence consistent with the impact of similar interventions on other diarrhoeal diseases. Colwell and colleagues reported a halving of cholera incidence in Bangladeshi villages after introduction of surface water filtration with local materials (saris) that removed zooplankton cholera hosts (copepods) from collected water **[115]**. Compliance to this simple filtration method was very high, with less than 1 % of the targeted population not performing it. Five years later however, Huq and colleagues found that only 18 % of the population initially targeted was still using that filtration method and another 13 % using another filtration method, and this translated into a 25 % reduction of cholera incidence lacking statistical significance **[116]**. In Kenya, solar water disinfection during a cholera outbreak showed a significant and major decrease (88 %) in cholera incidence for children less than 5 years of age, but no impact in other age groups **[117]**. Finally, in cholera endemic Indian slums, improved water storage containers (narrow-mouthed) reduced cholera incidence by 75 % and chlorination by households reduced cholera incidence by 58 %, both interventions also reducing the incidence of asymptomatic infections in contacts of cases **[118]**.

These three studies however suffered methodological flaws common to WASH or diarrhoeal diseases studies. In particular, allocation to intervention/groups was not randomized (2/3 studies), outcome measurement methods were prone to reporting or information bias (1/3 studies) and participants were not blinded to the intervention (3/3 studies).

The community water supply and sanitation improvement study by Azurin and Alvero reported a marked decrease in cholera incidence in the three communities which received a water supply improvement, a sanitation improvement or both compared to a control community, but no statistical test was reported, the allocation of the interventions was not randomized, and only four communities were studied in total **[119]**. These results should therefore be interpreted with caution.

Another review by Lantagne and Yates was conducted in 2018 on the impact of household water treatment interventions during cholera outbreaks, with a broader set of inclusion criteria – including grey literature. They concluded that moderate quality evidence suggests a beneficial impact of household water treatment on cholera, highlighting however the lack of evidence on sustained water treatment after the intervention **[120]**.

Results of a randomized controlled trial of a hospital-based hygiene and water treatment intervention (CHoBI7) in Dhaka, Bangladesh, have also been published since Taylor's review **[121]**. The CHoBI7 trial in Bangladesh reported a 47 % decrease in cholera infections in household contacts of cholera confirmed cases in the intervention arm, who received in the hospital and at home a 7-day program to promote handwashing with soap and drinking water treatment. This trial showed that the intervention had been particularly effective at improving stored drinking water quality **[122,123]**.

1.4.3 Delivering WASH interventions

Reviews of the impact of WASH interventions on diarrhoeal diseases highlight the role of intervention delivery in achieving health benefits, and variations in the effectiveness of WASH interventions is often attributed in part to differences in the way improvements were brought to or promoted in various communities.

Most WASH interventions require some degree of behaviour change from beneficiaries: washing hands with soap more often and at the right times, using and maintaining a new sanitation facility, treating drinking water at home, storing drinking water in a different way, using a new source of water **[124]**. Outcomes of WASH interventions are therefore highly dependent on the way they are delivered to the targeted population and on the uptake or compliance achieved. Intensive behaviour change interventions may be needed to achieve the level of compliance necessary to reduce exposure to faecal pathogens, compromising the cost-effectiveness of such interventions when scaling-up. Maintaining the changes over time is another major challenge **[125,126]**.

Improving water supply and sanitation usually involve some level of provision of facilities/infrastructure, for instance community safe water sources, latrine building materials and services, pit emptying equipment, waste treatment amenities. Sustaining these services often requires financial contributions from users, introducing additional difficulties in terms of willingness and capacity to pay and equity. Services also need to operate within a larger economic, social and political context, often with a complex set of public and private stakeholders **[127,128]**.

Cholera outbreaks often occur in acute emergency situations, after a natural disaster, population displacement and/or conflict, within disrupted and highly dynamic contexts. Short-term WASH interventions are then usually channelled through specific emergency funding and accountability systems, compromising sustainability of services to achieve immediate health gains **[129,130]**. High incidence of cholera and diarrhoeal diseases is however not limited to acute emergencies and many cholera hotspots in sub-Saharan Africa are located in areas affected by long-lasting humanitarian crises or in fragile states **[131,132]**. Achieving sustained improvements in water and sanitation services to prevent cholera and diarrhoeal diseases then requires in-depth understanding of the context and addressing a large range of development issues, from lack of coordination and leadership in governmental institutions to poor financial management or low human resource capacity **[133-136]**.

1.5 Origin, aims and objectives of the research

1.5.1 Origin of the research

The present research started in 2012 when the Veolia Foundation approached the LSHTM and the Environmental Health Group (EHG) for guidance on cholera epidemiology and WASH assessment while designing a proposal to improve tap water supply in Uvira in the South-Kivu province of DRC. The Veolia Foundation had been involved in cholera control activities for several years in eastern DRC.

As the funding for these improvements was being secured from the French Development Agency (AFD) and the European Union (EU) (a total of 10 million Euros), AFD's Research and Evaluation department and the Veolia Foundation commissioned the LSHTM and the EHG to design and conduct a prospective impact evaluation of the planned improvements on cholera and other diarrhoeal diseases.

Development of an evaluation protocol started in 2014 and was modified multiple times as the planned intervention was developed, refined and finalised. This protocol was approved in October 2015 and the evaluation itself was meant to be the subject of this PhD thesis, to address the lack of evidence on if and how tap water supply improvements may prevent and control cholera outbreaks in endemic settings (*see section 1.4.2 above*).

Improvement works were initially supposed to start in mid-2016 but delays in the financing negotiations with the Regideso (DRC National water agency) and the DRC government accumulated, and an international contractor was only awarded the contract for the improvement work in mid-2017, for an actual start of building activities in Uvira at the end of 2017 and with an expected duration of 24 months.

Data collection and analyses already initiated to inform the impact evaluation were however a great opportunity to strengthen our understanding of the burden of cholera and diarrhoeal diseases in Uvira and how it relates to current tap water service.

1.5.2 Aims and objectives

Focusing on the cholera-endemic town of Uvira in eastern DRC, the aim of the present research is to describe the burden of suspected and confirmed cholera in Uvira and to examine the role of tap water service on suspected cholera incidence in this urban setting.

The specific objectives of the research were the following:

- **Objective 1**: Describe the epidemiology of suspected and confirmed cholera based on admissions and rapid diagnostic testing confirmation data collected at the Uvira Cholera Treatment Centre (CTC);
- **Objective 2**: Investigate the relationship between tap water service and water-related practices in households;
- **Objective 3**: Examine the impact of tap water supply interruptions on CTC admissions;
- **Objective 4**: Investigate the influence of tap water service on temporal and spatial patterns of acute diarrhoeal diseases leading to CTC admissions.

1.6 Structure of the thesis

The present thesis is divided in seven chapters and research findings are presented in article format. They are organized as follows.

Chapter 1 provides an overview of diarrhoeal diseases epidemiology, control strategies and gaps in knowledge, with a special focus on cholera. It introduces the research rationale and its aims and objectives.

Chapter 2 introduces the study area as well as the data sources and methods used throughout the thesis.

Chapter 3 describes the epidemiology of suspected and confirmed cholera in Uvira (Objective 1), first with an article published in August 2018 in **PLoS One** based on data collected between April 2016 and November 2017 **[137]**. It then includes a further analysis of the data extended to April 2019, and a summary of results from a side study on other enteropathogens.

Chapter 4 examines the relationship between tap water infrastructure and water-related practices in households (Objective 2). A first part of this chapter presents in detail the findings of two household surveys conducted in 2016 and 2017, and a second part consists of an article published in December 2019 in the journal **npj Clean Water** on the prediction of drinking water contamination and domestic water consumption in households, based on a tap water service indicator and access to alternative water sources **[138]**.

Chapter 5 reports the results of a time-series analysis examining the relationship between daily variations in volume of tap water supplied by the town water treatment plant and the daily incidence of suspected cholera cases admitted to Uvira CTC (Objective 3). These results were published in **PLoS Medicine** in October 2015 **[139]**.

Chapter 6 explores the combined time and space patterns of CTC admissions in Uvira and their relationship with tap water availability using a discrete endemic-epidemic modelling framework (Objective 4).

Chapter 7 synthesizes the findings of the research and discusses its strengths and limitations. It also includes recommendations for further research and reflections on the PhD learning process.

Appendices complement the body of the thesis, with supporting information related to each chapter.

1.7 Chapter 1 - References

1. Ali M, Nelson AR, Lopez AL, Sack DA. *Updated Global Burden of Cholera in Endemic Countries*. **PLOS Neglected Tropical Diseases**. 2015;9(6):e0003832.
2. Chowdhury FR, Nur Z, Hassan N, von Seidlein L, Dunachie S. *Pandemics, pathogenicity and changing molecular epidemiology of cholera in the era of global warming*. **Annals of Clinical Microbiology and Antimicrobials**. 2017;16:10.
3. Global Health Observatory data repository. http://apps.who.int/gho/data/node.main.175. Accessed 21/03/2017.
4. Blackburn CC, Lenze Pe, Jr., Casey RP. *Conflict and Cholera: Yemen's Man-Made Public Health Crisis and the Global Implications of Weaponizing Health*. **Health Secur**. 2020;18(2):125-131.
5. Collaborators GBDDD. *Estimates of global, regional, and national morbidity, mortality, and aetiologies of diarrhoeal diseases: a systematic analysis for the Global Burden of Disease Study 2015*. **Lancet Infect Dis**. 2017;17(9):909-948.
6. Hodges K, Gill R. *Infectious diarrhea: Cellular and molecular mechanisms*. **Gut Microbes**. 2010;1(1):4-21.
7. Casburn-Jones AC, Farthing MJG. *Management of infectious diarrhoea*. **Gut**. 2004;53(2):296-305.
8. Learoyd TP, Gaut RM. *Cholera: under diagnosis and differentiation from other diarrhoeal diseases*. **Journal of travel medicine**. 2018;25(suppl_1):S46-s51.
9. Troeger C, Blacker BF, Khalil IA, et al. *Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a*

systematic analysis for the Global Burden of Disease Study 2016. **The Lancet Infectious Diseases**. 2018;18(11):1211-1228.

10. Kothary MH, Babu US. *Infective dose of foodborne pathogens in volunteers: a review*. **Journal of Food Safety**. 2001;21(1):49-68.
11. Gerba CP. *Environmentally Transmitted Pathogens*. **Environmental Microbiology**. 2015:509-550.
12. Julian TR. *Environmental transmission of diarrheal pathogens in low and middle income countries*. **Environmental Science: Processes & Impacts**. 2016;18(8):944-955.
13. World Health O. Handbook : IMCI integrated management of childhood illness. Geneva: World Health Organization; 2005.
14. Lopman BA, Baker JM. *Wading into the morass: natural immunity to enteropathogens*. **J Infect Dis**. 2020.
15. Rogawski McQuade ET, Liu J, Kang G, et al. *Protection From Natural Immunity Against Enteric Infections and Etiology-Specific Diarrhea in a Longitudinal Birth Cohort*. **The Journal of infectious diseases**. 2020.
16. Cohen D, Muhsen K. *Vaccines for enteric diseases*. **Hum Vaccin Immunother**. 2019;15(6):1205-1214.
17. Czerkinsky C, Holmgren J. *Vaccines against enteric infections for the developing world*. **Philosophical Transactions of the Royal Society B: Biological Sciences**. 2015;370(1671):20150142.
18. Barry E, Cassels F, Riddle M, Walker R, Wierzba T. *Vaccines Against Shigella and Enterotoxigenic Escherichia coli: A summary of the 2018 VASE Conference*. **Vaccine**. 2019;37(34):4768-4774.
19. Muhsin Ozel GH. Vibrio cholerae. Scanning electron microscopy. In: Robert Koch Institute, ed.
20. Cash RA, Music SI, Libonati JP, Schwartz AR, Hornick RB. *Live oral cholera vaccine: evaluation of the clinical effectiveness of two strains in humans*. **Infect Immun**. 1974;10(4):762-764.
21. Kaper JB, Morris JG, Jr., Levine MM. *Cholera*. **Clinical microbiology reviews**. 1995;8(1):48-86.
22. Watson AP, Armstrong AQ, White GH, Thran BH. *Health-based ingestion exposure guidelines for Vibrio cholerae: Technical basis for water reuse applications*. **Sci Total Environ**. 2018;613-614:379-387.
23. Barua D, Greenough WB. *Cholera.* New York: Plenum Medical Book Co; 1992.
24. Spira WM, Khan MU, Saeed YA, Sattar MA. *Microbiological surveillance of intra-neighbourhood E1 Tor cholera transmission in rural Bangladesh*. **Bull World Health Organ**. 1980;58(5):731-740.
25. Alam A, Larocque RC, Harris JB, et al. *Hyperinfectivity of human-passaged Vibrio cholerae can be modeled by growth in the infant mouse*. **Infect Immun**. 2005;73(10):6674-6679.
26. Butler SM, Nelson EJ, Chowdhury N, Faruque SM, Calderwood SB, Camilli A. *Cholera stool bacteria repress chemotaxis to increase infectivity*. **Mol Microbiol**. 2006;60(2):417-426.
27. Hartley DM, Morris JG, Jr., Smith DL. *Hyperinfectivity: a critical element in the ability of V. cholerae to cause epidemics?* **PLoS Med**. 2006;3(1):e7.
28. Harris JB, LaRocque RC, Chowdhury F, et al. *Susceptibility to Vibrio cholerae infection in a cohort of household contacts of patients with cholera in Bangladesh*. **PLoS Negl Trop Dis**. 2008;2(4):e221.
29. Harris JB, LaRocque RC. *Cholera and ABO Blood Group: Understanding an Ancient Association*. **Am J Trop Med Hyg**. 2016;95(2):263-264.

30. Chowdhury F, Begum YA, Alam MM, et al. *Concomitant enterotoxigenic Escherichia coli infection induces increased immune responses to Vibrio cholerae O1 antigens in patients with cholera in Bangladesh.* **Infect Immun**. 2010;78(5):2117-2124.

31. Harris JB, Podolsky MJ, Bhuiyan TR, et al. *Immunologic responses to Vibrio cholerae in patients co-infected with intestinal parasites in Bangladesh.* **PLoS Negl Trop Dis**. 2009;3(3):e403.

32. Clemens JD, Nair GB, Ahmed T, Qadri F, Holmgren J. *Cholera.* **Lancet**. 2017;390(10101):1539-1549.

33. Ali M, Emch M, Park JK, Yunus M, Clemens J. *Natural cholera infection-derived immunity in an endemic setting.* **J Infect Dis**. 2011;204(6):912-918.

34. King AA, Ionides EL, Pascual M, Bouma MJ. *Inapparent infections and cholera dynamics.* **Nature**. 2008;454(7206):877-880.

35. Woodward WE. *Cholera reinfection in man.* **J Infect Dis**. 1971;123(1):61-66.

36. Azman AS, Rudolph KE, Cummings DA, Lessler J. *The incubation period of cholera: a systematic review.* **The Journal of infection**. 2013;66(5):432-438.

37. Butler T. *Treatment of severe cholera: a review of strategies to reduce stool output and volumes of rehydration fluid.* **Transactions of the Royal Society of Tropical Medicine and Hygiene**. 2017;111(5):204-210.

38. Pietroni MAC. *Case management of cholera.* **Vaccine**. 2020;38 Suppl 1:A105-a109.

39. Williams PCM, Berkley JA. *Guidelines for the management of paediatric cholera infection: a systematic review of the evidence.* **Paediatrics and international child health**. 2018;38(sup1):S16-s31.

40. Nelson EJ, Nelson DS, Salam MA, Sack DA. *Antibiotics for both moderate and severe cholera.* **N Engl J Med**. 2011;364(1):5-7.

41. Kunkel A, Lewnard JA, Pitzer VE, Cohen T. *Antimicrobial Resistance Risks of Cholera Prophylaxis for United Nations Peacekeepers.* **Antimicrobial agents and chemotherapy**. 2017;61(8).

42. Gladkikh AS, Feranchuk SI, Ponomareva AS, Bochalgin NO, Mironova LV. *Antibiotic resistance in Vibrio cholerae El Tor strains isolated during cholera complications in Siberia and the Far East of Russia.* **Infect Genet Evol**. 2020;78:104096.

43. Mohammed Y, Aboderin AO, Okeke IN, Olayinka AT. *Antimicrobial resistance of Vibrio cholerae from sub-Saharan Africa: A systematic review.* **African journal of laboratory medicine**. 2018;7(2):778.

44. Parvin I, Shahunja KM, Khan SH, et al. *Changing Susceptibility Pattern of Vibrio cholerae O1 Isolates to Commonly Used Antibiotics in the Largest Diarrheal Disease Hospital in Bangladesh during 2000-2018.* **Am J Trop Med Hyg**. 2020.

45. Harris JB, LaRocque RC, Qadri F, Ryan ET, Calderwood SB. *Cholera.* **Lancet**. 2012;379(9835):2466-2476.

46. Utsalo SJ, Eko FO, Umoh F, Asindi AA. *Faecal excretion of Vibrio cholerae during convalescence of cholera patients in Calabar, Nigeria.* **European journal of epidemiology**. 1999;15(4):379-381.

47. Weil AA, Begum Y, Chowdhury F, et al. *Bacterial shedding in household contacts of cholera patients in Dhaka, Bangladesh.* **Am J Trop Med Hyg**. 2014;91(4):738-742.

48. Feachem RG. *Environmental aspects of cholera epidemiology. III. Transmission and control.* **Tropical diseases bulletin**. 1982;79(1):1-47.

49. Nelson EJ, Harris JB, Morris JG, Jr., Calderwood SB, Camilli A. *Cholera transmission: the host, pathogen and bacteriophage dynamic*. **Nat Rev Microbiol**. 2009;7(10):693-702.

50. Wagner EG, Lanoix JN. *Excreta disposal for rural areas and small communities*. **Monograph series World Health Organization**. 1958;39:1-182.

51. World Health Organization. *Boil water.* Geneva: World Health Organization;2015.

52. Cohen A, Colford JM. *Effects of Boiling Drinking Water on Diarrhea and Pathogen-Specific Infections in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis*. **Am J Trop Med Hyg**. 2017;97(5):1362-1377.

53. World Health Organization. Guidelines for drinking-water quality. 4th ed ed. Geneva: World Health Organization; 2011.

54. LeChevallier MW, Au K-K. *Water treatment and pathogen control.* Iwa Publishing; 2004.

55. Boone SA, Gerba CP. *Significance of fomites in the spread of respiratory and enteric viral disease*. **Applied and environmental microbiology**. 2007;73(6):1687-1696.

56. Rusin P, Maxwell S, Gerba C. *Comparative surface-to-hand and fingertip-to-mouth transfer efficiency of gram-positive bacteria, gram-negative bacteria, and phage*. **Journal of applied microbiology**. 2002;93(4):585-592.

57. Scott E, Bloomfield SF. *The survival and transfer of microbial contamination via cloths, hands and utensils*. **Journal of Applied Bacteriology**. 1990;68(3):271-278.

58. Burgess CM, Gianotti A, Gruzdev N, et al. *The response of foodborne pathogens to osmotic and desiccation stresses in the food chain*. **International Journal of Food Microbiology**. 2016;221:37-53.

59. Islam MS, Hossain MA, Khan SI, et al. *Survival of Shigella dysenteriae Type 1 on Fomites*. **Journal of Health, Population and Nutrition**. 2001;19(3):177-182.

60. Nocker A, Fernández PS, Montijn R, Schuren F. *Effect of air drying on bacterial viability: a multiparameter viability assessment*. **Journal of microbiological methods**. 2012;90(2):86-95.

61. Delahoy MJ, Wodnik B, McAliley L, et al. *Pathogens transmitted in animal feces in low- and middle-income countries*. **International journal of hygiene and environmental health**. 2018;221(4):661-676.

62. Zambrano LD, Levy K, Menezes NP, Freeman MC. *Human diarrhea infections associated with domestic animal husbandry: a systematic review and meta-analysis*. **Transactions of the Royal Society of Tropical Medicine and Hygiene**. 2014;108(6):313-325.

63. Khamesipour F, Lankarani KB, Honarvar B, Kwenti TE. *A systematic review of human pathogens carried by the housefly (Musca domestica L.)*. **BMC public health**. 2018;18(1):1049.

64. Islam MS, Zaman MH, Islam MS, Ahmed N, Clemens JD. *Environmental reservoirs of Vibrio cholerae*. **Vaccine**. 2019.

65. Lutz C, Erken M, Noorian P, Sun S, McDougald D. *Environmental reservoirs and mechanisms of persistence of Vibrio cholerae*. **Frontiers in Microbiology**. 2013;4:375.

66. Colwell RR, Brayton P, Herrington D, Tall B, Huq A, Levine MM. *Viable but non-culturable Vibrio cholerae O1 revert to a cultivable state in the human intestine*. **World J Microbiol Biotechnol**. 1996;12(1):28-31.

67. Gallego-Hernandez AL, DePas WH, Park JH, et al. *Upregulation of virulence genes promotes Vibrio cholerae biofilm hyperinfectivity*. **Proceedings of the**

National Academy of Sciences of the United States of America. 2020;117(20):11010-11017.

68. Mitterer F, Pombo JP, Schild S. *Vibrio cholerae Released by Protozoa are Hyperinfectious*. **Trends in microbiology**. 2020;28(1):4-6.

69. Jutla A, Khan R, Colwell R. *Natural Disasters and Cholera Outbreaks: Current Understanding and Future Outlook*. **Current environmental health reports**. 2017;4(1):99-107.

70. Jutla A, Whitcombe E, Hasan N, et al. *Environmental factors influencing epidemic cholera*. **Am J Trop Med Hyg**. 2013;89(3):597-607.

71. Moore SM, Azman AS, Zaitchik BF, et al. *El Nino and the shifting geography of cholera in Africa*. **Proceedings of the National Academy of Sciences of the United States of America**. 2017;114(17):4436-4441.

72. Ericsson CD, Steffen R, Backer H. *Water Disinfection for International and Wilderness Travelers*. **Clinical Infectious Diseases**. 2002;34(3):355-364.

73. Feachem RG. *Sanitation and disease : health aspects of excreta and wastewater management*. Chichester West Sussex ; New York: Published for the World Bank by Wiley; 1983.

74. Morris JG, Jr., Sztein MB, Rice EW, et al. *Vibrio cholerae O1 can assume a chlorine-resistant rugose survival form that is virulent for humans*. **J Infect Dis**. 1996;174(6):1364-1368.

75. Rice EW, Johnson CH, Clark RM, et al. *Vibrio cholerae 01 can assume a 'rugose' survival form that resists killing by chlorine, yet retains virulence*. **International Journal of Environmental Health Research**. 1993;3(2):89-98.

76. Rahman M, Jubair M, Alam MT, et al. *High-frequency rugose exopolysaccharide production by Vibrio cholerae strains isolated in Haiti*. **PLoS One**. 2014;9(11):e112853.

77. Clark Robert M, Rice Eugene W, Pierce Bennett K, Johnson Clifford H, Fox Kim R. *Effect of Aggregation on Vibrio Cholerae Inactivation*. **Journal of Environmental Engineering**. 1994;120(4):875-887.

78. Geldreich EE. *Microbial quality of water supply in distribution systems*. Boca Raton: CRC Lewis Publishers; 1996.

79. Sousa OV, Vieira RHSF, Patel TR, Hofer E, Mesquita VP. *Effects of chlorine on cells of Vibrio cholerae*. **Food Microbiology**. 2001;18(3):355-359.

80. Kolvin JL, Roberts D. *Studies on the growth of Vibrio cholerae biotype eltor and biotype classical in foods*. **J Hyg (Lond)**. 1982;89(2):243-252.

81. Barua D. *Survival of cholera vibrios in food, water and fomites*. **Public health papers**. 1970;40:29-31.

82. Farhana I, Hossain ZZ, Tulsiani SM, Jensen PK, Begum A. *Survival of Vibrio cholerae O1 on fomites*. **World J Microbiol Biotechnol**. 2016;32(9):146.

83. Calfee MW, Wendling M. *Inactivation of vegetative bacterial threat agents on environmental surfaces*. **Sci Total Environ**. 2013;443:387-396.

84. El-Bassiony GM, Luizzi V, Nguyen D, Stoffolano Jr JG, Purdy AE. *Vibrio cholerae laboratory infection of the adult house fly Musca domestica*. **Medical and veterinary entomology**. 2016;30(4):392-402.

85. Hossain Z, Farhana I, Lindeberg Y, et al. *Can cholera fly?* 2016.

86. Kotenok Ia F, Chicherin Iu V. *[Houseflies (M. domestica L.) as transmitters of the agent of cholera]*. **Zh Mikrobiol Epidemiol Immunobiol**. 1977(12):23-27.

87. Purdy AE. *Fly Models of Vibrio cholerae Infection and Colonization*. **Methods in molecular biology (Clifton, NJ)**. 2018;1839:77-96.

88. Bi Q, Ferreras E, Pezzoli L, et al. *Protection against cholera from killed whole-cell oral cholera vaccines: a systematic review and meta-analysis.* **Lancet Infect Dis**. 2017;17(10):1080-1088.

89. Clemens JD, Sack DA, Harris JR, et al. *Field trial of oral cholera vaccines in Bangladesh: results from three-year follow-up.* **Lancet**. 1990;335(8684):270-273.

90. Sur D, Lopez AL, Kanungo S, et al. *Efficacy and safety of a modified killed-whole-cell oral cholera vaccine in India: an interim analysis of a cluster-randomised, double-blind, placebo-controlled trial.* **Lancet**. 2009;374(9702):1694-1702.

91. Fong Y, Halloran ME, Park JK, Marks F, Clemens JD, Chao DL. *Efficacy of a bivalent killed whole-cell cholera vaccine over five years: a re-analysis of a cluster-randomized trial.* **BMC Infect Dis**. 2018;18(1):84.

92. Teoh SL, Kotirum S, Hutubessy RCW, Chaiyakunapruk N. *Global economic evaluation of oral cholera vaccine: A systematic review.* **Hum Vaccin Immunother**. 2018;14(2):420-429.

93. Mogasale V, Ramani E, Wee H, Kim JH. *Oral Cholera Vaccination Delivery Cost in Low- and Middle-Income Countries: An Analysis Based on Systematic Review.* **PLoS Negl Trop Dis**. 2016;10(12):e0005124.

94. Peprah D, Palmer JJ, Rubin GJ, et al. *Perceptions of oral cholera vaccine and reasons for full, partial and non-acceptance during a humanitarian crisis in South Sudan.* **Vaccine**. 2016;34(33):3823-3827.

95. Azman AS, Luquero FJ, Ciglenecki I, Grais RF, Sack DA, Lessler J. *The Impact of a One-Dose versus Two-Dose Oral Cholera Vaccine Regimen in Outbreak Settings: A Modeling Study.* **PLoS Med**. 2015;12(8):e1001867.

96. World Health Organisation. *Cholera vaccines: WHO position paper - August 2017.* **Releve epidemiologique hebdomadaire**. 2017;92(34):477-498.

97. Pezzoli L. *Global oral cholera vaccine use, 2013-2018.* **Vaccine**. 2019.

98. Wolf J, Hunter PR, Freeman MC, et al. *Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression.* **Tropical Medicine & International Health**. 2018;23(5):508-525.

99. Wolf J, Pruss-Ustun A, Cumming O, et al. *Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression.* **Trop Med Int Health**. 2014;19(8):928-942.

100. Cairncross S, Hunt C, Boisson S, et al. *Water, sanitation and hygiene for the prevention of diarrhoea.* **International journal of epidemiology**. 2010;39 Suppl 1:i193-205.

101. Clasen Thomas F, Roberts Ian G, Rabie T, Schmidt W-P, Cairncross S. Interventions to improve water quality for preventing diarrhoea. *Cochrane Database of Systematic Reviews.* 2006(3). http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004794.pub2/abstract.

102. Curtis V, Cairncross S. *Effect of washing hands with soap on diarrhoea risk in the community: a systematic review.* **Lancet Infect Dis**. 2003;3(5):275-281.

103. Ejemot-Nwadiaro Regina I, Ehiri John E, Meremikwu Martin M, Critchley Julia A. Hand washing for preventing diarrhoea. *Cochrane Database of Systematic Reviews.* 2008(1). http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD004265.pub2/abstract.

104. Esrey SA, Feachem RG, Hughes JM. *Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities*. **Bull World Health Organ**. 1985;63(4):757-772.

105. Esrey SA, Potash JB, Roberts L, Shiff C. *Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma*. **Bull World Health Organ**. 1991;69(5):609-621.

106. Fewtrell L, Kaufmann RB, Kay D, Enanoria W, Haller L, Colford JM, Jr. *Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis*. **Lancet Infect Dis**. 2005;5(1):42-52.

107. Freeman MC, Stocks ME, Cumming O, et al. *Hygiene and health: systematic review of handwashing practices worldwide and update of health effects*. **Trop Med Int Health**. 2014;19(8):906-916.

108. Waddington H, Snilstveit B. *Effectiveness and sustainability of water, sanitation, and hygiene interventions in combating diarrhoea*. **Journal of Development Effectiveness**. 2009;1(3):295-335.

109. Najnin N, Leder K, Forbes A, et al. *Inconsistency in Diarrhea Measurements when Assessing Intervention Impact in a Non-Blinded Cluster-Randomized Controlled Trial*. **Am J Trop Med Hyg**. 2019;101(1):51-58.

110. Schmidt WP. *The elusive effect of water and sanitation on the global burden of disease*. **Trop Med Int Health**. 2014;19(5):522-527.

111. Wood L, Egger M, Gluud LL, et al. *Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study*. **BMJ (Clinical research ed)**. 2008;336(7644):601-605.

112. Brown J, Cumming O. *Stool-Based Pathogen Detection Offers Advantages as an Outcome Measure for Water, Sanitation, and Hygiene Trials*. **Am J Trop Med Hyg**. 2020;102(2):260-261.

113. Pickering AJ, Null C, Winch PJ, et al. *The WASH Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea*. **The Lancet Global Health**. 2019;7(8):e1139-e1146.

114. Taylor DL, Kahawita TM, Cairncross S, Ensink JH. *The Impact of Water, Sanitation and Hygiene Interventions to Control Cholera: A Systematic Review*. **PLoS One**. 2015;10(8):e0135676.

115. Colwell RR, Huq A, Islam MS, et al. *Reduction of cholera in Bangladeshi villages by simple filtration*. **Proceedings of the National Academy of Sciences of the United States of America**. 2003;100(3):1051-1055.

116. Huq A, Yunus M, Sohel SS, et al. *Simple sari cloth filtration of water is sustainable and continues to protect villagers from cholera in Matlab, Bangladesh*. **mBio**. 2010;1(1).

117. Conroy RM, Meegan ME, Joyce T, McGuigan K, Barnes J. *Solar disinfection of drinking water protects against cholera in children under 6 years of age*. **Archives of disease in childhood**. 2001;85(4):293-295.

118. Deb BC, Sircar BK, Sengupta PG, et al. *Studies on interventions to prevent eltor cholera transmission in urban slums*. **Bull World Health Organ**. 1986;64(1):127-131.

119. Azurin JC, Alvero M. *Field evaluation of environmental sanitation measures against cholera*. **Bull World Health Organ**. 1974;51(1):19-26.

120. Lantagne D, Yates T. *Household Water Treatment and Cholera Control*. **J Infect Dis**. 2018;218(suppl_3):S147-s153.

121. George CM, Monira S, Sack DA, et al. *Randomized Controlled Trial of Hospital-Based Hygiene and Water Treatment Intervention (CHoBI7) to Reduce Cholera*. **Emerg Infect Dis**. 2016;22(2):233-241.

122. Burrowes V, Perin J, Monira S, et al. *Risk Factors for Household Transmission of Vibrio cholerae in Dhaka, Bangladesh (CHoBI7 Trial).* **Am J Trop Med Hyg**. 2017;96(6):1382-1387.

123. Rafique R, Rashid MU, Monira S, et al. *Transmission of Infectious Vibrio cholerae through Drinking Water among the Household Contacts of Cholera Patients (CHoBI7 Trial).* **Front Microbiol**. 2016;7:1635.

124. Dreibelbis R, Winch PJ, Leontsini E, et al. *The Integrated Behavioural Model for Water, Sanitation, and Hygiene: a systematic review of behavioural models and a framework for designing and evaluating behaviour change interventions in infrastructure-restricted settings.* **BMC public health**. 2013;13(1):1015.

125. Ginja S, Gallagher S, Keenan M. *Water, sanitation and hygiene (WASH) behaviour change research: why an analysis of contingencies of reinforcement is needed.* **International Journal of Environmental Health Research**. 2019:1-14.

126. Martin NA, Hulland KRS, Dreibelbis R, Sultana F, Winch PJ. *Sustained adoption of water, sanitation and hygiene interventions: systematic review.* **Tropical Medicine & International Health**. 2018;23(2):122-135.

127. Valcourt N, Javernick-Will A, Walters J, Linden K. *System Approaches to Water, Sanitation, and Hygiene: A Systematic Literature Review.* **Int J Environ Res Public Health**. 2020;17(3).

128. Tutusaus M, Schwartz K. *Water services in small towns in developing countries: at the tail end of development.* **Water Policy**. 2018;20(S1):1-11.

129. Hsiao A, Desai SN, Mogasale V, Excler JL, Digilio L. *Lessons learnt from 12 oral cholera vaccine campaigns in resource-poor settings.* **Bull World Health Organ**. 2017;95(4):303-312.

130. Shannon K, Hast M, Azman AS, Legros D, McKay H, Lessler J. *Cholera prevention and control in refugee settings: Successes and continued challenges.* **PLoS Negl Trop Dis**. 2019;13(6):e0007347.

131. Bruckner C, Checchi F. *Detection of infectious disease outbreaks in twenty-two fragile states, 2000-2010: a systematic review.* **Conflict and health**. 2011;5(1):13.

132. Campuzano Cuadrado P, Arcos González P. *[Epidemic cholera in complex emergencies].* **Rev Esp Salud Publica**. 2014;88(2):191-201.

133. Baird M. *Service delivery in fragile and conflict-affected states*. 2011.

134. Boot NL, Chen Y, Cohen S, Khayat W, Steele A. *Delivering sustainable water supply in fragile and conflict affected states: experiences from Syria.* 2015.

135. Batley R, Mcloughlin C. *State capacity and non-state service provision in fragile and conflict-affected states*. 2009.

136. Kennedy J, Harmer A, McCoy D. *The political determinants of the cholera outbreak in Yemen.* **Lancet Glob Health**. 2017;5(10):e970-e971.

137. Jeandron A, Cumming O, Rumedeka BB, Saidi JM, Cousens S. *Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo.* **PLoS One**. 2018;13(8):e0201306.

138. Jeandron A, Cumming O, Kapepula L, Cousens S. *Predicting quality and quantity of water used by urban households based on tap water service.* **npj Clean Water**. 2019;2(1):23.

139. Jeandron A, Saidi JM, Kapama A, et al. *Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo.* **PLoS Med**. 2015;12(10):e1001893.

2

Data and methods: an overview

The present research in Uvira required a variety of data and estimates to be collected, compiled and generated from different sources – this chapter aims to introduce the study location and context, and describe in detail the sources and methodologies involved in creating the datasets used in further chapters. It also provides explanations for methodological choices and compromises made.

2.1 Study area

Uvira is the second largest city of the South-Kivu province in the Democratic Republic of the Congo (DRC) after the provincial capital Bukavu, with approximately 250,000 inhabitants in 2017. Uvira is located on the shores of the northern extremity of Lake Tanganyika, at the border with Burundi, between the following coordinates (North/West/South/East): 3.338° S, 29.123° E, 3.437° S, 29.2° E (**Figure 2-1**). Uvira is delimited in the north by the Ruzizi plain, in the South by Fizi territory, and in the West by the Mounts Mitumba.

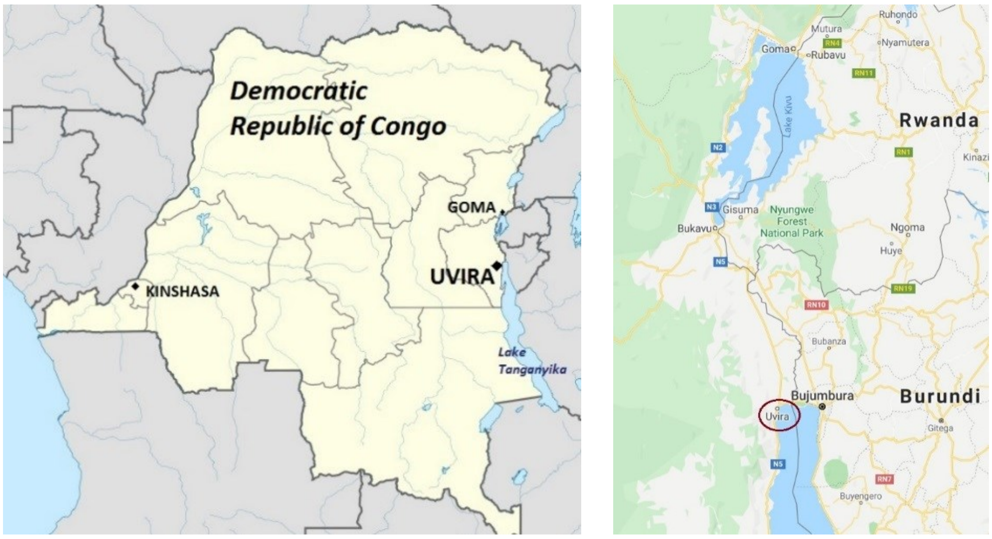


Figure 2-1 Location of Uvira within the Democratic Republic of the Congo (left – adapted from [1]) and relative to Rwanda and Burundi (right [2])

The town is located at an average altitude of approximately 800 m and stretches over nearly 10 km between north and south and a maximum of 2 km between the lake shores and Mounts Mitumba steep foothills. According to the Köppen and Geiger classification, Uvira enjoys a tropical savannah climate, with an average monthly temperature ranging between 21.9 °C and 23.4 °C and an average annual precipitation of 1,132 mm **[3]**. The months of June, July and August are usually much drier and slightly cooler than the rest of the year (**Figure 2-2**).

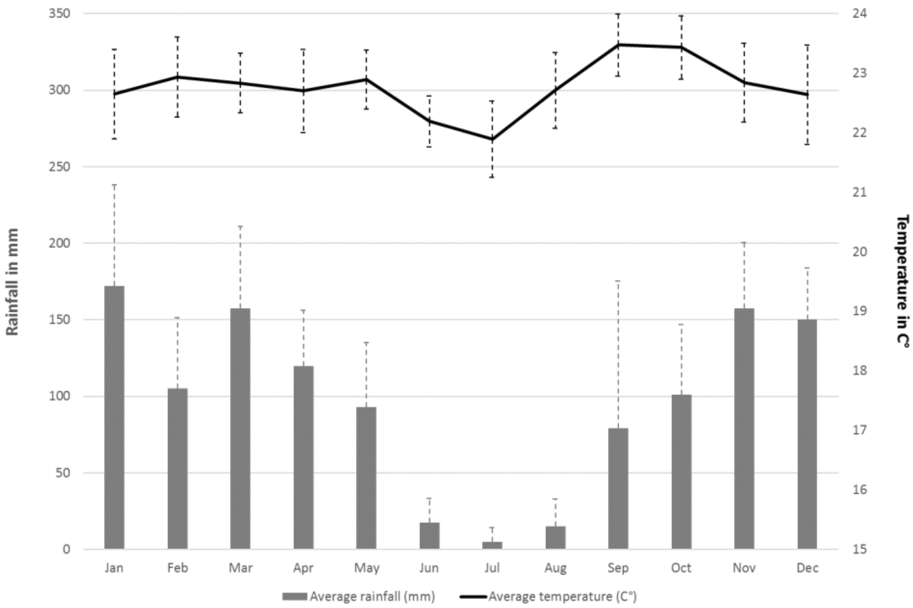


Figure 2-2 Average monthly rainfall (bottom)and temperature (top) for years 1991 to 2016 in Uvira, at coordinates 3.37°S,29.13°E [adapted from [3]). Dotted lines represent standard deviation from mean values.

Three large rivers cross Uvira before flowing into Lake Tanganyika. These are, from North to South: river Kavimvira, river Mulongwe and river Kalimabenge. Two smaller rivers also cross the southernmost neighbourhoods: river Kamongola and river Karigo. At the northern end of town, Nyangara is a large waterlogged area. Uvira spans along two sporadically asphalted roads: the RN5 that links Bukavu to Lubumbashi, and the RN4 towards the East and Bujumbura. The port of Kalundu, at the southern end of the town, operates passenger and cargo services on the lake to Kalemie and Moba in the Katanga province in southern DRC, and to Burundi, Tanzania and Zambia. Some key topographical elements of Uvira and its surroundings are represented in **Figure 2-3**.

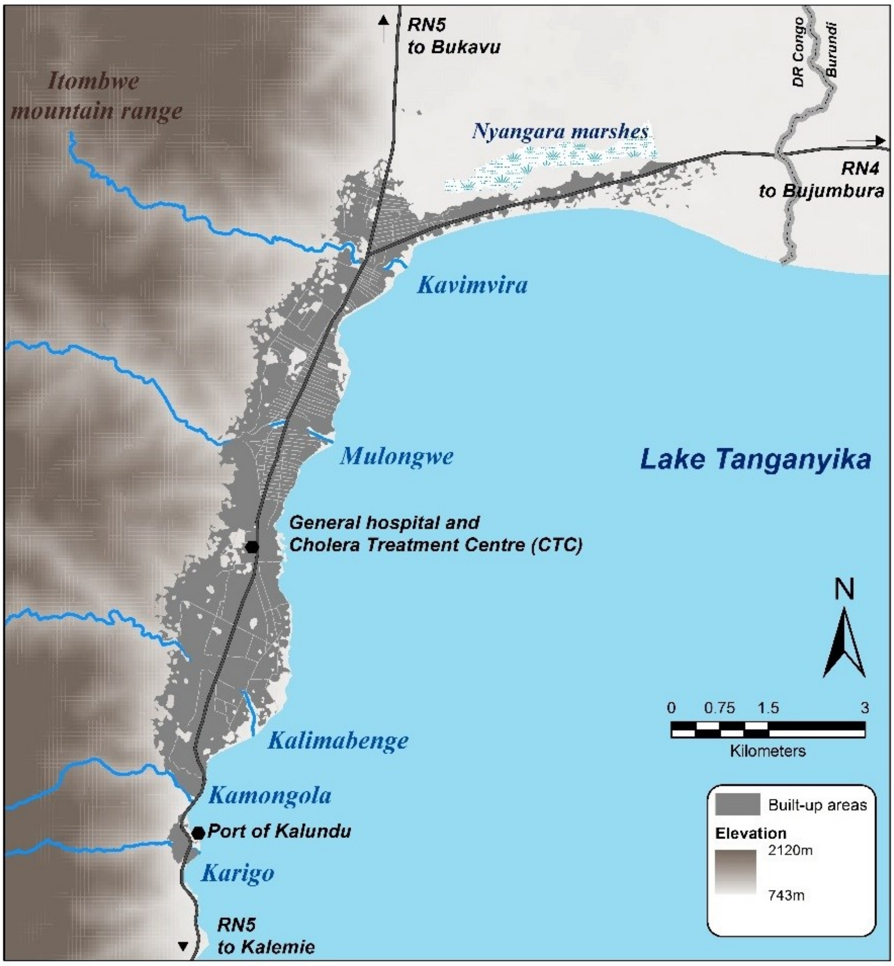


Figure 2-3 Uvira and its surroundings

Economic infrastructure in the Uvira area has been in a derelict condition for several decades and most of the population in town relies on subsistence farming on the outskirts, fishing on Lake Tanganyika, unskilled labour, small craft businesses and small trade. Electricity in Uvira is distributed by the Societe Nationale de l'Electricité (SNEL) and produced by two hydro-electric plants on the Ruzizi river that also serve Bukavu, and the capital of the North-Kivu province, Goma. Production capacity is frequently insufficient for the regional needs, especially during the dry season **[4]**. This results in frequent power interruptions in Uvira, also exacerbated by the poor condition of the distribution network.

Despite improvements in the political and economic situation of the country since 2005, DRC still has the 2nd highest incidence of extreme poverty in sub-Saharan Africa after Madagascar and ranks 179 out of 189 on the 2019 Human Development Index (HDI)**[5,6]**.

The eastern provinces of DRC are particularly affected by what is often referred to as one of the world's largest and most complex humanitarian crises, and Uvira is no exception. The Office for Coordination of Humanitarian Affairs (OCHA) estimated that during the second half of 2019, nearly half a million people would be targeted for emergency humanitarian assistance in South-Kivu alone, through nearly 50 different operational partners (mostly national and international non-governmental organisations [NGO and INGO] and United Nations agencies/programs), requiring an estimated 68.3 million USD **[7]**. A comparison with the official provincial budget of 111 million USD for 2019 illustrates the ongoing reliance of the province upon external assistance and humanitarian aid.

This situation has been fuelled for decades by conflict and insecurity. Uvira and its surroundings have been particularly affected by conflict since the mid-nineties and the First and Second Congo wars (1996-2003). Uvira territory still has one of the highest concentrations of armed groups in the Eastern Congo, with violence destabilizing all aspects of society, and influencing local power structures and governance **[8]**. A contingent of the United Nations (UN) peacekeeping forces (Mission de l'Organisation des Nations unies pour la stabilisation en République démocratique du Congo or MONUSCO) has been based in Uvira since 1999 with the mandate of protecting civilians and promoting human rights.

Although the overall socio-economic and humanitarian situation in the town of Uvira itself is less precarious than in the surrounding rural areas, it remains volatile and is regularly disrupted by the influx of internally displaced people (IDP) or refugees – for example thousands of people fleeing Burundi following the 2015 presidential elections and attempted coup - international tensions with neighbouring Burundi, and even heavy artillery fighting like that in September 2017 when the Mai-Mai rebel group Yakutumba attempted to seize the town only to be repelled by MONUSCO intervention.

2.2 Mapping the study area

2.2.1 Official administrative and health divisions

During most of the research period, the town of Uvira had the administrative status of a "cité", divided into 14 neighbourhoods ("quartiers"), and administered by the town hall ("Bureau de cité"). Each of these neighbourhoods are divided into smaller geographical units usually around a main street ("avenue") and supervised by a chief ("chef d'avenue"). Since January 2019, Uvira has the status of a "ville" administered by a mayor, divided into three "communes". Neighbourhood divisions within these communes are believed to remain the same. The town of Uvira is administered separately from the Uvira territory ("territoire d'Uvira"), that governs areas surrounding Uvira.

Health authorities use a different division system, that partially overlaps the administrative one. Within South-Kivu province, four health zones ("Zone de Santé" or ZS) are defined, one being Uvira ZS. Uvira ZS is divided into 21 health areas ("Aire de santé" or AS), and 16 of these AS overlap with one or more of Uvira city neighbourhoods according to the administrative division system. Population figures for ZS and AS are not regularly compiled or centralised.

2.2.2 Delineating the administrative and health divisions

As of mid-2018, no official geographical delineation of either administrative or health divisions at the scale of Uvira was available from sources outside the present research study. Although street names and house numbers are seen in some areas of town, they are not systematically defined or used in all neighbourhoods.

A mapping effort was started by the Veolia Foundation in 2011 with the health authorities, in order to locate retrospectively the residence of patients having attended the Uvira Cholera Treatment Centre (CTC). AS were the first geographical units to be roughly delimited, followed by the streets they included, and large visual references on Google Earth maps were used along local knowledge to draw the units' limits. The initial map was then refined over time by walking streets and neighbourhoods with their respective chiefs and following the limits of units with a Global Positioning System (GPS) device. This step had to be repeated several times to adjust for large streets being officially divided up, and to update street names to match those used in population censuses of administrative divisions.

The most recent and significant changes in the delineation of streets and neighbourhoods occurred in spring 2018, when a long lasting dispute on the official limits between the town and the territory came to light in two areas, on the outskirts of neighbourhoods Kabindula and Songo, and on the outskirts of neighbourhoods Kakombe and Kibondwe.

Several streets initially considered in our study as part of Uvira town were then excluded, as they are not actually administered by the town but rather by the territory.

Figure 2-4 shows the 14 neighbourhoods, 16 AS and 258 streets across the town of Uvira, including the disputed areas. Details of the 258 streets and their neighbourhoods are given in **Appendix 2-1**.

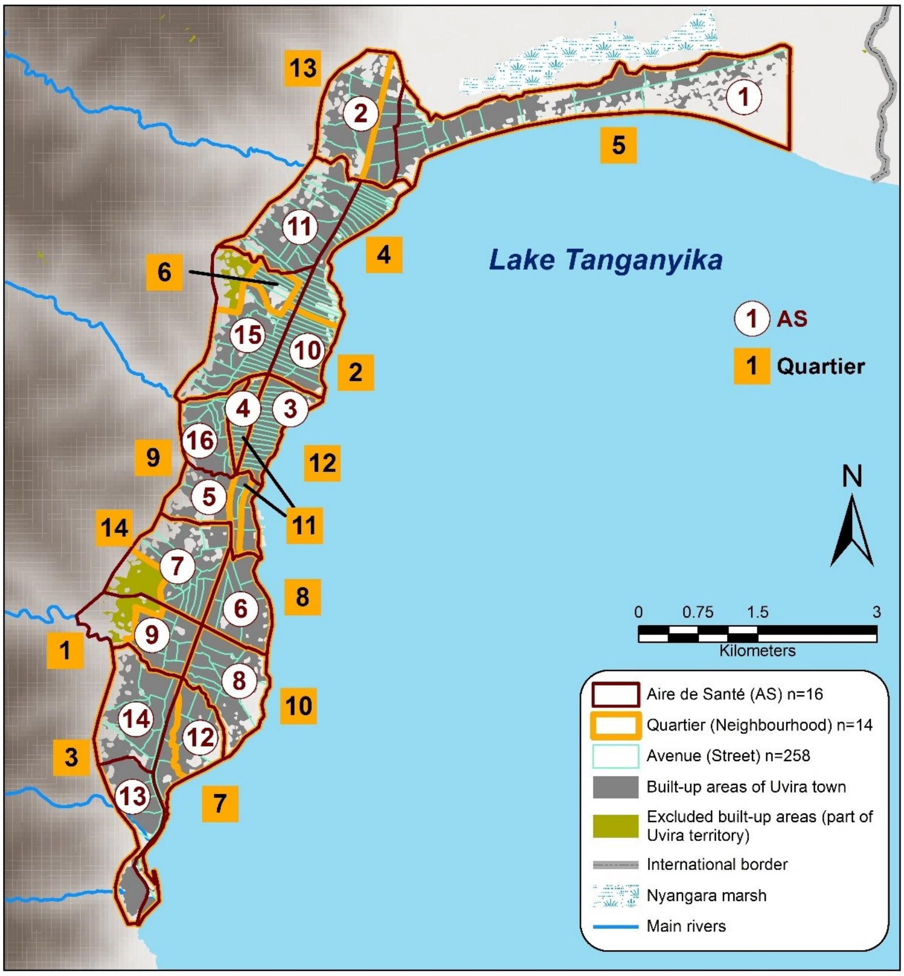


Figure 2-4 Health and administrative divisions of Uvira

<u>AS names</u>: 1 AS Kilomoni; 2 AS Kavimvira; 3 AS Mulongwe; 4 AS Rombe 1; 5 AS Tanganyika; 6 AS Kimanga; 7 AS Saint-Paul; 8 AS Nyamianda; 9 AS Kabindula; 10 AS Kasenga Etat; 11 AS Kiyaya; 12 AS Kalundu Etat; 13 AS Kalundu Catholique; 14 AS Kalundu Cepac; 15 AS Kasenga Cepac; 16 AS Mitumba.

<u>Quartiers names</u>: 1 Q. Songo; 2 Q. Kakombe; 3 Q. Kalundu; 4. Q. Kasenga; 5 Q. Kavimvira; 6 Q. Kibondwe; 7 Q. Kilibula; 8 Q. Kimanga; 9 Q. Mulongwe; 10 Q. Nyamianda; 11 Q. Rombe 1; 12 Q. Rombe 2; 13 Q. Rugenge; 14 Q. Songo.

2.2.3 Geographical units of analysis for the study

In the absence of an official address system used by the authorities and population alike, a major challenge for the present study was to identify, geolocate and characterise geographical units for which data – related to population, CTC admissions, tap water access and other potential risk factors such as population density - was most complete and accurate. The resolution of spatial units - their size and number - also had to be considered for the specific analysis needs.

One of the main constraints was the availability of population estimates, which is limited to the administrative authorities and the geographical units they use for population censuses. The other was the precision with which people admitted the CTC reported their place of residence especially when streets have the same name but a different number – for example streets Shishi 1, Shishi 2, Shishi 3 and Shishi 4, often recorded as Shishi without further precision. Another difficulty was to account for streets with exactly the same name – for instance street "avenue Kamanyola" found in 4 distinct neighbourhoods and AS. Finally, some streets are divided across neighbourhoods and AS, making the localisation of a residence only possible with the combination of street name, neighbourhood and AS altogether.

The smallest geographical unit for which CTC patient residence, population and geographical boundaries were regarded as sufficiently accurate for the present study is referred to as an Analysis Base Unit (ABU) and is either a street or a grouping of streets. Ultimately the 203 ABU were grouped into lower resolution units for analysis, at the levels of the entire town, 3 areas (North, Centre and South), 16 AS, 14 neighbourhoods, and 37 sub-neighbourhood divisions.

2.2.4 Built-up areas delineation

Although most data were available at a discrete geographical level (street, AS or neighbourhood), the collection of point data such as georeferenced interviewed households or tap water connections made it desirable to use a continuous geographical mapping of the population. This would avoid considering population being homogeneously spread across simplified shapes of geographical units, ignoring large uninhabited areas at the margins of town or within some units. Most buildings in Uvira are single storey constructions, and few are not used for accomodation, and a proxy for mapping population was therefore chosen as built-up areas.

An increasing number of geographical open source data has become available during the present study, with crowdsourcing initiatives – for example Open Street Map[1] or Missing Maps[2] - and remote sensing data analysis such as Global Urban Footprint[3] or WorldPop[4]. None of these sources provided however sufficient up-to-date information at a high enough resolution to use directly.

---

[1] Available at http://www.openstreetmap.org
[2] Available at http://www.missingmaps.org
[3] Available at https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-9628/16557_read-40454/
[4] Available at http://www.worldpop.org

Satellite image classification and early attempt

An attempt to identify buildings and built-up areas across town was made in 2015 by classifying a high-resolution multi-spectral (4 bands – 3 colours + infrared) satellite map of Uvira. A high-resolution multispectral and panchromatic satellite image was purchased from the commercial provider Mapmart (Lone Tree, Colorado, USA). It was taken on the 23rd of September 2015, with a resolution of 2 m and was ortho-corrected for altitude.

In brief, the spectral signature of 23 different classes (for example roof / blue roof / red roof, lake shore / lake) was manually sampled over areas totalling 0.13 km^2 across town. Examination of spectral values distribution and covariance over the 4 bands allowed a grouping of training areas into 7 classes with minimal signature overlapping. An automated maximum likelihood function was then used to classify each 2 x 2 m cell of the entire satellite image into one of the seven trained classes. To reduce misclassification of single cells, the classified image was divided into hexagonal tiles with 10 m sides (approximately 260 m^2) which were classified as built upon if the majority of the cells overlapped by each hexagon was classified as "roof". Once intersected with town boundaries, the merged surface of all built-up hexagons totalled 9.4 km^2 (out of 18.3 km^2 taken into consideration). This process is illustrated in **Appendix 2-2**. This built-up area delineation was used for sampling households in a survey of water-related practices in 2016 (see chapter *Tap water service and households' water-related practices*).

Manual building identification

In 2018, the need for a more reliable and reproducible mapping of built-up areas became even greater and the decision was made to actually map every single identifiable construction in the town, with an adaptation of the methods used by Missing Maps and other crowd sourcing mapping initiatives. For the entire town area, a "buildings" geospatial layer for Uvira and its surroundings was downloaded from Open Street Map (OSM). With a last recorded update dating from 2011, it provided a partial mapping of buildings across the town. It was then complemented by the high-resolution satellite image purchased in 2015. OSM building polygons were converted into points at their centroid location (N = 30,346) and superimposed on the satellite image. Points were manually verified and updated with points created at the centre of each discernible building (N = 49,857).

The mean surface area of buildings drawn as polygons in OSM was 52 m^2, and this was reflected by buffering each point with a radius 4 m, ignoring overlap. The buffered points were then used with the "delineate built-up areas" function in ArcMap (ESRI, Aylesbury, UK), to delineate densely clustered arrangements of buildings into polygons. Parameters for the function were chosen to consider groupings of buildings within a distance of 50 m, a minimum length of non-built area ("hole") of 10 m, and a minimum number of buildings per cluster of three. The built-areas polygons' boundaries were constrained by the inclusion of edge features: lake shore, five permanent rivers and main roads or streets. The process is represented in **Figure 2-5**.

The delineating function identified 127 independent buildings not included in the resulting built areas polygons. Considering their small number, these buildings were excluded from all analyses involving built-up areas. The total built-up surface for Uvira town (excluding disputed areas with the territory) was estimated to 11.96 km^2, out of a surface of 17.88 km^2 when using hand-drawn boundaries.

Built-up areas were combined with the boundaries of neighbourhoods and streets. For each street a constant population density was assumed across built-up areas and a continuous map of population density across town was created using street level population estimates (see section *2.3 Population estimates over time* below).

This method would ideally have been performed by several individuals to enable quality checks, over images taken at different times during the study period and outlining each building shape instead of representing it by a point. Resources were unfortunately not available for these labour-intensive enhancements.

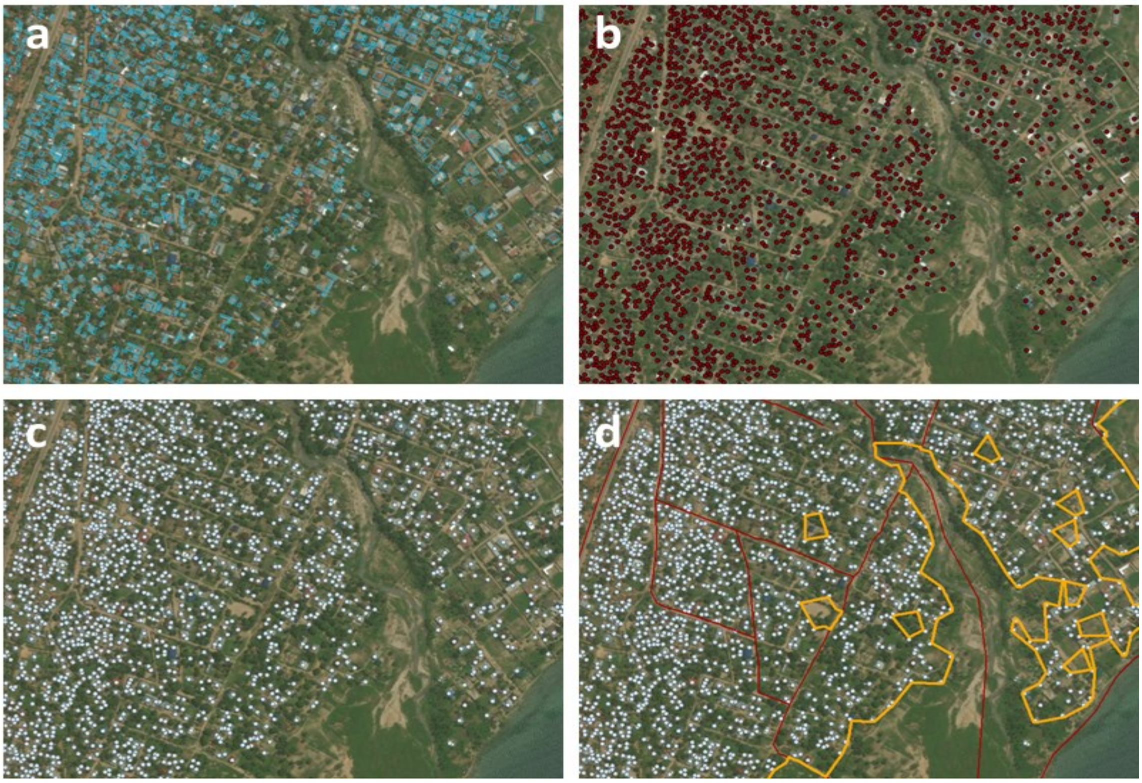


Figure 2-5 Built-up areas delineation process: a) Open Street Map buildings layer; b) Open Street Map buildings converted to points and checked/augmented; c) Building points buffered with a radius 4m; d) built-up areas delineated (yellow) as a function of building clusterings and inclusion of edge features (red)

2.3 Population estimates over time

Quarterly population census records were available at the town hall for 10 time points between June 2008 and November 2017 at the neighbourhood level (14 quartiers). The only comprehensive population census at the higher resolution of streets ("avenues") was obtained in November 2017. A full list of streets and neighbourhoods with population census data for November 2017 is available in **Appendix 2-3**.

In order to estimate populations over time at the street level, we first linearly interpolated the neighbourhood estimates for each week between the available time points. For predictions beyond November 2017, we postulated that the population growth would be equal to the mean growth observed between the three last recorded data points (November 2015, May 2016 and November 2017). Data over time for neighbourhoods is summarised in **Appendix 2-4**.

We then assumed a constant population growth in all streets within a neighbourhood and calculated the proportion of a neighbourhood's population living in each street as reported in the November 2017 census. These weights were then applied to the weekly neighbourhood estimates to obtain the corresponding weekly estimates of streets' populations (**Figure 2-6**).

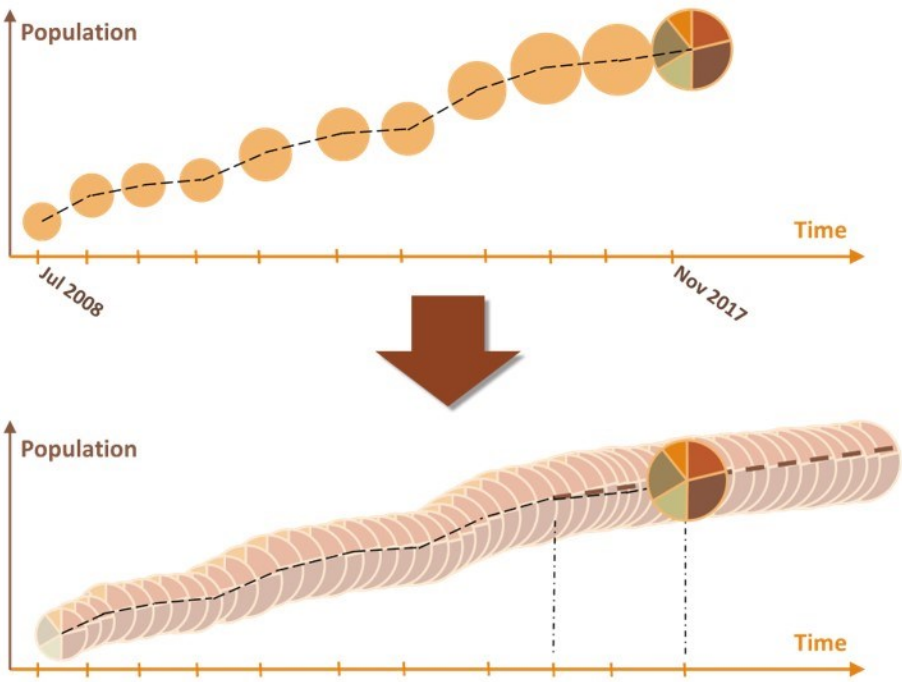


Figure 2-6 Schematic illustration of population estimates over time for a hypothetical neighbourhood (circle) and its streets (circle sections)

2.4 Suspected and confirmed cholera incidence data

2.4.1 The Uvira Cholera Treatment Centre

Since 2004, the Cholera Treatment Centre (CTC) of Uvira is located in the grounds of the Uvira District Hospital ("Hôpital Général de Référence" or HGR - location depicted in **Figure 2-3**). Patients presenting directly to the hospital or referred by local health posts with acute diarrhoea, defined as three or more loose or liquid stools in 24 h, are assessed and admitted as required. They are treated for dehydration with oral rehydration salts (ORS) and/or intravenous (IV) Ringer's lactate solution. Children under 5 years of age are occasionally administered zinc as well, while patients whose symptoms do not resolve after three days may be prescribed broad-spectrum antibiotics (single dose 300 mg doxycycline mostly). Albendazole is also occasionally distributed upon discharge during mass drug administration campaigns against intestinal parasites. Admission to and treatment at the CTC are free of charge for all patients, although payments for IV materials purchased at the hospital pharmacy have been reported when the CTC has run out of stock.

Patients admitted to the CTC may be referred to other hospital departments if they also present complicating medical conditions requiring additional care, such as diabetes, malaria or pregnancy.

All patients admitted to the CTC are considered as suspected cholera cases and are reported as such weekly to the district health office ("Bureau Central de Zone" or BCZ) for further reporting to the national level. At least one admission has been recorded by the Uvira CTC for every week except three between 2004 and May 2019. When CTC admissions exceed 25 patients per week, the BCZ and the provincial health office ("Direction Provinciale de la Santé" or DPS) are alerted and additional resources mobilised to strengthen treatment capacity and potentially initiate a broader outbreak response. National and international NGOs may be involved.

Figure 2-7 Cholera treatment centre of Uvira: communal areas for carers (left) and pre-admission observation area

2.4.2 Patient admissions data

Details of each patient admitted to the CTC are recorded on a paper register by nursing staff. These include a unique ID, the date of admission, patient name, age and sex, residence location and status at discharge. Other information was less systematically recorded over the study duration: CTC shift at admission, patient occupation, rehydration plan(s) administered, other treatments given, particularities such as pregnancy, date of discharge.

During epidemiological investigations and scoping visits in 2010 and 2011, the Veolia Foundation team started entering the data available in the CTC registers from 2009, and retrospectively reconciling recorded residence locations with their streets and AS nomenclature.

In November 2015, a formatted A3 sized register with standardised information to be recorded was agreed with the BCZ and introduced to the CTC along with staff training, to improve data completeness and quality for research purposes. In particular, residence location recording was strengthened by adding a column for neighbourhood ("Quartier"), to ensure that a residence location is identified by a combination of street, neighbourhood and AS. Emphasis was also placed on recording exact age of the patient and date / status at discharge.

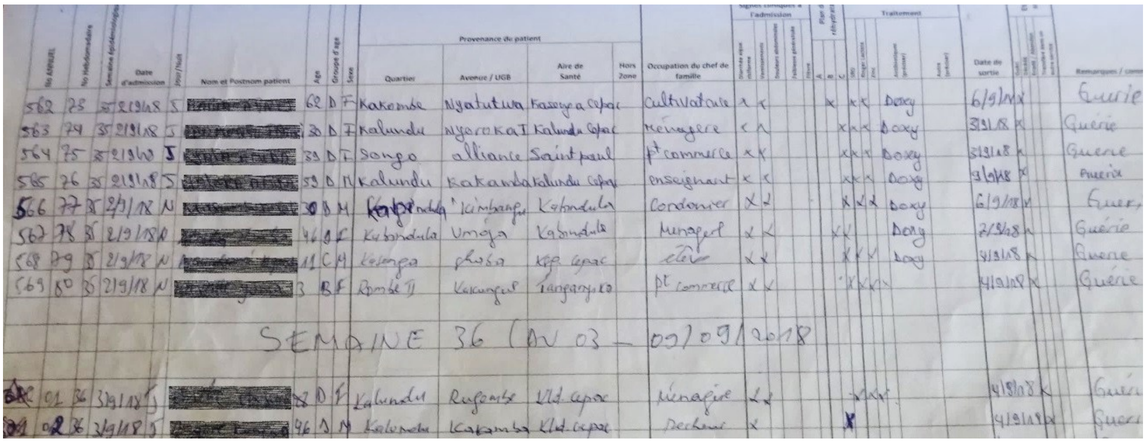


Figure 2-8 Example of entries in the CTC register in 2018

2.4.3 Cholera confirmation with rapid diagnostic tests

From April 2016, a laboratory technician at Uvira District Hospital was trained to collect a rectal swab from all consenting newly admitted patients present at the CTC during daily morning visits 7 days per week. The swab was then enriched in a vial containing 4 ml of Alkaline Peptone Water (APW) for approximately 6 h at ambient temperature (approximately 25 °C +/- 2 °C). The sample was then tested by dipping a Crystal® VC rapid diagnostic test (RDT) dipstick (Span Diagnostics, Surat, India) into a few drops of the upper portion of the enriched APW and reading the results after about 5 min.

In this setting, a rectal swab was preferred to a stool sample to reduce the risk of sample contamination, either across patients or with disinfectant, and to limit stool handling by clinical staff. The use of Crystal® VC RDTs on enriched rectal swabs was evaluated to have a sensitivity and specificity of 92 % and 91 % respectively in a previous study **[9]**. APW is the recommended enrichment medium for all vibrio species and the enrichment step has been shown to improve significantly the specificity of Crystal® VC RDTs estimated at 71 % when used directly on stool samples **[10]**.

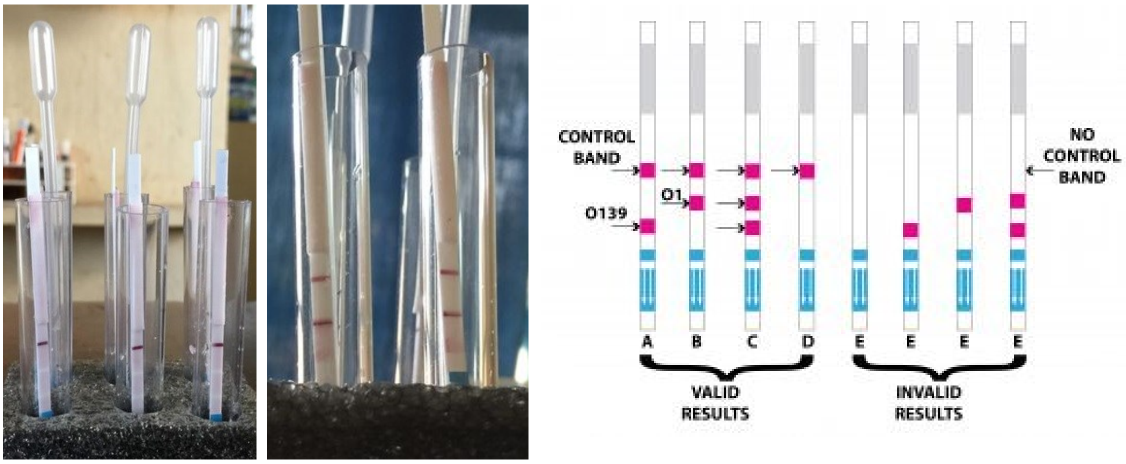


Figure 2-9 Crystal® VC dipstick negative results (left), positive for *vibrio cholerae* O_1 (middle) and interpretation instructions (right) [11]

In its recommendations for cholera surveillance, the Global Task Force on Cholera Control (GTFCC) defines confirmed cholera cases as any suspected case with *Vibrio cholerae* O_1 or O_{139} confirmed by culture or polymerase chain reaction (PCR) **[12]**. In the present case, we chose to refer to participants tested positive with cholera RDTs as *RDT-positive cases*.

Demographic and clinical information was collected from the patient or a family member and the CTC patient register at the time of rectal swab sampling. If antibiotic treatment had already been initiated, the number of hours between the first dose (usually a single dose 300 mg doxycycline when administered at the CTC) and sampling was also recorded.

RDT results were not revealed to the CTC clinical staff to ensure that the same standard of care was delivered for all patients, confirmed or not.

2.4.4 Data management

For the present study, anonymised CTC admissions data were entered periodically from the CTC registrer into a spreadsheet, starting from the 1st of January 2009. Until April 2016 and the beginning of the cholera confirmation study, completeness and accuracy of data entry were verified against the register hard copy through counts and data validation checks. Data between 2009 and 2013 were systematically checked retrospectively when the definition of geographical units was strengthened in 2014.

From April 2016, anonymous data collected for the confirmation study were both recorded in a dedicated paper register and directly entered onto a tablet through digital forms linked up with an online Open Data Kit (ODK) database hosted at the LSHTM. Any discrepancies between the ODK data and the admissions database were checked against both CTC and confirmation paper registers. Approximately 10% of the confirmation records were verified against the test picture collected via ODK.

Information on the symptoms at presentation and treatment received at the CTC (rehydration plan A, B or C, antibiotics) recorded in the CTC register were deemed too unreliable to use in absence of a possible cross-checking with individual patient files. Efforts to strengthen completeness and accuracy of data collected in the CTC register were generally hampered by the fast turn-over of clinical staff, often appointed to the CTC as part of a rolling posting across hospital departments.

As of the 20th of May 2019, a total of 13,506 admissions were recorded at the CTC since the 1st of January 2009 and included in the study database. This database includes the date of admission, sex and age of the patient, residence as a combination of street, neighbourhood and AS information, end of CTC treatment date, status when leaving the CTC (discharged, transferred or deceased). From April 2016 it also records participation in the confirmation study, RDT results and time between antibiotics first dose and sampling.

Table 2-1 Missing information from CTC records, before and after the provision of an A3 formatted register and related training

	Missing information (% of total number of records)	
	January 2009 to November 2015 n = 9490	November 2015 to May 2019 n = 3935
Admission date	0 (0%)	0 (0%)
Sex	19 (0.2%)	2 (0.1%)
Exact age	4497 (47.4%)	32 (0.8%)
Age group	3 (0.03%)	10 (0.3%)
Discharge date	7889 (83.1%)	345 (8.8%)
Status at discharge	349 (3.7%)	259 (6.6%)
Patient residence		
Aire de Santé AS	45 (0.5%)	12 (0.3%)
Neighbourhood	185 (1.9%)	9 (0.2%)
Street	772 (8.1%)	227 (5.8%)
Analysis Base Unit ABU	214 (2.3%)	46 (1.2%)

2.5 Piped water supply infrastructure and data

2.5.1 Overview of the piped water system

The population relies on both surface water sources, and the tap water system managed by the national water agency Regideso. Water from the river Mulongwe is drawn upstream of inhabited areas, treated by sedimentation, flocculation and chlorination at the Regideso water treatment plant before being fed into a single 1,600 m^3 water reservoir, from which it is distributed to private and shared taps across town by gravity. A small 160 m^3 reservoir, located at the extreme south of the town, is used to draw water from two ground water sources and distribute it after minimal treatment by gravity to the local tap connections, independently of the rest of the distribution network.

The current piped water distribution system, initially built in the late eighties for approximately 62,000 inhabitants **[13]**, has been extended over time to reach more areas of Uvira, although in practice its gravity design fails to serve adequately the taps located further away from the reservoir or at higher altitude due to pressure losses. The amount of water treated at the plant is also irregular, due to the intermittent power supply from the local grid, limited resources for generator fuel, stockouts of chemicals for water treatment and frequent downtime for preventive maintenance or repairs of equipment. The geological characteristics of the area are not favourable for groundwater sources, and there are no wells or boreholes in Uvira.

2.5.2 Water treatment station

The water treatment station is located on the river Mulongwe in the centre of town (altitude 810 m). The river water intake chamber is located upstream of the station and most inhabited areas (altitude 817 m), with a pre-treatment stage consisting of a simple grid and sand/mud removal by settlement. At the entrance of the station, water goes through an aeration cascade and a coagulation/flocculation/decantation process with addition of aluminium sulphate. After sand filtration, lime is added for softening/remineralisation, and water is finally chlorinated.

Figure 2-10 Water intake chamber on the river Mulongwe (left) and decantation reservoirs at the water treatment station (right)

Most of the water is then pumped from a temporary treated water reservoir (capacity 2 x 200 m^3) to a single water reservoir of 1,600 m^3 located on the higher ground of the Tanganyika neighbourhood, at an altitude of 890 m (tank bottom). In 2010, three pumps were dedicated to that purpose, with maximum flow capacity of 157 m^3/h and 2 x 80 m^3/h, although only the two smaller pumps were then functional. One of the 80 m^3/h pumps was also used intermittently to feed a small water reservoir (capacity 450 m^3) for a distribution system located 17 km north of Uvira, in the municipality of Kiliba.

The water treatment plant is supplied with electricity by the SNEL and relies on its own generator to cover the frequent interruptions in SNEL supply, although at lower power output and much higher cost. On grid supply, it was estimated that the treatment plant could deliver approximately 155 m^3/h to the reservoir, while on the generator it would deliver only two thirds of that output.

Data on the treatment station operations (inputs and outputs) are collated monthly into production and quality summary sheets. The production report contains information on the volume of water treated and fed to Tanganyika and Kiliba reservoirs, the amount of chemicals consumed, and the sources of power used. It also lists the interruption times dedicated to recurrent maintenance (sand filter washing for example). The quality report includes information on the quality of raw water at intake (pH and turbidity) and quality of treated water (pH, turbidity and residual chlorine), with parameters measured in the station laboratory when measuring equipment and consumables are available.

For the present research, data was collected on the daily volume and average concentration of residual chlorine in treated water produced and pumped up to the Tanganiyka reservoir. Out of 3,770 daily records between the 1st of January 2009 and the 28th of April 2019, 217 (5.8 %) records of residual chlorine were missing, and none for the volume of water produced.

Over the study period, the daily volume of water delivered to the Tanganyika reservoir ranged from 0 m^3 to 6,341 m^3 (mean 3,122 m^3; standard deviation SD 1,058 m^3). The recorded concentration of residual chlorine in treated water pumped to the reservoir - average of up to three daily measures measured by means of DPD (N,N diethyl-p-phenylene diamine) colorimetry – ranged from 0.1 to 3.3 mg/l (median 0.7; Interquartile range IQR 0.6 – 0.8).

2.5.3 Water distribution network and taps

From the Tanganyika reservoir, tap water is distributed by gravity throughout the town of Uvira via more than 70 km of cast iron, steel and PVC pipes. The secondary network – defined as pipe network of diameters <= 90 mm, excluding the very final connection to each individual tap – represents approximately 75 % of the total network length.

In August 2012, 3,016 functional taps were georeferenced across this network, mapped by the Veolia Foundation and the Regideso. Those taps were located at a distance from the reservoir ranging from 15 m to more than 9 km, and a maximum elevation difference of 147 m. Terrain and reservoir location tend to favour distribution to taps located in the southern part of town, which is also the area that historically was first densely populated in the 1970s.

In April 2018, a new detailed georeferencing of taps was undertaken as the Regideso service database had recently become accessible to the Uvira Regideso office. A total of 5,678 taps recorded in the Regideso database were geolocalised, among which a large majority (4,065 or 71.6 %) were considered as active tap connections. About half these connections are invoiced based on the consumption measured by a meter; the others do not have a functional meter and are invoiced based on past consumption, estimated number of users and consumption variation at neighbouring taps with a meter. A map of georeferenced taps in April 2018 is available in **Appendix 2-5**.

Regideso database use is highly restricted, and only invoicing data for the previous month (March 2018) could be extracted and analysed. It indicated that 60,925 m^3 were invoiced to customers, out of 75,548 m^3 treated. This 80 % ratio is similar to a figure estimated in 2010 during a feasibility study. The 20 % difference between production and distribution is attributable to water leakage from damaged pipe network and inaccurate invoicing, especially for those based on estimates rather than meter readings.

Figure 2-11 Typical tap connection with the tap protected by a lock and meter by a plastic cover (left); exposed water pipe following soil erosion (right)

2.5.4 Improvements of the piped water infrastructure 2013-2020

In the frame of the National Multi-sectorial Plan for the Elimination of Cholera 2013-2017, it was announced that the water supply and distribution system in Uvira would be substantially improved thanks to a €10 million grant from the French Development Agency (AFD), the European Union (EU), the Veolia Environment Foundation (VEF) and OXFAM UK.

These works aimed at improving the population of Uvira’s tap water access by substantially refurbishing the entire tap water supply system operated by the national water agency Regideso. The project specifically aimed at 1) increasing water treatment capacity, 2) strengthening tap water distribution with the partial refurbishment and improvement of the pipe network and the building of a second water reservoir and 3) the set-up or renovation of 115 community-managed public taps and installation or refurbishment of 5,000 private taps across the city.

Although these activities were initially planned to start in mid-2016 and last 18 months, many administrative and logistical challenges postponed the beginning of the works to the end of 2017, and as of end-2019, completion was expected at the end of 2020.

Figure 2-12 Construction of a new reservoir on the hills of Kiyaya overlooking northern Uvira (left); new water pipes being laid in Kasenga neighbourhood (right)

2.6 Drinking and environmental water quality analyses

The quality of drinking water in households and of surface water (lake and rivers) was evaluated on several occasions at the Centre de Recherche en Hydrobiologie (CRH) of Uvira. Although the laboratory facilities are rudimentary, especially for the microbiological analyses, the CRH's relatively central location, its mandate for research on water and readiess to collaborate made it an excellent partner for this work.

2.6.1 Households' stored drinking water

Water used for drinking in households was sampled and analysed on two separate occasions during surveys of households' water-related practices in October 2016 and October 2017. The objectives were 1) to assess the levels of faecal contamination as a proxy for diarrhoeal disease risk; 2) to check whether the water still had sufficient levels of residual chlorine if ever chlorinated and whether the physico-chemical characteristics of the water would be suitable for point-of-use chlorination.

Samples collection

During both surveys, 150 ml samples were collected by trained interviewers in sterile bags Whirl-pak® (Nasco, Fort Atkinson, WI, US) from enrolled households. Instructions was given to sample water that would be used for drinking at the time of interview, and that it should be poured into the bag with the utensils (cup, jug, ladle) commonly used for serving. Sample bags were kept by interviewers in a cool bag without ice until collected for analysis at the CRH within 6 h of sampling.

Microbial quality assessment

Detection and/or quantification of Faecal Indicator Bacteria (FIB) instead of actual waterborne pathogens is a standard method for assessing the microbial safety of drinking water **[14]**. Amongst other characteristics, FIB are chosen to be representative of the risk of presence of waterborne pathogens and to react similarly to water treatment **[15]**. *E. coli*, and to a lesser degree thermotolerant coliforms (TTC, also referred to as faecal coliforms FC), are FIB of choice for drinking water quality surveillance, despite conflicting results of studies evaluating the relationship between FIB in drinking water and risk of diarrhoeal diseases **[16-18]**. Laboratory methods using FIB are based on the presence or number of growing FIB colonies on a specific culture medium at specific incubation temperature and duration. Samples to be tested can be directly inoculated onto the medium at various dilutions, with the Most Probable Number (MPN) procedure **[19]**. Extraction of micro-organisms from the sample can also be used: a known sample volume is filtered through a membrane with a specific pore size, and the membrane is inoculated onto a pad saturated with culture medium (membrane filtration procedure)**[20]**. Microbial analyses must be performed in aseptic conditions to avoid sample contamination and incubation of inoculated media should be uninterrupted for 18 to 24 h at a stable (+/- 0.5 °C) temperature.

Figure 2-13 Water analysis laboratory at CRH (left); steam sterilisation set-up (right)

For the first household survey in October 2016, and with little information on the human resources and equipment available at the CRH, it was decided to use a battery-powered incubator (Wagtech™ Potatest – Palintest, Tyne & Wear, UK)) dedicated to field testing of water microbial quality by membrane filtration, with lauryl sulphate medium incubated for 18 h at 44 °C for TTC detection and quantification. Lauryl sulphate medium was prepared from powder with distilled water and sterilised in a pressure cooker as per manufacturer instructions. Taking into consideration the maximum incubation capacity of 39 samples per day (+ one negative control plate) and expecting high levels of contamination, it was decided to filter a single volume of 50 ml of each sample, to maximise daily throughput while reducing the probability of being unable to count the number of colonies on incubated plates (referred to as “too numerous to count” or TNTC). To limit costs and waste, reusable aluminium petri dishes were sterilised by pressure cooker every day, and the membrane filtration unit was sterilized between each sample by means of formaldehyde produced by incomplete combustion of methanol, as per manufacturer instructions.

Multiple issues with the above protocol were noticed at the data analysis stage:

- Several containers of powdered lauryl sulphate were probably damaged during transport, possibly by too high or too low temperature. Although the prepared liquid did not turn red as instructed and in absence of extra supply or positive controls, it was used on several batches of analysis, and yellow/translucent colonies counted as TTC. We confirmed later that those results were invalid.

- Despite using a brand new Potatest incubating kit, several batches of 20 samples were lost to the unexplained failure of one of the incubators to function on battery instead of mains.

- Nearly 170 samples had a higher turbidity than expected for drinking water (>5 Nephelometric Turbidity Units or NTU; up to 77 NTU) but a lower rate of faecal contamination than expected considering they were mostly reported as originating from surface water sources. These unusual findings led us to review the literature in detail about possible elevated turbidity interference with membrane filtration detection methods. Although no clear guidelines were identified on necessary sample dilution before filtration at elevated turbidity, Lechevallier et al. reported in 1981 that coliforms detection by membrane filtration decreased as turbidity increased, as a result of coliform entrapment in particles and/or a higher number of bacteria antagonistic to coliforms and/or clogged membrane pores, all of which plausibly hindering coliform colony growth **[21]**.

The microbial results from the 2016 survey samples were therefore considered unreliable and the data discarded. The protocol used for the 2017 survey was modified to address these issues and was adjusted to laboratory conditions.

Figure 2-14 Aluminium petri dishes after incubation on lauryl sulphate medium: membrane clogging preventing bacterial growth (left); thermotolerant colonies are colored in yellow (right)

Major practical changes in 2017 included improvements in the equipment and materials used in the laboratory, to increase workflow and reliability. First, single-use sterile petri dishes were used in combination with prepared m-ColiBlue24® (Hach, Loveland, CO, US) 2 ml doses as culture medium, to avoid the lengthy steam sterilisation stages each day. m-ColiBlue24® medium is a nutritive, lactosebased medium, containing specific inhibitors that selectively eliminate the growth of noncoliform bacteria. Total coliform colonies become visible as they process a red dye present in the medium, while *E. coli* colonies become distinguishable with a blue dye they selectively produce. Membrane filters grown on m-ColiBlue24® require 24 hours of incubation at 35 °C. In addition, we purchased a robust field incubator (Ranger MX45 Mini, Lynd Products, Harrogate, UK) that we set up to operate only from two large vehicle batteries that we recharged and used in turn. The incubator's larger size made it more flexible in terms of number of samples processed each day, and the custom battery set-up gave us more control over power supply conditions.

We also changed the protocol to ensure high turbidity would not bias the results. Turbidity was measured before membrane filtering of samples, and the volume of sample filtered was adjusted with bottled water to reach approximately 3 NTU over 50 ml filtered. This meant a compromise on the method sensitivity, as we did not have the capacity to accordingly multiply the number of replicates for each sample and reach a standard amount of sample filtered to 50 or 100 ml.

Physico-chemical parameters

The main physico-chemical parameters of interest for household drinking water were related to the chlorination process. Water treatment with chlorine is a widely used and promoted method to remove microbiological contaminants, either prior to consumption (“point-of-use water treatment” or PoUWT) or prior to distribution in a water supply system.

When chlorine is added to water, part of it (called residual chlorine below, but also commonly referred to as free chlorine or free residual chlorine) is available to inactivate most pathogens causing diarrhoea in humans (viruses, bacteria and protozoa), and part of it is transformed by the organic materials and metals present in the water into compounds with no disinfection properties (combined chlorine). During the water treatment process, the quantity of residual chlorine and its disinfection efficiency depend largely on the type of water being treated – particularly turbidity, pH and temperature. These factors and the chlorine-resistance properties of the pathogens targeted will determine the chlorine concentration and contact time needed to achieve disinfection **[22]**. Efficiency of chlorine is optimal at turbidity <= 5 NTU an pH between 6.8 and 7.2. Deviations from recommended pH and turbidity in the water to be treated should be addressed with either pre-treatment or increasing chlorine concentration and contact time **[23]**.

Remaining residual chlorine after treatment will indicate that 1) there was initially enough chlorine to inactivate the pathogens present in the water; 2) pathogens introduced into the water after treatment can still be inactivated **[24]**. WHO recommends a residual chlorine concentration between 0.2 and 0.5 mg/l where users collect drinking water, to allow for protection against re-contamination during transport and storage time – during which residual chlorine will naturally disappear **[25]**.

During both household surveys, we therefore measured the following parameters from drinking water samples collected in households: pH, turbidity, residual and total chlorine concentrations. pH was measured with a digital multiparameter analyser (Hanna Instruments, Rhode Island, US) and turbidity with a digital turbidimeter (Eutech Instruments, Singapore). Residual and total chlorine concentrations were measured by DPD colorimetry, with a digital reader in the laboratory in 2016, and a visual reading with a pool tester in the households in 2017.

In 2016 and with little prior knowledge of the acceptability of drinking water sampling in households or the survey staff capacity, we chose to perform all parameter measurements at the CRH as soon as the samples of the day were delivered by the data collection team. The delay between sample collection and measurement however meant that chlorine concentrations, residual and total, would continue to decay, and that residual chlorine would continue inactivating the coliforms potentially present. In 2017, we chose to use sampling bags containing a sodium thiosulphate tablet, that neutralises residual chlorine present in the sample and interrupts its potential inactivation of pathogens present. It meant that chlorine levels had to be measured directly at the time of sampling by each interviewing team.

2.6.2 Environmental waters

During the household survey in 2016 and over 20 weeks of the summer of 2017, surface water in several locations across Uvira was analysed in an attempt to establish whether simplified laboratory methods would be suitable to explore time and/or space variations in surface water contamination by both *E. coli* and *V. cholerae* O_1. In November 2016, samples were submitted to the cholera detection method only, while between July and September 2017, samples were both analysed for microbial quality and cholera presence.

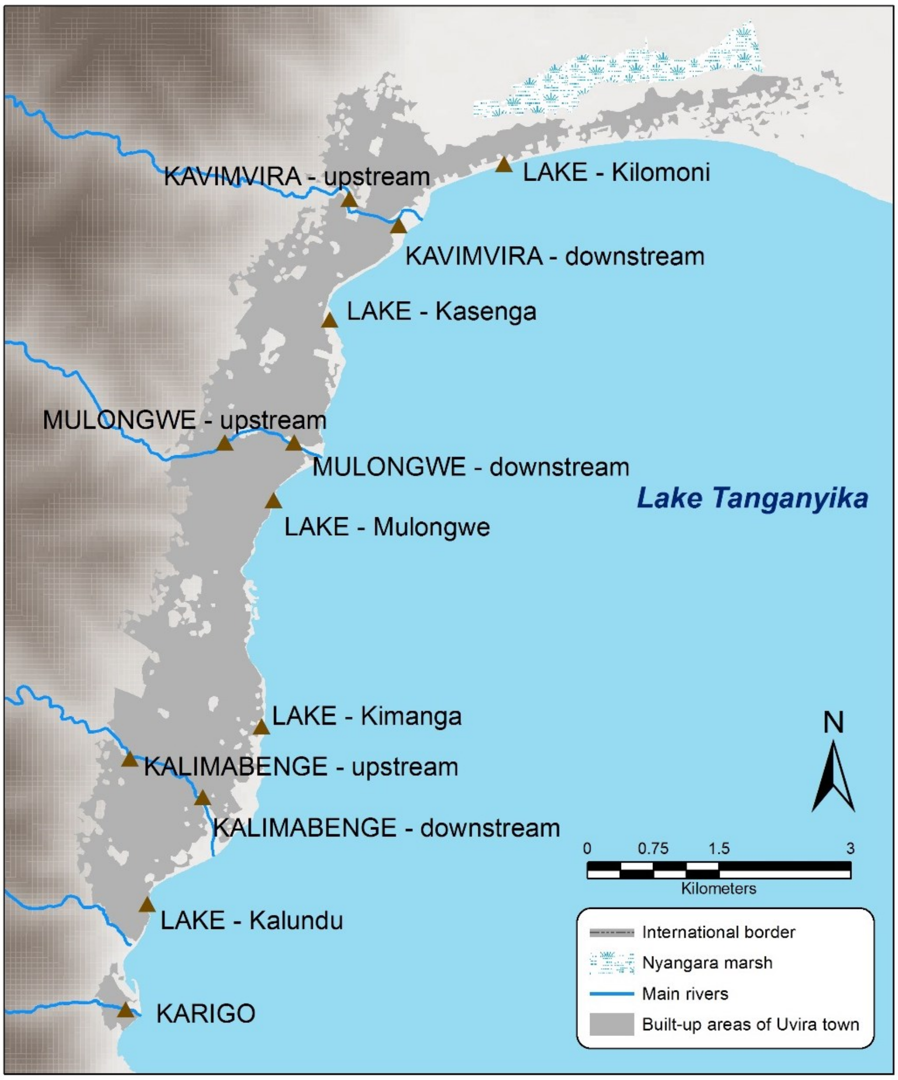


Figure 2-15 Map of environmental water sampling locations

Samples collection

Environmental water was sampled from 12 and 14 locations across town in 2016 and 2017 respectively. In 2016, sampling was repeated between 6 and 8 times over a 3-week period during the household survey implementation. In 2017, sampling was repeated

10 times between the 28th of June and the 8th of November. Locations were purposefully selected for accessibility and use by the population for domestic activities. They are mapped in **Figure 2-15**. Water was sampled at 1.5 m minimum from shore and 25 m away from human activities, between 7 am and 11 am. For lake locations, water was sampled approximately 25 cm below the surface.

2.2 liters of water were collected in a large sample Whirl-pak® and were brought back to the CRH laboratory within 3 h in large coolboxes without ice, protected from the sun-light. Physico-chemical measurements on-site were performed with a portable multi-parameter analyser, although data were discarded due to multiple calibration failures of several probes.

Microbial quality assessment

Turbidity and contamination levels were expected to be much higher than in drinking water samples. In order to avoid the need for additional sterilisation procedures, pre-prepared CompactDry™ EC (Nissui Pharmaceutical Co. Ltd, Ibaraki, Japan) were used. These plates are ready to receive 1 ml of sample before incubation for 24 hours at 37 °C. As for m-ColiBlue24® medium, red and blue colonies of fecal coliforms and *E. coli* respectively can then be directly observed and counted.

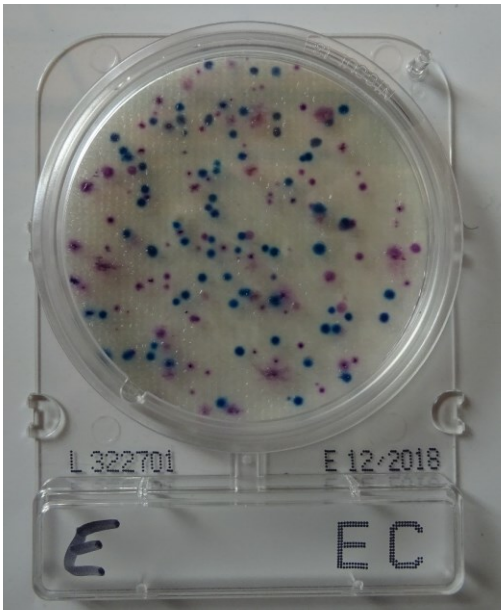



Figure 2-16 Left: Sample collection upstream of human activities on Kavimvira river by the CRH team (left to right: Clement, Charlotte, Vercus); Right: Compact Dry EC plate after incubation – red-pink dots are FC colonies and blue dots are *E. coli* colonies.

An initial assessment suggested that a minimum 1:5,000 dilution was necessary for river water samples to be able to count colonies in a 1 ml incubated sample, while a 1:60 dilution was sufficient for lake water samples. These dilutions were performed in sterile 50 ml and 10 ml tubes with sterile Pasteur pipettes and bottled water. The number of CompactDry™ EC plates available did not allow us to perform multiple dilutions for each sample.

V. cholerae detection

In 2016, we decided to explore whether *V. cholerae* O_1 presence could be detected in environmental water samples by a field method requiring few laboratory resources. Without the possibility of performing either microbiological cultures or PCR locally, we chose to use a method based on rapid diagnostic tests that had been used previously in Cameroon **[26]**.

Two liters of sampled water were slowly filtered on a 25 x 25 cm piece of sterile gauze folded and twisted into the neck of sterilised plastic funnels. The piece of gauze was then transferred into a 50 ml tube. A volume of 2X APW, equivalent to that of the soaked gauze (between 15 and 25 ml) was then added, before approximately 22 h of incubation at ambient temperature (~26 °C) with an unscrewed cap to allow for oxygen. The apical portion of the enriched sample was then tested for *V. cholerae* O_1 and/or O_{139} presence with a Crystal® VC RDT. To prevent environmental contamination with enriched cholera from our laboratory analyses, special attention was paid to boiling the APW samples after analysis before discarding them in the laboratory dedicated pit.

The absence of *V. cholerae* O_1/O_{139} in environmental samples tested by Debes et al. in Cameroon and later by Bwire et al. in Uganda, has prevented a full evaluation of the performance of this low-resource method; however, no false positives were detected in comparison with PCR results, even though *V. cholerae* non-O_1 non-O_{139} was identified in several samples by PCR **[26,27]**. A more robust method would have required multiple filtration of the sample on decreasing filter sizes, down to 0.22 µm under vacuum, instead of using gauze. The sensitivity of this method followed by APW enrichment and Crystal® VC testing has been reported to have an 86 % sensitivity and 100 % specificity against culture **[28]**.

Beyond the limitations related to the concentration/extraction of potential *V. cholerae* from the samples, the APW enrichment step, necessary to reach the detection limit of RDTs, is not appropriate to promote the growth of viable but non-culturable (VBNC) *V.cholerae*, a form frequently encountered in the environment. Further, using RDTs for *V. cholerae* O_1/O_{139} detection prevents distinguishing between toxigenic and non-toxigenic strains of the O_1/O_{139} serogroups. In absence of further laboratory testing involving cultures and/or PCR, results of this method must be interpreted with caution in terms of public health significance.

Figure 2-17 Set-up for sample filtration through gauze and APW tube with gauze (left)

2.7 Research framework and ethical considerations

The present research was part of a larger collaboration initiated in 2012 between the Veolia Foundation (VF), the French Development Agency (AFD), OXFAM, the Ministry of Health of the Democratic Republic of the Congo, Regideso and the LSHTM in parallel with the plans to improve tap water supply in Uvira. This collaboration is broadly aimed at evaluating the impact of the tap water infrastructure improvements in Uvira on water-related practices, diarrhoeal diseases and cholera. It was formalised by a Memorandum of Understanding (MoU) between all parties agreed in 2015, and a steering committee, composed of a member of each above institution and an additional representant of the University of Kinshasa – School of Public Health (UNIKIN ESP), was established at the same time to monitor research activities.

A favourable review from the LSHTM ethics committee (# 8913 /RR/3514) was obtained in January 2015 by the principal investigator Dr Jeroen Ensink for the overall research activities plan, conditional on approval from the ethics committee of the UNIKIN ESP, that was received in September 2015 (# ESP/CE/088/2015).

Two subsequent ethics applications received a favourable review by both LSHTM and UNIKIN ESP at a later stage: 1) to conduct a confirmation of cholera by rapid diagnostic tests amongst patients admitted to the CTC (LSHTM # 10603 /RR/4020; UNIKIN ESP # ESP/CE/088b/2016) ; 2) to preserve and analyse by PCR at the LSHTM a selection of rectal swabs sampled from CTC patients (LSHTM # 15193 /RR/10475 and UNIKIN ESP # ESP/CE/088c/2017).

In accordance with research ethics principles, all individual data collected at the CTC were anonymised. Household interviews were conducted after an informed and witnessed consent process that included written consent given by the main respondent willing to participate. For cholera confirmation, rectal swabs were only collected after a witnessed written consent was given by willing participants or legal guardians of children younger than 15 years. Informed consent forms used are available in **Appendix 2-6** and **Appendix 2-7**.

2.8 Chapter 2 - References

1. Location of Province South-Kivu in the Democratic Republic of the Congo. *Wikipedia* 2016; https://upload.wikimedia.org/wikipedia/commons/f/f7/Democratic_Republic_of_the_Congo_%2826_provinces%29_-_Sud-Kivu.svg.
2. Google. Google Maps. https://www.google.com/maps/@-2.7399972,28.9626296,9z. Accessed 19/05/2020.
3. Climate Change Knowledge Portal - DRC profile. The World Bank Group. https://climateknowledgeportal.worldbank.org/country/congo-democratic-republic/climate-data-historical. Accessed 11/05/2020.
4. L'Essentiel RDC. Urgent ! Situation énergétique alarmante dans la province du Sud-Kivu, alerte la Snel. 2019; https://lessentielrdc.info/urgent-situation-energetique-alarmante-dans-la-province-du-sud-kivu-alerte-la-snel/. Accessed 20/05/2020.
5. International Monetary Fund. *Democratic Republic of the Congo.* Washington DC,2019. 19/286.
6. United Nations Development Program. *Human Development Report 2019.* 2019.
7. Office of Coordination For Humanitarian Affairs. *Plan de réponse humanitaire RDC - mise à jour 2019.* 2019.
8. Verweijen J. *A Microcosm of Militarization: conflict, governance and armed mobilization in Uvira, South Kivu.* 2016.
9. Bhuiyan NA, Qadri F, Faruque ASG, et al. *Use of Dipsticks for Rapid Diagnosis of Cholera Caused by Vibrio cholerae O1 and O139 from Rectal Swabs.* **Journal of Clinical Microbiology**. 2003;41(8):3939-3941.
10. Page A-L, Alberti KP, Mondonge V, Rauzier J, Quilici M-L, Guerin PJ. *Evaluation of a Rapid Test for the Diagnosis of Cholera in the Absence of a Gold Standard.* **PLOS ONE**. 2012;7(5):e37360.
11. Center for Disease Control. Crystal® VC Rapid Diagnostic Test (RDT) Procedure. 2018; https://www.cdc.gov/cholera/crystal-vc.html. Accessed 20/05/2020.
12. Global Task Force For Cholera Control - Surveillance WG. *Interim Guidance Document on Cholera Surveillance.* 2017.

13. Collectif. *L'eau et l'aménagement dans l'Afrique des Grands lacs.* Bordeaux 1993.

14. Organization WH. *Guidelines for drinking-water quality: first addendum to the fourth edition.* 2017.

15. Payment P, Waite M, Dufour A. *Introducing parameters for the assessment of drinking water quality*. **Assessing microbial safety of drinking water**. 2003;4:47-77.

16. Gruber JS, Ercumen A, Colford JM, Jr. *Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis*. **PloS one**. 2014;9(9):e107429-e107429.

17. Luby SP, Halder AK, Huda TM, et al. *Microbiological Contamination of Drinking Water Associated with Subsequent Child Diarrhea*. **Am J Trop Med Hyg**. 2015;93(5):904-911.

18. Wright J, Gundry S, Conroy R. *Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use*. **Trop Med Int Health**. 2004;9(1):106-117.

19. Cochran WG. *Estimation of Bacterial Densities by Means of the "Most Probable Number"*. **Biometrics**. 1950;6(2):105-116.

20. Association APH, Association AWW, Federation WE. *Standard methods for the examination of water and wastewater.* American Public Health Association; 2017.

21. LeChevallier MW, Evans TM, Seidler RJ. *Effect of turbidity on chlorination efficiency and bacterial persistence in drinking water*. **Applied and environmental microbiology**. 1981;42(1):159-167.

22. LeChevallier MW, Au K-K. *Water treatment and pathogen control.* Iwa Publishing; 2004.

23. Lantagne DS, Blount BC, Cardinali F, Quick R. *Disinfection by-product formation and mitigation strategies in point-of-use chlorination of turbid and non-turbid waters in western Kenya*. **J Water Health**. 2008;6(1):67-82.

24. Robert R, Sam G, Sam K, et al. *Technical notes on drinking-water, sanitation and hygiene in emergencies.* 2014.

25. World Health O. Guidelines for drinking-water quality. 4th ed ed. Geneva: World Health Organization; 2011.

26. Debes AK, Ateudjieu J, Guenou E, et al. *Clinical and Environmental Surveillance for Vibrio cholerae in Resource Constrained Areas: Application During a 1-Year Surveillance in the Far North Region of Cameroon*. **Am J Trop Med Hyg**. 2016;94(3):537-543.

27. Bwire G, Debes AK, Orach CG, et al. *Environmental Surveillance of Vibrio cholerae O1/O139 in the Five African Great Lakes and Other Major Surface Water Sources in Uganda*. **Front Microbiol**. 2018;9:1560.

28. Chakraborty S, Alam M, Scobie HM, Sack DA. *Adaptation of a simple dipstick test for detection of Vibrio cholerae O1 and O139 in environmental water*. **Front Microbiol**. 2013;4:320.

3

Cholera and other enteric infections amongst patients admitted to the cholera treatment centre

Located within the Uvira district hospital, the Cholera Treatment Centre (CTC) provides free treatment for patients presenting with acute diarrhoea, defined as three or more loose stools within 24 hours. All patients admitted to the CTC are considered as suspected cholera cases and reported as such to the national health information system. In an area where cholera is considered endemic but where diarrhoeal diseases of other aetiologies are likely to be highly prevalent as well, the absence of other clinical criteria for admission raises the issues of the actual incidence rate of cholera amongst CTC patients and of the other enteropathogens responsible for non-cholera diarrhoeal episodes.

This chapter presents the results of an ongoing cholera confirmation study started in 2016 at the Uvira CTC, first with an article published in the journal PLOS ONE in August 2018 based on data to November 2017, followed by the analysis of the same dataset extended to April 2019. It also reports a summary of the results of a side study that used molecular methods to detect the presence of a range of enteropathogens in preserved stools collected from CTC patients.

3.1 Research paper 1

Title: Confirmation of cholera by rapid diagnostic test among patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo

Authors: Aurelie Jeandron[1], Oliver Cumming[1], Baron Bashige Rumedeka[2], Jaime Mufitini Saidi[2], Simon Cousens[3]

[1] Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[2] Ministère de la Santé Publique, Division Provinciale de la Santé Publique, District Sanitaire d'Uvira, Uvira, Sud-Kivu, République Démocratique du Congo

[3] Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

Status: Published in PLoS One in August 2018

Citation: Jeandron A, Cumming O, Rumedeka BB, Saidi JM, Cousens S (2018) Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo. PLoS ONE 13(8): e0201306.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	173940	**Title**	Ms
First Name(s)	Aurélie		
Surname/Family Name	Jeandron		
Thesis Title	Tap water access and its relationship with cholera and other diarrhoeal diseases in an urban, cholera-endemic setting in the Democratic Republic of the Congo		
Primary Supervisor	Prof. Simon Cousens		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	PLOS ONE		
When was the work published?	AUGUST 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Conceptualization; Data curation; Formal analysis; Investigation; Writing - original draft; Writing - review & editing

SECTION E

Student Signature	Aurelie Jeandron
Date	21/05/2020

Supervisor Signature	Simon Cousens
Date	22/05/2020

PLOS ONE

RESEARCH ARTICLE

Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo

Aurelie Jeandron[1]*, Oliver Cumming[1], Baron Bashige Rumedeka[2], Jaime Mufitini Saidi[2], Simon Cousens[3]

1 Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **2** Ministère de la Santé Publique, Division Provinciale de la Santé Publique, District Sanitaire d'Uvira, Uvira, Sud-Kivu, République Démocratique du Congo, **3** Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

* aurelie.jeandron@lshtm.ac.uk

OPEN ACCESS

Citation: Jeandron A, Cumming O, Rumedeka BB, Saidi JM, Cousens S (2018) Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo. PLoS ONE 13 (8): e0201306. https://doi.org/10.1371/journal.pone.0201306

Editor: Andrew S. Azman, Johns Hopkins Bloomberg School of Public Health, UNITED STATES

Received: February 1, 2018

Accepted: July 11, 2018

Published: August 1, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was funded by the Veolia Foundation (www.fondation.veolia.com; grant No 13/14 HD1123) and the French Development Agency (AFD www.afd.fr; grant No ERS-EVA 364-2015). The funders had no role in study design,

Abstract

Introduction

Cholera is endemic in the Eastern provinces of the Democratic Republic of the Congo since 1978, and Uvira in South-Kivu has been reporting suspected cholera cases nearly every week for over a decade. The clinical case definition for suspected cholera is relatively non-specific, and cases are rarely confirmed by laboratory methods, especially in endemic settings. This may lead to over-estimation of cholera cases and limit effective public health responses.

Methods and results

Between April 2016 and November 2017, 69% of the 2,059 patients admitted to the Uvira Cholera Treatment Centre (CTC) were tested for cholera with rapid diagnostic tests (RDTs). Of those admitted as suspected cholera cases, only 40% tested positive for cholera, equivalent to an estimated annual incidence of suspected/confirmed cholera in Uvira of 43.8 and 16.3 cases per 10,000 inhabitants respectively. A multivariable logistic regression indicates that boys aged 2 to 4 years, girls aged 5 to 15 years and adult men are respectively 1.9, 2.1 and 1.8 times more likely to test positive than adult women. On the contrary, boys under 2 are 10 times less likely to test positive. The odds of testing positive also increase as weekly admissions to the CTC rise, with up to a 5-fold increase observed during the weeks with the highest numbers of admissions compared to the lowest ones. Other predictors of cholera confirmation include duration of stay at the CTC, clinical outcome of admission, lower weekly rainfall and area of residence in Uvira, with the northern part of town having the highest confirmation rate.

Conclusion

Cholera is an on-going public health problem in Uvira but the majority of suspected cases admitted to the CTC were found to be negative for cholera after RDT testing. These findings

data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

may have important implications for cholera control strategies in favour of interventions that address cholera and other diarrhoeal diseases alike.

Introduction

In 2016, more than 130,000 cases of cholera were reported to the World Health Organisation (WHO) globally [1]. These figures are likely incomplete due to weak national surveillance systems and the actual global number of cholera cases is estimated to be closer to 3 million every year [2]. Comprehensive surveillance is also hampered by the relatively non-specific WHO standard case definition of suspected cholera cases, especially in endemic areas [3]. Enhanced surveillance data on the burden and characteristics of confirmed cholera across 11 study sites in seven countries of Sub-Saharan Africa by the AFRICHOL consortium show large variations in incidence and clinical presentation, with the proportion of confirmed cases amongst suspected cases ranging from 10% to 60% and a specificity of the WHO standard case definition of 8% [4, 5]. In Haiti, the specificity of this case definition in an outbreak setting was estimated to be only 43% over a two-year study in 4 hospitals [6]. Over-reporting of suspected cholera cases is therefore possible in areas with a long history of cholera outbreaks, when laboratory confirmation is only done on initial outbreak cases.

On the African continent, reported cholera incidence is generally higher inland than in coastal areas and, in the past two decades, outbreaks have been mostly clustered around the Great Lakes and Lake Chad regions [7]. Globally, increased cholera incidence and outbreaks tend to coincide with increased rainfall, although other climatic factors also appear to play a role, with seasonal patterns more pronounced further from the Equator [8].

With 28,093 suspected cases reported, the Democratic Republic of the Congo (DRC) was, in absolute numbers, the second most affected country after Haiti in 2016 [1, 9]. Provinces in the east of the country, particularly South-Kivu province, have been identified as a stable cholera transmission focus since 1978 [9, 10].

Our study describes the epidemiology of suspected and confirmed cholera based on admissions and results of rapid diagnostic testing collected over 18 months at the Cholera Treatment Centre (CTC) of Uvira in eastern DRC.

Methods

Study area

The city of Uvira is located in South Kivu Province in eastern DRC on the shores of Lake Tanganyika and had an estimated population of 233,000 inhabitants in June 2016 (based on municipal authority figures). Uvira is divided into 16 neighbourhoods ("aires de santé" or AS), with estimated populations ranging from 6,500 to 24,000 people. These neighbourhoods can be grouped into three areas: North, Centre and South (Fig 1). The CTC at Uvira District Hospital admits all patients with acute diarrhoea, defined as 3 or more loose or liquid stools in 24h, presenting directly to the hospital or referred by local health posts. Patients are treated for dehydration and occasionally administered broad-spectrum antibiotics, zinc or albendazole. Admission and treatment at the CTC is free of charge for all patients. Patients admitted to the CTC may be referred to other hospital departments if they also present complicating medical conditions requiring additional care, such as diabetes, malaria or pregnancy.

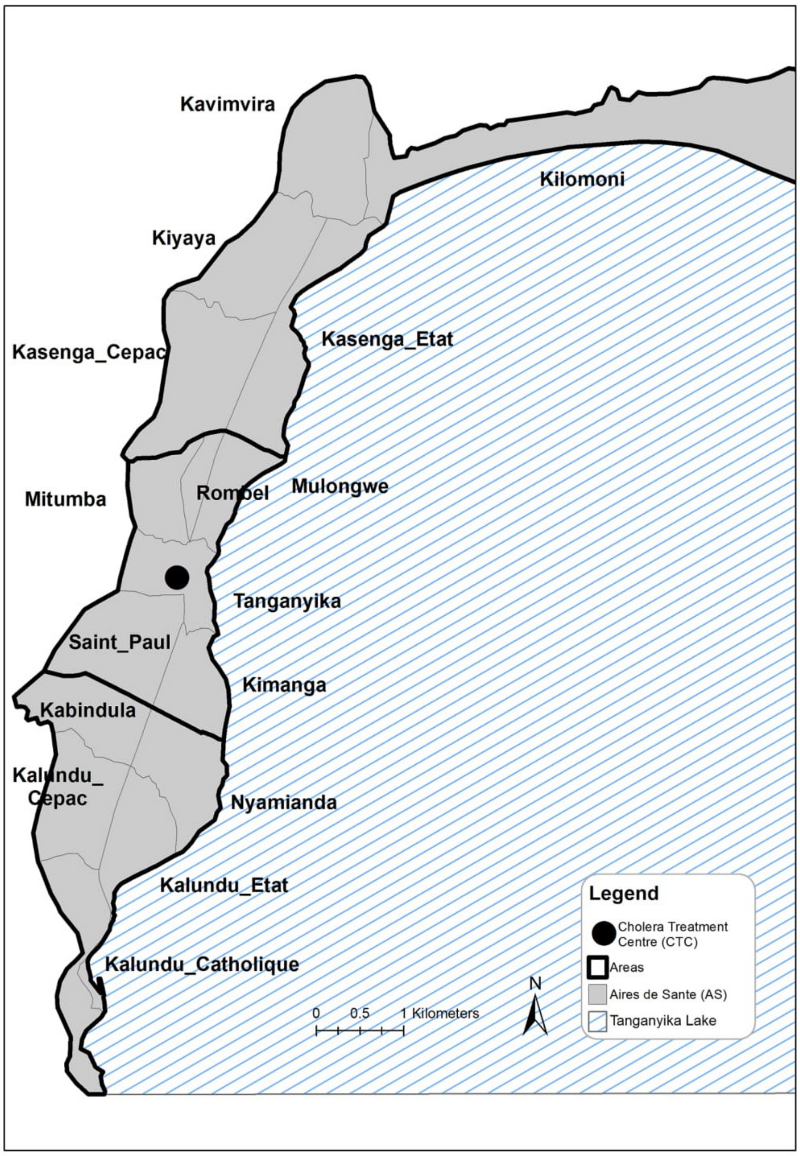


Fig 1. Map of Uvira.

https://doi.org/10.1371/journal.pone.0201306.g001

At least one admission has been reported by the Uvira CTC to the District Health Office for every week except two between 2004 and 2017. All CTC admissions are considered as suspected cholera cases in Uvira and reported as such to the national level.

Data and sample collection and laboratory analysis

From April 2016, a laboratory technician at Uvira District Hospital was trained to collect a rectal swab from all consenting newly admitted patients present at the CTC during daily morning visits 7 days per week. Demographic and clinical information was collected from the patient or a family member and the CTC patient register. The number of hours between first administration of antibiotic treatment (single-dose 300mg doxycycline) and sampling was also recorded.

The rectal swab was enriched in 4 ml of Alkaline Peptone Water (APW) for approximately six hours at ambient temperature (approximately 25˚C +/- 2˚C), and the apical portion of the unshaken APW vial was then tested for cholera with a rapid diagnostic test (RDT) (Crystal VC, Span Diagnostics, Surat, India), according to the manufacturer's instructions.

Although stool samples are recommended by the manufacturer, rectal swabs were preferred to stool samples to reduce the risk of sample contamination, either across patients or with disinfectant, and to limit stool handling by clinical staff. The use of Crystal VC RDTs on enriched rectal swabs was however evaluated to have a sensitivity and specificity of 92% and 91% respectively in a previous study [11]. The enrichment step in APW has been shown to improve significantly the specificity of VC Crystal RDTs estimated at 71% when used directly on stool samples [12].

Data collection

Data was collected electronically using Open Data Kit (ODK) on an Android tablet and stored remotely on a dedicated ODK server at the London School of Hygiene and Tropical Medicine [13]. Complementary information post-confirmation, such as discharge date and outcome were added later from the CTC patient register. Population estimates for each AS were obtained from Uvira Town Hall in June 2016 and, in the absence of estimates at other time-points, populations were assumed to be stable over the study period. Average daily rainfall over the Uvira area was estimated using data from the Goddard Earth Sciences Data and Information Services Center (GES DISC) [14].Weekly rainfall was then calculated and each week categorized as dry or rainy with the median weekly rainfall of 3.6mm used as the cut-off. Weeks classified as "rainy" coincided mostly with the definition of rainy season for the area, between the months of October and May with an interruption in January [15].

Data analysis

Data was analysed with STATA 14 (StataCorp, College Station, TX) and R [16].

Case definitions

In this study, all patients admitted to the Uvira CTC were considered as suspected cholera cases, as per the definition recommended by the Global Task Force for Cholera Control (GTFCC) once an outbreak is declared [17]. They are referred to as suspected cases or admitted patients interchangeably. The same set of recommendations for cholera surveillance defines confirmed cholera cases as any suspected case from which Vibrio cholerae O1 or O139 is isolated. As no isolation of cholera strains were performed during the present study, we refer to participants tested positive with cholera RDTs as RDT-positive cases.

Characteristics of RDT-positive cholera cases

Univariable logistic regression was used to assess whether age, sex, antibiotics administration before sampling, weekly number of admissions, area of residence, weekly rainfall category and clinical outcome (duration of stay and outcome of admission) were associated with the probability of a patient being enrolled or RDT positive. Predictors that were associated ($p<0.1$) in univariable models were included in a multivariable model and 2-way interactions were assessed by comparing likelihood ratios of models with or without interaction term. The final multivariable model was used to predict the proportion of untested patients that would have been confirmed as cholera cases by RDT. In order to investigate geographical variation in cholera confirmation rates at a smaller scale, AS of residence was introduced in the final model as a random effect to avoid coefficient estimate bias attributable to a high number of parameters [18]. Associations were reported as crude odds ratios (OR) for the univariable models, or adjusted odds ratios (aOR) for the multivariable models, along with 95% confidence intervals (CI).

Time and space distribution of suspected and RDT-positive cholera cases

Weekly admissions per area and per AS were calculated over the entire study period and plotted. The number and proportion of weeks with suspected and RDT-positive cases was used as an indicator of the endemicity of suspected and RDT-positive cholera. Annual numbers of CTC admissions per area or AS were estimated as the mean number of admissions over 52 consecutive weeks (32 overlapping periods covered by the study), and annual number of RDT-positive cholera cases were derived by applying the confirmed/admission ratios estimated from the multivariable model for each area or AS.

Funding and ethical considerations

This study was funded by the Agence Française de Développement and the Veolia Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Demographic and clinical data for each patient is routinely collected in the CTC patient registry. Patients were only enrolled for RDT confirmation once a written informed consent was obtained, from either the patient or from the legal guardian for patients aged under 15. If illiterate, an oral consent was obtained and witnessed by an impartial third party. All data was anonymized before collection and analysis. In order to ensure that clinical staff deliver the same standard of care for all patients, the results of individual RDT results were not communicated to the clinical team.

The study was approved by the ethics committees of the School of Public Health at the University of Kinshasa, Democratic Republic of the Congo (ESP/CE/088c/2017), and of the London School of Hygiene and Tropical Medicine, United Kingdom (No 10603).

Results

Characteristics of admitted patients

Between the 4th of April 2016 and the 5th of November 2017 (83 weeks), a total of 2,059 patients were admitted to the CTC. A majority (54.6%) of patients admitted to the CTC were 16 or older, while 336 (16.3%) were under 5 years of age. Similar numbers of female and male patients were admitted (1,019 and 1,038 respectively). Weekly admissions to the CTC ranged from 3 to 103 patients per week, with a median of 15 admissions per week (Fig 2). The 83 weeks of the study period were stratified into 4 categories of admission incidence–low, moderate, high and very high—defined by arbitrary cut-offs at 10, 20 and 50 admissions per week. These categories represented respectively 156 (7.6%), 302 (14.7%), 571 (27.7%) and 1,030 (50%) admissions over respectively 22 (26.5%), 22 (26.5%), 19 (22.9%) and 15 (18.1%) weeks. Date of exit and clinical outcome were missing for 9.3% and 7% of the admitted patients respectively. A majority (52.7%) of the patients admitted to the CTC stayed one or two nights, while 17.2% did not stay overnight and 30.1% stayed for 3 nights or more. The vast majority (87.1%) of the patients were discharged from the CTC, but 5.4% of the patients were

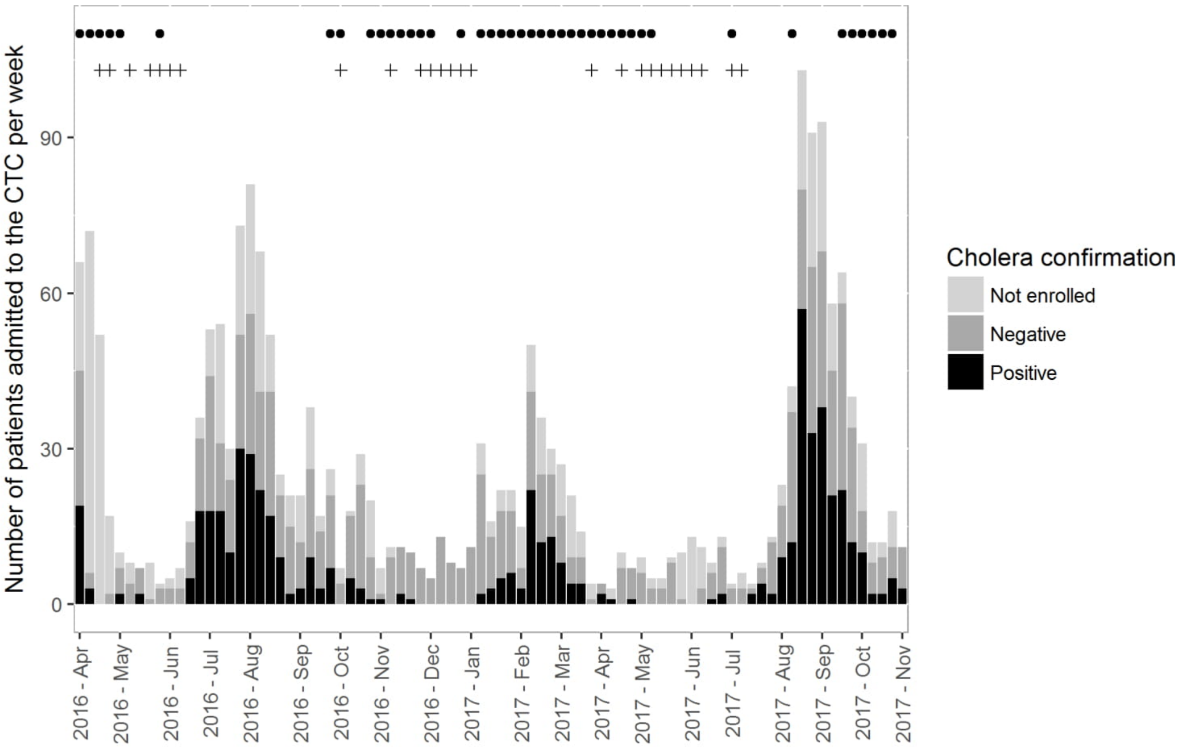


Fig 2. Weekly number of patients admitted to the CTC during the study period. The dots represent weeks classified as "rainy" (weekly rainfall estimated to be above 3.6 mm). The plus signs represent weeks with no cholera confirmed cases.

https://doi.org/10.1371/journal.pone.0201306.g002

transferred to other hospital departments. 10 deaths were recorded at the CTC during the study period, representing a case fatality ratio (CFR) of 0.5%.

Enrolment and characteristics of enrolled patients

Rectal swabs were collected from 1,419 patients (68.9% of all those admitted). Of the 640 patients not enrolled, 39.5% refused to participate and 60.5% were not recruited due to their short stay at the CTC or the absence of the laboratory technician. The distribution of age and sex was similar among those recruited and not recruited. Patients admitted during weeks of very high admission incidence were slightly less likely to be enrolled in the study than patients admitted during other weeks. Similarly, patients leaving on the same day as admission were less likely to be enrolled, as were transferred or deceased patients. Characteristics of admitted and enrolled patients are summarized in Table 1.

Cholera confirmation

562 out of 1,419 rectal swabs (39.6%) were positive for cholera O1 after enrichment. None of the patients enrolled reported having taken antibiotics between symptom onset and CTC admission, and only 30 patients had been administered a single dose of doxycycline within the 12 hours preceding rectal sampling. Although *Vibrio cholerae* shedding was unlikely to have

Table 1. Characteristics of patients admitted to the CTC and enrolled into the confirmation study.

	Admitted	Enrolled
	n (% of total)	n (% enrolled)
Total	2059 (100%)	1419 (68.9%)
Sex		
Male	1038 (50.4%)	713 (68.7%)
Female	1019 (49.5%)	705 (69.2%)
Missing	2 (0.1%)	1 (50%)
Age group		
Under 2 year old	62 (3%)	37 (59.7%)
2 to 4 year old	274 (13.3%)	182 (66.4%)
5 to 15 year old	594 (28.8%)	412 (69.4%)
16 and older	1124 (54.6%)	787 (70%)
Missing	5 (0.2%)	1 (20%)
Weekly admissions incidence		
Low (3 to 9 admissions)	156 (7.6%)	109 (69.9%)
Moderate (10 to 19 admissions)	302 (14.7%)	208 (68.9%)
High (20 to 49 admissions)	571 (27.7%)	428 (75%)
Very High (50 to 103 admissions)	1030 (50%)	674 (65.4%)
Duration of stay		
Less than 24h	160 (7.8%)	84 (52.5%)
1 to 2 nights	1086 (52.7%)	771 (71%)
3 to 4 nights	501 (24.3%)	357 (71.3%)
5 nights or more	120 (5.8%)	81 (67.5%)
Missing	192 (9.3%)	126 (65.6%)
Outcome		
Discharged	1793 (87.1%)	1267 (70.7%)
Transferred	111 (5.4%)	47 (42.3%)
Death	10 (0.5%)	3 (30%)
Missing	145 (7%)	102 (70.3%)
Area of residence		
North (est. population 75,046)	891 (43.3%)	649 (72.8%)
Centre (est. population 98,384)	474 (23%)	316 (66.7%)
South (est. population 59,733)	611 (29.7%)	395 (64.6%)
Outside Uvira or missing	83 (4%)	59 (71.1%)
Weekly rain category		
Dry ($< = 3.6$ mm)	1178 (57.2%)	857 (72.8%)
Rainy (>3.6mm)	881 (42.8%)	562 (63.8%)

https://doi.org/10.1371/journal.pone.0201306.t001

been affected in such a short time frame, antibiotics administration before sampling was included as a potential predictor in the analysis [19]. The proportions of positive samples by age, sex, weekly incidence category, antibiotics administration, clinical outcome, area of residence and weekly rainfall category are shown in Table 2, along with crude and adjusted odds ratios from the univariable and multivariable models.

Although sex did not appear to be associated independently with cholera confirmation ($p = 0.49$), there was evidence for interaction between sex and age group ($p<0.001$). Girls aged 5 to 15 (OR 2.1; 95% CI: 1.5–3), boys aged 2 to 4 (OR 1.9; 95% CI: 1.2–3) and men aged 16 or older (OR 1.5; 95% CI: 1.1–2) were more likely to test positive than women aged 16 or older. Boys under 2 years of age were much less likely to test positive than women aged 16 or older

Table 2. Predictors for cholera confirmation amongst patients admitted to the CTC.

	Enrolled (N = 1'419)	Positive	Univariable logistic regression		Multivariable logistic regression	
	n	n (% positive)	Crude OR (95% CI)	*p-value**	Adjusted OR (95% CI)	*p-value**
Total	1419	562 (39.6%)				
Sex						
Male	713	289 (40.5%)	1.08 (0.87–1.33)	*0.49*		
Female	705	273 (38.7%)	reference			
Missing	1	0 (0%)				
Age group						
Under 2 year-old	37	7 (18.9%)	0.38 (0.17–0.88)	*0.007*		
2 to 4 year old	182	81 (44.5%)	1.32 (0.95–1.82)			
5 to 15 year old	412	176 (42.7%)	1.23 (0.96–1.57)			
16 and older	787	298 (37.9%)	reference			
Missing	1	0 (0%)				
Demographic group						
Girls—under 2	16	6 (37.5%)	1.22 (0.44–3.44)	*p<0.001*	1.45 (0.49–4.3)	*p<0.001*
Girls—2 to 4 years	87	35 (40.2%)	1.37 (0.85–2.21)		1.21 (0.68–2.14)	
Girls—5 to 15 years	197	99 (50.3%)	2.08 (1.47–2.95)		2.12 (1.42–3.15)	
Women—16 and older	404	133 (32.9%)	reference		reference	
Boys—under 2	21	1 (4.8%)	0.1 (0.01–0.77)		0.12 (0.01–0.93)	
Boys—2 to 4 years	95	46 (48.4%)	1.91 (1.22–3.01)		1.85 (1.11–3.08)	
Boys—5 to 15 years	214	77 (36%)	1.15 (0.81–1.62)		1.04 (0.7–1.56)	
Men—15 and older	383	165 (43.1%)	1.54 (1.15–2.06)		1.78 (1.28–2.49)	
Missing	1	0 (0%)	excluded			
Antibiotics treatment before sampling						
No	1389	20 (1.4%)	0.76 (0.35–1.63)	*0.47*		
Yes	30	10 (33.3%)	reference			
Weekly admissions incidence						
Low (3 to 9 admissions)	109	14 (12.8%)	reference	*p<0.001*	reference	*p<0.001*
Moderate (10 to 19 admissions)	208	44 (21.2%)	1.82 (0.95–3.5)		1.83 (0.87–3.83)	
High (20 to 49 admissions)	428	155 (36.2%)	3.85 (2.13–6.98)		3.35 (1.69–6.66)	
Very High (50 to 103 admissions)	674	349 (51.8%)	7.29 (4.08–13.03)		5.33 (2.71–10.5)	
Duration of stay						
Less than 24h	84	26 (31%)	reference	*p<0.001*	reference	*p<0.001*
1 to 2 nights	771	227 (29.4%)	0.93 (0.57–1.52)		0.87 (0.5–1.53)	
3 to 4 nights	357	204 (57.1%)	2.97 (1.79–4.94)		2.44 (1.36–4.4)	
5 nights or more	81	53 (65.4%)	4.22 (2.2–8.1)		3.5 (1.67–7.34)	
Missing	126	52 (41.3%)	excluded			
Outcome						
Discharged	1267	493 (38.9%)	reference	*0.05*	reference	*0.02*
Transferred	47	25 (53.2%)	1.78 (0.99–3.2)		2.33 (1.14–4.77)	
Death	3	3 (100%)	excluded		excluded	
Missing	102	41 (40.2%)				
Area of residence						
North (est. population 75,046)	649	275 (42.4%)	1.5 (1.13–1.99)	*0.035*	1.57 (1.14–2.17)	*0.05*
Centre (est. population 98,384)	316	104 (32.9%)	reference		reference	
South (est. population 59,733)	395	157 (39.7%)	1.35 (0.99–1.84)		1.3 (0.91–1.86)	
Outside Uvira or missing	59	26 (44.1%)	1.61 (0.91–2.83)		1.57 (0.82–3.01)	
Weekly rain category						

(*Continued*)

Table 2. (Continued)

	Enrolled (N = 1'419)	Positive	Univariable logistic regression		Multivariable logistic regression	
	n	n (% positive)	Crude OR (95% CI)	*p-value**	Adjusted OR (95% CI)	*p-value**
Dry (< = 3.6 mm)	857	375 (43.8%)	1.56 (1.25–1.95)	*p*<0.001	1.34 (1.01–1.78)	*0.04*
Rainy (> 3.6 mm)	562	187 (33.3%)	reference		reference	

* χ2 significance test of likelihood ratio (LR test)

https://doi.org/10.1371/journal.pone.0201306.t002

(OR 0.1; 95% CI: 0.01–0.8). The proportion of positive samples was higher in weeks with higher numbers of admissions to the CTC, with the odds of a patient testing positive being more than 7-fold higher in very high incidence weeks compared to low incidence weeks. Similarly, the proportion of positive samples increased with the patient's duration of stay. Patients staying 5 nights or more had more than 4 times the odds of testing positive compared to those leaving on the day of admission. All three patients tested who later died were confirmed cholera cases. There was weak evidence in univariable logistic regression analyses that transferred patients were more likely to be positive than discharged patients. Finally, patients admitted during dry weeks had higher odds of testing positive to cholera than those admitted during rainy weeks.

In a multivariable model including demographic group (combination of age and sex), admissions incidence category, duration of stay and outcome, area of residence and weekly rainfall category, girls aged 5 to 15, boys between 2 and 4 years of age and men aged 16 or older had respectively 2.1 (95% CI 1.4–3.2), 1.9 (95% CI 1.1–3.1) and 1.8 (95% CI 1.3–2.5) higher odds of testing positive compared to women aged 16 or older. Boys under 2 had 10 times lower odds of testing positive (aOR 0.1; 95% CI: 0.01–0.9). Patients admitted during weeks of high and very high admissions incidence had 3.4 (95% CI 1.7–6.7) and 5.3 (95% CI 2.7–10.5) times the odds respectively of testing positive compared to those admitted during low admissions weeks. Similarly, patients staying 3 or 4 nights or more than 5 nights had 2.4 (95% CI 1.4–4.4) and 3.5 (95% CI 1.7–7.3) times higher odds of being confirmed respectively than patients leaving on the day of admission. Patients transferred to another department were also more likely to be confirmed than discharged patients (aOR = 2.3; 95% CI 1.1–4.7). Patients residing in the North area of town were also more likely to be confirmed as cases than patients living in other areas of town (aOR = 1.6; 95%CI 1.1–2.2 compared to patients living in the centre), while patients admitted during dry weeks were more likely to be confirmed than during other weeks (aOR = 1.4; 95% CI 1–1.8).

There was no evidence of any 2-way interactions between demographic group, admissions category, clinical outcomes, area of residence and weekly rainfall. There was no evidence of a variation in confirmation rates amongst CTC admitted patients across AS after accounting for area (Likelihood ratio test of multivariable logistic model with AS as random effect vs multivariable model p = 1).

We used the multivariable model to predict the probability that non-enrolled patients would have tested positive. An estimated 39.6% of the 563 non-enrolled patients with complete information (age and sex group, duration of stay and outcome) are predicted to have been "true" cholera cases, very similar to the proportion of positives among those enrolled. The proportion of "true" cholera cases amongst all admissions (excluding 77 with incomplete information) is thus estimated to be 39.5% over the study period.

Time and space distribution of CTC admissions and RDT-positive cholera cases. 1,976 patients admitted to the CTC over the study period reported residing in one of the 16 neighbourhoods of Uvira. Information on neighbourhood of residence was missing for 3 patients and 80 resided outside of Uvira municipality. For the entire city, this represented an annual incidence rate of 41.8 and 16.3 suspected and confirmed cases per 10,000 respectively (Table 3). Over the study period, weekly admissions incidence ranged from 0.13 to 4.3 per 10,000, with a median of 0.64 per 10,000.

The northern area had the highest annual incidence of both suspected and confirmed cases, 62.3 and 26.6 cases per 10,000 respectively. At the AS level, the highest annual incidence rate for suspected and confirmed cases were found in AS Kalundu Cepac (South), AS Kavimvira and AS Kilomoni (North). In a given week, the highest weekly incidence rate for admissions by AS over the entire study period was found in AS Kiyaya (North), with 26 cases per 10,000 per week, followed closely by AS Kavimvira and AS Kalundu Cepac. The lowest annual incidence rates by AS for both suspected and confirmed cases were found in AS Kimanga and AS Rombe I (Centre).

At least one admission to the CTC was recorded every week throughout the study, with at least one RDT-positive case in more than two thirds of the weeks. At the AS level, the proportion of weeks with admissions ranged from 27% for AS Kimanga, to 77% for AS Kasenga Cepac.

The distribution of cases over time shows similar patterns over the three areas of Uvira, although with different amplitudes in variations in the number of both suspected and RDT-positive cases (Fig 3). The number of AS contributing admissions in any given week ranged from 3 to 16 (median 8 AS affected, IQR 5–11 AS). Patients resided in 12 to 16 AS during the weeks in the very high admissions category.

Discussion

The epidemiology of cholera, particularly in endemic areas, is often described based on suspected rather than confirmed cases despite the lack of specificity of the standard case definition. This is likely to limit our understanding of cholera transmission and of how cholera prevention and outbreak response activities should be designed.

Using rapid diagnostic tests to confirm cholera amongst patients admitted to the Uvira CTC over an eighteen-month period, this study reveals that only about 40% of patients admitted to the CTC were confirmed as infected with cholera, giving an estimated city-wide incidence of confirmed cholera of sufficient severity to seek healthcare at the CTC of 16.5 cases per 10,000 annually. The data show an increasing probability of confirmed cholera when CTC admission rates are higher, with the proportion of confirmed cases reaching 52% during weeks with 50 CTC admissions or more. The data also suggest the occurrence of two distinct peaks each year, one around weeks 10 to 15 (March–April), and one around weeks 30 to 38 (August to October) of the calendar year. During these peaks, incidence of both admissions and suspected cases increased across all three areas of Uvira, although to a different degree. Both the incidence of CTC admissions and the case confirmation ratio were higher during weeks with lower rainfall. This study however confirms the endemic nature of cholera in Uvira, with confirmed cholera cases identified during 70% of the weeks observed—57 out of 81 with at least one patient enrolled. The incidence of CTC admissions, case confirmation ratio and the number of weeks with suspected and confirmed cases are all higher in the northern area of town, compared to the central area, and to a lesser degree to the southern area. This study also shows that at the individual level, female patients aged 5 to 15, male patients between 2 and 4 and over 15 years of age, patients admitted for 3 nights or more, and patients transferred to other departments are more likely to be RDT-positive. Boys under two years of age are however less likely to test positive with RDT.

Table 3. CTC admissions and RDT-positive cholera incidence rates and endemicity by neighbourhood and area of residence.

	Estimated population in June 2016	Number of admissions / patients enrolled / RDT-positive over study period	Confirmation ratio estimated from multivariable model	Annual CTC admissions / confirmed cholera incidence* per 10'000	Distribution of weekly CTC admissions incidence per 10'000 over study period		Number of weeks with at least one admission	Number of weeks with patients enrolled / at least one confirmed case
					Median (IQR)	Maximum	n (% over study period)	
AS Kasenga_Cepac	19182	241 / 158 / 60	41.6%	63.5 / 26.4	0.52 (0.52–2.09)	7.82	64 (77%)	46 / 25
AS Kasenga_Etat	24200	210 / 159 / 71	44.9%	46 / 20.6	0.41 (0.21–1.24)	6.2	62 (75%)	54 / 30
AS Kavimvira	13959	195 / 156 / 78	45.1%	70.6 / 31.8	0.72 (0–1.43)	24.36	51 (61%)	43 / 23
AS Kilomoni	11162	164 / 117 / 38	38.4%	83.7 / 32.1	0.9 (0–1.79)	12.54	53 (64%)	40 / 20
AS Kiyaya	6543	81 / 59 / 28	41.6%	65.4 / 27.2	0 (0–1.53)	25.98	34 (41%)	27 / 16
NORTH	**75046**	**891 / 649 / 275**	**42.6%**	**62.3 / 26.6**	**0.8 (0.4–1.6)**	**10.66**	**81 (98%)**	**73 / 50**
AS Mitumba	19610	137 / 24 / 5	30.9%	39.5 / 12.2	0.51 (0–1.02)	6.63	47 (57%)	42 / 19
AS Mulongwe	12815	98 / 90 / 32	32.7%	40.8 / 13.4	0 (0–1.56)	5.46	38 (46%)	32 / 14
AS Kimanga	13879	35 / 72 / 19	30.6%	12.9 / 3.9	0 (0–0.72)	2.88	22 (27%)	18 / 4
AS RombeI	15092	53 / 35 / 14	32.6%	15.3 / 5	0 (0–0.66)	3.98	26 (31%)	21 / 12
AS Saint_Paul	21114	95 / 61 / 22	31.9%	25.6 / 8.2	0.47 (0–0.95)	3.32	51 (61%)	36 / 13
AS Tanganyika	15874	56 / 34 / 12	29.7%	20.6 / 6.1	0 (0–0.63)	2.52	36 (43%)	26 / 11
CENTRE	**98384**	**474 / 316 / 104**	**31.5%**	**26.2 / 8.2**	**0.41 (0.2–0.81)**	**2.54**	**74 (89%)**	**68 / 37**
AS Kabindula	15525	149 / 83 / 33	42.4%	42.8 / 18.1	0 (0–1.29)	12.88	39 (47%)	35 / 14
AS Kalundu_Catholique	7845	50 / 41 / 14	34.5%	36.4 / 12.5	0 (0–1.27)	5.1	36 (43%)	31 / 12
AS Kalundu_Cepac	10783	180 / 103 / 42	41.4%	82.9 / 34.3	0.93 (0–1.85)	22.26	51 (61%)	40 / 20
AS Kalundu_Etat	14378	160 / 121 / 54	39.4%	60.8 / 24	0.7 (0–2.09)	7.65	53 (64%)	43 / 24
AS Nyamianda	11202	72 / 47 / 14	38.9%	36.8 / 14.3	0 (0–0.89)	4.46	35 (42%)	29 / 11
SOUTH	**59733**	**611 / 395 / 157**	**40.3%**	**52.4 / 21.1**	**0.67 (0.33–1.34)**	**9.38**	**78 (94%)**	**69 / 37**
Total Uvira municipality	248476	1976 / 1360 / 536	39.4%	41.8 / 16.5	0.64 (0.34–1.27)	4.29	83 (100%)	81 / 57
Unknown residence location or outside Uvira		83 / 59 / 26						

* based on the mean number of admissions over 52 consecutive weeks (32 periods) and the confirmed/admission ratios estimated by the multivariable logistic regression mode

https://doi.org/10.1371/journal.pone.0201306.t003

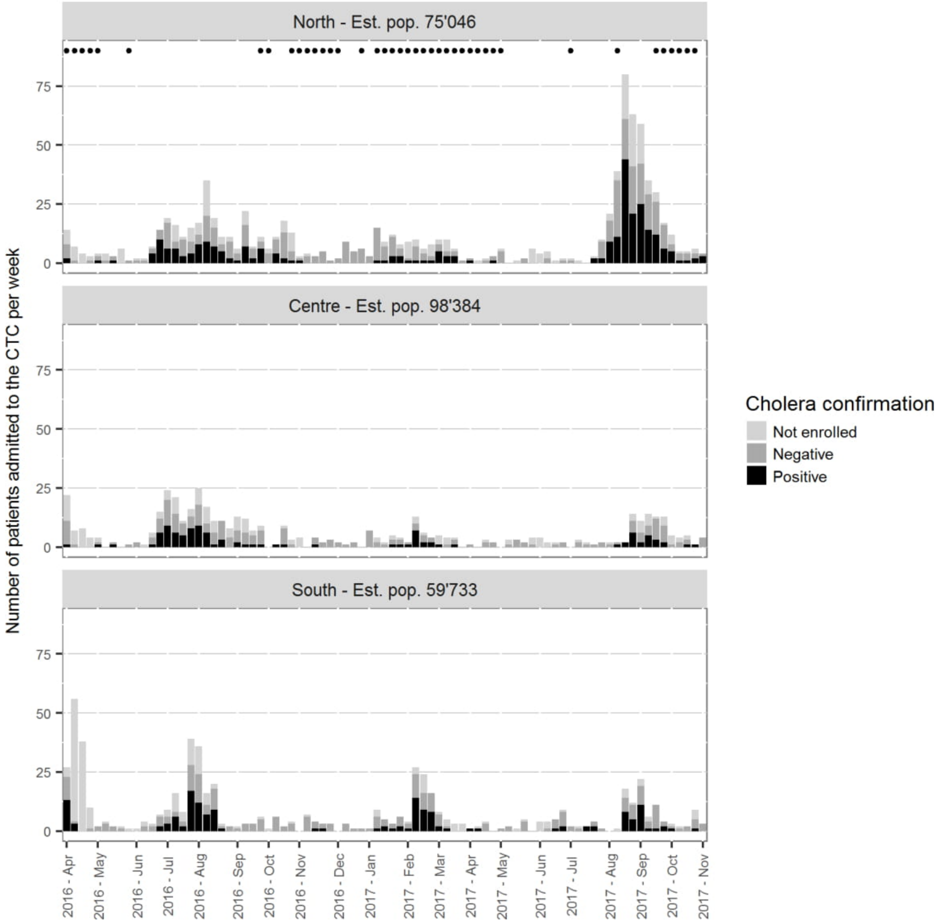


Fig 3. Weekly number of suspected, RDT-positive and RDT-negative cholera cases admitted to the CTC, by area of residence.

https://doi.org/10.1371/journal.pone.0201306.g003

The incidence rates of suspected and RDT-positive cholera cases as well as the confirmation ratio observed in this study are broadly similar to those reported for the Goma enhanced surveillance site of the AFRICHOL study, where annual incidence rates for suspected / confirmed cholera were estimated to 40.5 / 13.6 per 10,000 and the confirmation ratio to 50% [4]. Although estimated over a larger geographical area, suspected cholera incidence rates in the most affected province of Mozambique (Cabo Delgado) in 2009–2010 were much lower than those in Uvira with 14.2 cases per 10,000 per year, while cumulative suspected cholera incidence rates over 6 years (2011–2016) in cholera "hot spot" districts in Uganda did not exceed 100 per 10,000, suggesting an average annual rate of 16.7 suspected cases per 10,000 [20, 21]. The confirmed cholera incidence rate in Kolkata, India, was broadly similar to that observed here, with 22 confirmed cases per 10,000 per year[22]. Suspected cholera incidence rates in Uvira are however of a much lower magnitude than those observed during the first year of the Haiti cholera epidemic that started in October 2010, where the yearly incidence rate of suspected cholera reached 3,000 per 10,000 in some districts, and 490 per 10,000 over the entire country [23].

The higher incidence of both suspected and confirmed cases during the dry season observed in our study is in contrast to findings by Bompangue et al. in the region [9]. Data collected between 2002 and 2008 in Uvira showed a positive correlation between rainfall and number of suspected cholera cases reported in Uvira with two to five weeks latency. However, these results are not incompatible with a previous study showing that tap water supply interruptions were associated with a 2.5-fold increase in admissions to the CTC in the 12 following days as these tap water supply interruptions are longer and more frequent during the dry season due to more power outage [24]. Further investigations of seasonal patterns of admissions and RDT-positive cases, in relation with tap water availability and tap water supply interruptions will require specific spatio-temporal methods and a longer dataset.

Our study has various limitations. First, it is based on passive surveillance data collected at a healthcare facility. It is therefore limited to those seeking health care, and this decision is likely to depend on several factors including the perceived severity of symptoms, the monetary and opportunity costs of seeking care, and the perceived quality of care provided at the CTC. Even though all health posts and community health workers across town are instructed to refer acute diarrhoea cases to the CTC, where patients are treated for free, it is not possible to exclude the possibility that some groups of the population choose not to seek treatment if symptoms are deemed non-severe, or to seek treatment elsewhere in a private facility or from traditional healers. Healthcare seeking behaviours may also vary over time, especially during an outbreak. Further, it is likely that CTC admission criteria may be more stringent during an outbreak when capacity is already stretched. It is also worth noting that the clinical presentation of cholera infection can vary from no symptoms at all, to mild and short-lasting gastrointestinal symptoms and, in less than 10% of the cases, to severe diarrhoea with vomiting leading rapidly to dehydration and sometimes death [25]. Although it is accepted that symptom severity–and therefore the probability of a case being admitted to a health facility—is dependent on the inoculum ingested, infectious dose is influenced by a number of factors such as hypochlorhydria, concurrent ingestion with food, retinol A deficiency or blood group O [26–28]. Thus, clinical cholera illness incidence rates are only a proxy for the level of exposure of an individual to *Vibrio cholerae* and intensity of cholera transmission within a community.

Another important limitation of this study is the way cholera cases were confirmed. Although the use of RDTs on stool samples after an enrichment stage has been shown to perform as well as traditional culture methods, the sensitivity and specificity of this method were 86% and 100% respectively in comparison with PCR in another study [29]. The amount of stool collected with a rectal swab may also be less standardized, and the use of enriched rectal swabs with Crystal VC RDTs had a sensitivity and specificity of 92% and 91% respectively in comparison with culture methods [11]. Although WHO's standard definition of a confirmed cholera case is based on the isolation of *Vibrio cholerae* O1 or O139 in a suspected case's stools, this is not sufficient to attribute diarrhoea aetiology to cholera only, as acute diarrhoea may actually be caused by other enteric infections concurrently present, or by non-infectious causes. Asymptomatic cholera infections and associated *Vibrio cholerae* shedding in stools are common in endemic areas [26, 30]. A case-control or longitudinal study design would be needed to estimate in a robust manner the true cholera attributable diarrhoea burden in Uvira CTC patients. Finally, these results stem from only 69% of the patients admitted, and although those enrolled do not appear to differ importantly from those not participating on measured socio-demographic and clinical factors, selection bias associated with unmeasured factors amongst those enrolled cannot be excluded.

Our study confirms that cholera is indeed a public health concern in Uvira. Our findings substantiate the endemicity and incidence of cholera, with RDT-positive cholera cases admitted to the CTC nearly all year round, and from all neighbourhoods. Although a non-specific case definition is valuable for timely and sensitive outbreak detection, they highlight that a substantial proportion of suspected cases are not testing positive with RDTs, and that confirmation of suspected cases is needed to understand the true burden of cholera and to investigate risk factors for cholera, especially in endemic areas. This need to improve cholera surveillance with a more specific case definition was recently highlighted by another study in seven African countries [5]. This would also contribute to the planning, targeting and evaluation of cholera vaccine campaigns. In Uvira, deployment of an oral cholera vaccine alone would only address a fraction of the diarrhoea cases admitted to the CTC which are currently suspected to be cholera cases. Our study also demonstrates the feasibility of enhanced surveillance of cholera with RDTs over 18 months under challenging circumstances and limited human resources. Although a different clinical management is not called for, the high proportion of negative results amongst CTC admitted patients underlines as well the importance of high quality infection prevention and control (IPC) measures, to avoid cholera-negative patients becoming infected with cholera during their stay at the CTC. This is also true for patients transferred to other hospital departments–often paediatrics or obstetrics—where there may be less stringent infection control measures in place. The high proportion of negative results also raises the question of non-cholera aetiology of diarrhoea in CTC admitted patients, and of the burden attributable to other entero-pathogens that could potentially require specific control measures, such as rotavirus vaccination. Further, we found that only boys younger than 2 are less likely to be confirmed as cholera cases than patients from other older age groups, which raises questions about the exclusion of under 2s from the WHO suspected case definition outside of an outbreak [3].

The tap water supply improvements currently being implemented in Uvira provide an opportunity to study the effect of such improvements on both cholera and non-cholera diarrhoea. Exploration of time and space clustering of suspected and confirmed cases, in relation to the changes in tap water access and reliability, may provide explanations for the substantial variations in incidence rates observed between different areas of town and over time. An impact evaluation of these tap water supply improvements in Uvira is under way [31]. In addition, identification of other entero-pathogens shed by suspected cases will provide a better understanding of the non-cholera diarrhoeal aetiology in patients admitted to the CTC. The planned genomic characterization of cholera strains isolated from confirmed cholera cases in Uvira will provide insight into the regional dynamics of cholera transmission.

In summary, our data show that cholera is endemic throughout Uvira, but only 40% of admissions to Uvira CTC have a detectable cholera infection. Interventions aimed at reducing the burden of cholera and other diarrhoeal diseases should be a public health priority.

Supporting information

S1 Table. STROBE checklist.
(DOC)

S2 Table. Dataset.
(CSV)

Acknowledgments

We thank all the CTC staff who contributed to this study, as well as the Uvira district head doctor (MCZ) in Uvira, Dr P. Nimi, and the provincial health division (DPS) in Bukavu. We are also grateful to OXFAM GB staff for their invaluable support in Uvira. In addition, we truly appreciate the valuable inputs provided by M. Lemenager and F. Bedecarrats from the Agence

Francaise de Development and T. Vandevelde, T. Constant and D. Machuel from the Veolia Foundation during the study.

Author Contributions

Conceptualization: Aurelie Jeandron, Simon Cousens.

Data curation: Aurelie Jeandron, Baron Bashige Rumedeka, Jaime Mufitini Saidi.

Formal analysis: Aurelie Jeandron.

Investigation: Aurelie Jeandron, Baron Bashige Rumedeka, Jaime Mufitini Saidi.

Supervision: Oliver Cumming, Simon Cousens.

Writing – original draft: Aurelie Jeandron, Simon Cousens.

Writing – review & editing: Aurelie Jeandron, Oliver Cumming, Simon Cousens.

References

1. World Health Organisation. WER Volume 92, 2017, No. 36. Wkly Epidemiol Rec. 2017;92(36):521–30.
2. Ali M, Nelson AR, Lopez AL, Sack DA. Updated global burden of cholera in endemic countries. PLoS Negl Trop Dis. 2015; 9(6):e0003832. https://doi.org/10.1371/journal.pntd.0003832 PMID: 26043000
3. World Health Organisation. Outbreak surveillance and response in humanitarian emergencies—WHO guidelines for EWARN implementation. Geneva: World Health Organisation, Emergencies DCiH; 2012.
4. Sauvageot D, Njanpop-Lafourcade BM, Akilimali L, Anne JC, Bidjada P, Bompangue D, et al. Cholera Incidence and Mortality in Sub-Saharan African Sites during Multi-country Surveillance. PLoS Negl Trop Dis. 2016; 10(5):e0004679. https://doi.org/10.1371/journal.pntd.0004679 PMID: 27186885
5. Nadri J, Sauvageot D, Njanpop-Lafourcade BM, Baltazar CS, Banla Kere A, Bwire G, et al. Sensitivity, Specificity, and Public-Health Utility of Clinical Case Definitions Based on the Signs and Symptoms of Cholera in Africa. Am J Trop Med Hyg. 2018; 98(4):1021–30. https://doi.org/10.4269/ajtmh.16-0523 PMID: 29488455
6. Lucien MA, Schaad N, Steenland MW, Mintz ED, Emmanuel R, Freeman N, et al. Identifying the most sensitive and specific sign and symptom combinations for cholera: results from an analysis of laboratory-based surveillance data from Haiti, 2012–2013. Am J Trop Med Hyg. 2015; 92(4):758–64. https://doi.org/10.4269/ajtmh.14-0429 PMID: 25732682
7. Rebaudet S, Sudre B, Faucher B, Piarroux R. Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. J Infect Dis. 2013; 208 Suppl 1:S46–54.
8. Emch M, Feldacker C, Islam MS, Ali M. Seasonality of cholera from 1974 to 2005: a review of global patterns. Int J Health Geogr. 2008; 7:31. https://doi.org/10.1186/1476-072X-7-31 PMID: 18570659
9. Bompangue D, Giraudoux P, Handschumacher P, Piarroux M, Sudre B, Ekwanzala M, et al. Lakes as source of cholera outbreaks, Democratic Republic of Congo. Emerg Infect Dis. 2008; 14(5):798–800. https://doi.org/10.3201/eid1405.071260 PMID: 18439365
10. Bompangue D, Giraudoux P, Piarroux M, Mutombo G, Shamavu R, Sudre B, et al. Cholera epidemics, war and disasters around Goma and Lake Kivu: an eight-year survey. PLoS Negl Trop Dis. 2009; 3(5): e436. https://doi.org/10.1371/journal.pntd.0000436 PMID: 19436726
11. Bhuiyan NA, Qadri F, Faruque ASG, Malek MA, Salam MA, Nato F, et al. Use of Dipsticks for Rapid Diagnosis of Cholera Caused by Vibrio cholerae O1 and O139 from Rectal Swabs. Journal of Clinical Microbiology. 2003; 41(8):3939–41. https://doi.org/10.1128/JCM.41.8.3939-3941.2003 PMID: 12904424
12. Page A-L, Alberti KP, Mondonge V, Rauzier J, Quilici M-L, Guerin PJ. Evaluation of a Rapid Test for the Diagnosis of Cholera in the Absence of a Gold Standard. PLOS ONE. 2012; 7(5):e37360. https://doi.org/10.1371/journal.pone.0037360 PMID: 22666350
13. Hartung C, Lerer A, Anokwa Y, Tseng C, Brunette W, Borriello G. Open data kit: tools to build information services for developing regions. Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development; London, United Kingdom. 2369236: ACM; 2010. p. 1–12.

14. GPM IMERG Late Precipitation L3 1 day 0.1 degree x 0.1 degree V04 [Internet]. Goddard Earth Sciences Data and Information Services Center (GES DISC). 2016 [cited 12/12/2017]. Available from: https://disc.gsfc.nasa.gov/datacollection/GPM_3IMERGDL_04.html.

15. Climate Change Knowledge Portal—Burundi Dashboard [Internet]. The World Bank Group. [cited 12/01/2018]. Available from: http://sdwebx.worldbank.org/climateportal/countryprofile/home.cfm?page=country_profile&CCode=BDI&ThisTab=ClimateBaseline.

16. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Computing; 2016.

17. Global Task Force For Cholera Control—Surveillance WG. Interim Guidance Document on Cholera Surveillance. 2017.

18. Katz E. Bias in Conditional and Unconditional Fixed Effects Logit Estimation. Political Analysis. 2001; 9 (4):379–84.

19. Leibovici-Weissman Ya, Neuberger A, Bitterman R, Sinclair D, Salam MA, Paul M. Antimicrobial drugs for treating cholera. Cochrane Database of Systematic Reviews. 2014(6).

20. Gujral L, Sema C, Rebaudet S, Taibo CL, Manjate AA, Piarroux R, et al. Cholera epidemiology in Mozambique using national surveillance data. J Infect Dis. 2013; 208 Suppl 1:S107–14.

21. Bwire G, Ali M, Sack DA, Nakinsige A, Naigaga M, Debes AK, et al. Identifying cholera "hotspots" in Uganda: An analysis of cholera surveillance data from 2011 to 2016. PLoS Negl Trop Dis. 2017; 11(12): e0006118. https://doi.org/10.1371/journal.pntd.0006118 PMID: 29284003

22. Sur D, Deen J, Manna B, Niyogi S, Deb A, Kanungo S, et al. The burden of cholera in the slums of Kolkata, India: data from a prospective, community based study. Archives of Disease in Childhood. 2005; 90(11):1175–81. https://doi.org/10.1136/adc.2004.071316 PMID: 15964861

23. Gaudart J, Rebaudet S, Barrais R, Boncy J, Faucher B, Piarroux M, et al. Spatio-Temporal Dynamics of Cholera during the First Year of the Epidemic in Haiti. PLOS Neglected Tropical Diseases. 2013; 7(4): e2145. https://doi.org/10.1371/journal.pntd.0002145 PMID: 23593516

24. Jeandron A, Saidi JM, Kapama A, Burhole M, Birembano F, Vandevelde T, et al. Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo. PLoS Med. 2015; 12(10):e1001893. https://doi.org/10.1371/journal.pmed.1001893 PMID: 26506001

25. Kaper JB, Morris JG Jr., Levine MM. Cholera. Clinical microbiology reviews. 1995; 8(1):48–86. PMID: 7704895

26. Harris JB, LaRocque RC, Chowdhury F, Khan AI, Logvinenko T, Faruque AS, et al. Susceptibility to Vibrio cholerae infection in a cohort of household contacts of patients with cholera in Bangladesh. PLoS Negl Trop Dis. 2008; 2(4):e221. https://doi.org/10.1371/journal.pntd.0000221 PMID: 18398491

27. Harris JB, LaRocque RC, Qadri F, Ryan ET, Calderwood SB. Cholera. Lancet. 2012; 379(9835):2466–76. https://doi.org/10.1016/S0140-6736(12)60436-X PMID: 22748592

28. Valcin CL, Severe K, Riche CT, Anglade BS, Moise CG, Woodworth M, et al. Predictors of disease severity in patients admitted to a cholera treatment center in urban Haiti. Am J Trop Med Hyg. 2013; 89 (4):625–32. https://doi.org/10.4269/ajtmh.13-0170 PMID: 24106188

29. Ontweka LN, Deng LO, Rauzier J, Debes AK, Tadesse F, Parker LA, et al. Cholera Rapid Test with Enrichment Step Has Diagnostic Performance Equivalent to Culture. PLoS One. 2016; 11(12): e0168257. https://doi.org/10.1371/journal.pone.0168257 PMID: 27992488

30. Nelson EJ, Harris JB, Morris JG, Calderwood SB, Camilli A. Cholera transmission: the host, pathogen and bacteriophage dynamic. Nature reviews Microbiology. 2009; 7(10): https://doi.org/10.1038/nrmicro2204 PMID: 19756008

31. ClinicalTrials.gov: Impact Evaluation of Urban Water Supply Improvements on Cholera and Other Diarrhoeal Diseases in Uvira, Democratic Republic of Congo—Bethesda (MD): National Library of Medicine (US); [Available from: https://ClinicalTrials.gov/show/NCT02928341.

3.2 Confirmation of cholera by rapid diagnostic test amongst patients admitted to the CTC: analysis of a dataset extended to April 2019

The analysis performed in 2018, on the data related to the confirmation of cholera by rapid diagnostic tests amongst patients admitted to the Uvira CTC between April 2016 and November 2017, was repeated on a second dataset, running from November 2017 to April 2019. The objective was to confirm the results previously obtained on demographic, clinical, geographical and time patterns of cholera confirmation amongst CTC admitted patients.

3.2.1 Data and methods

Data sources and methods were the same as in the published study, except for population estimates and rainfall estimates data **[1]**.

Population census estimates obtained from the Uvira town hall in November 2017 highlighted a boundary dispute between the municipality and the territory of Uvira on the outskirts of AS Kabindula and AS Saint-Paul, and on the outskirts of AS Kasenga Cepac. In order to avoid inaccuracies in population estimates, we classified as "residing outside Uvira or missing" the CTC patients residing in areas of AS Kabindula, AS Saint-Paul and AS Kasenga for which no population estimates were included in the overlapping neighbourhoods of Kabindula, Saint-Paul, Kakombe and Kibondwe (see chapter *Data and methods: an overview*, section *2.2.2 Delineating the administrative and health divisions*). The same correction was performed on the first study period, leading to the re-classification of 11 patients as "residing outside Uvira or missing" (4 in AS Saint-Paul and 7 in AS Kabindula), hence affecting very slightly initial incidence rates estimates by AS.

The remote sensing data (GPM Level 3 IMERG Late Daily 0.1-degree x 0.1-degree v04) used initially to categorise weeks of admission as dry/rainy were discontinued by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC) and replaced by later versions of the dataset **[2]**. In 2018, the CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) rainfall estimates were satisfactorily validated against meteorological station data over Eastern Africa **[3,4]**. These estimates were therefore used here over the extended period analyses.

Weekly rainfall estimates in Uvira for the first analysis period differ significantly between IMERG and CHIRPS, with CHIRPS typically estimating substantially higher rainfall. Using the overall median weekly cumulated rainfall (21.2 mm/week) for CHIRPS as cut-off between rainy/dry weeks, as was done initially for IMERG with a median of 3.65

mm/week, leads to 10 and 2 weeks in the initial study being classified as “dry” instead of “rainy” and as “rainy” instead of “dry” respectively (**Figure 3-1**).

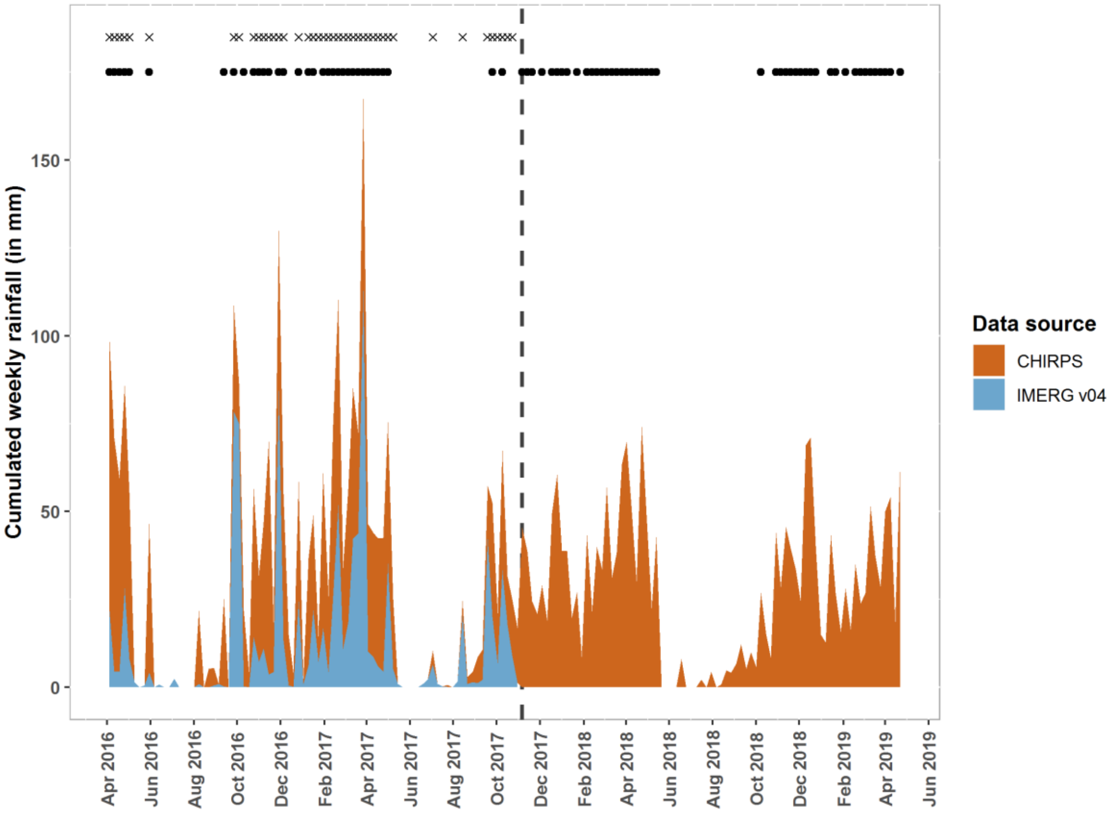


Figure 3-1 Cumulated weekly rainfall estimates for IMERG v04 and CHIRPS over initial and extended study periods. (X) and (•) represent weeks classified as rainy, based on IMERG v04 (X) or CHIRPS (•). The two study periods are indicated by the dashed line.

3.2.2 Results

Admissions and enrolment

Between the 6th of November 2017 and the 28th of April 2019 (77 weeks), 1,320 patients were admitted to the CTC. Of those, 777 patients (58.9 %) participated to the cholera confirmation study, against 69 % during the initial study period. Overall, between the 4th of April 2016 and the 28th of April 2019 (160 weeks), 3,379 patients were admitted and 2,195 patients (65 %) participated to the study. Characteristics of admitted patients differed marginally between the 2 study periods. The distribution of admissions across incidence categories – ie lower than 10, between 10 and 19, between 20 and 49 and more than 50 patients per week - differs, reflecting a lower weekly rate of admissions overall during the 2nd study period. Indeed while the weeks were nearly equally distributed across the four categories during the initial period - 22 (26.5 %), 22 (26.5 %), 19 (22.9 %) and 15 (18.1 %) weeks in low, moderate, high and very high categories respectively, - the distribution for the 2nd period was markedly skewed towards the lower incidence category with 34 (44 %), 20 (26 %), 18 (23 %) and 5 (7 %) weeks respectively. Despite being lower generally, the rate of enrolment for the 2nd period is well balanced across patient characteristics. As for the initial study, patients admitted during high incidence weeks, staying less than 24 hours at the CTC, deceased or transferred to another hospital department or under the age of 2, were less likely to be enrolled in the study. Characteristics of admitted and enrolled patients during the 1st and 2nd study periods and overall, are shown in **Table 3-1**.

Table 3-1 Characteristics of patients admitted to the Uvira CTC and enrolled in the cholera confirmation study for both study periods

	FIRST STUDY PERIOD		SECOND STUDY PERIOD		OVERALL	
	Admitted n (% of total)	Enrolled n (% of admitted)	Admitted n (% of total)	Enrolled n (% of admitted)	Admitted n (% of total)	Enrolled n (% of admitted)
Total	2059 (100%)	1419 (68.9%)	1320 (100%)	777 (58.9%)	3379 (100%)	2195 (65%)
Sex						
Male	1038 (50.4%)	713 (68.7%)	653 (49.5%)	380 (58.2%)	1691 (50%)	1101 (65.1%)
Female	1019 (49.5%)	705 (69.2%)	667 (50.5%)	397 (59.5%)	1686 (49.9%)	1093 (64.8%)
Missing	2 (0.1%)	1 (50%)	0 (0%)	0 (0%)	2 (0.1%)	1 (50%)
Age group						
Under 2-year-old	62 (3%)	37 (59.7%)	35 (2.7%)	14 (40%)	97 (2.9%)	51 (52.6%)
2 to 4-year-old	274 (13.3%)	182 (66.4%)	134 (10.2%)	60 (44.8%)	408 (12.1%)	242 (59.3%)
5 to 15-year-old	594 (28.8%)	412 (69.4%)	333 (25.2%)	198 (59.5%)	927 (27.4%)	609 (65.7%)
16 and older	1124 (54.6%)	787 (70%)	818 (62%)	505 (61.7%)	1942 (57.5%)	1292 (66.5%)
Missing	5 (0.2%)	1 (20%)	0 (0%)	0 (0%)	5 (0.1%)	1 (20%)
Weekly admissions incidence						
Low (1 to 9 admissions)	156 (7.6%)	109 (69.9%)	158 (12%)	114 (72.2%)	314 (9.3%)	223 (71%)
Moderate (10 to 19 admissions)	302 (14.7%)	208 (68.9%)	268 (20.3%)	191 (71.3%)	570 (16.9%)	399 (70%)
High (20 to 49 admissions)	571 (27.7%)	428 (75%)	521 (39.5%)	317 (60.8%)	1092 (32.3%)	745 (68.2%)
Very High (50 to 103 admissions)	1030 (50%)	674 (65.4%)	373 (28.3%)	155 (41.6%)	1403 (41.5%)	828 (59%)
Duration of stay						
Less than 24h	160 (7.8%)	84 (52.5%)	134 (10.2%)	53 (39.6%)	294 (8.7%)	137 (46.6%)
1 to 2 nights	1086 (52.7%)	771 (71%)	716 (54.2%)	463 (64.7%)	1802 (53.3%)	1233 (68.4%)
3 to 4 nights	501 (24.3%)	357 (71.3%)	318 (24.1%)	206 (64.8%)	819 (24.2%)	563 (68.7%)
5 nights or more	120 (5.8%)	81 (67.5%)	67 (5.1%)	33 (49.3%)	187 (5.5%)	114 (61%)
Missing	192 (9.3%)	126 (65.6%)	85 (6.4%)	22 (25.9%)	277 (8.2%)	148 (53.4%)
Outcome						
Discharged	1793 (87.1%)	1267 (70.7%)	1167 (88.4%)	721 (61.8%)	2960 (87.6%)	1987 (67.1%)
Transferred	111 (5.4%)	47 (42.3%)	83 (6.3%)	26 (31.3%)	194 (5.7%)	73 (37.6%)
Death	10 (0.5%)	3 (30%)	10 (0.8%)	1 (10%)	20 (0.6%)	4 (20%)
Missing	145 (7%)	102 (70.3%)	60 (4.5%)	29 (48.3%)	205 (6.1%)	131 (63.9%)
Area of residence						
North	891 (43.3%)	649 (72.8%)	597 (45.2%)	375 (62.8%)	1488 (44%)	1024 (68.8%)
Centre	474 (23%)	316 (66.7%)	263 (19.9%)	162 (61.6%)	737 (21.8%)	474 (64.3%)
South	611 (29.7%)	395 (64.6%)	403 (30.5%)	208 (51.6%)	1014 (30%)	598 (59%)
Outside Uvira/missing	83 (4%)	59 (71.1%)	57 (4.3%)	32 (56.1%)	140 (4.1%)	99 (70.7%)
Weekly rain category (CHIRPS data)						
Dry (<= 21.2 mm)	1337 (64.9%)	984 (73.6%)	624 (47.3%)	324 (51.9%)	1961 (58%)	1308 (66.7%)
Rainy (> 21.2 mm)	722 (35.1%)	434 (60.1%)	696 (52.7%)	453 (65.1%)	1418 (42%)	887 (62.6%)

RDT-confirmation results

During the 2nd study period, 335 (43.1 %) of the 777 patients enrolled tested positive to *V. cholerae* O_1 with the Crystal® VC RDT following enrichment in APW for about 6 h. Overall, 897 (40.9 %) of the 2,195 patients enrolled tested positive.

69 (3.1 %) participants in total had received antibiotics before rectal swab sampling. Antibiotic administration before rectal sampling did not appear to be associated with a lower rate of confirmation during either study period, with a mean time between treatment and sampling of 6.2 h. It is worth noting however that among those who received antibiotics the mean time since antibiotic administration is lower for patients testing positive (5.3 h vs 6.8 h, p-value of t-test for equal means = 0.01).

When restricted to the 2nd period data, univariable logistic regression only highlights 3 variables independently associated with RDT-confirmation, instead of 6 variables for the initial analysis: age group, weekly admissions incidence and duration of stay at the CTC. There is no evidence of an independent association with demographic group (defined as age group stratified by sex), outcome of admission, area of residence or weekly rain category.

A multivariable model run on the 2nd period data and including only the above three variables finds that patients aged 5 to 15 are less likely to test positive than those older than 15 (adjusted odds ratio aOR 0.66; 95% CI 0.46 – 0.94). It also confirms that increasing weekly admissions incidence and duration of stay at the CTC for 3 days or more are strongly associated with higher odds of participants testing positive. Results of univariable and multivariable models fitted to the 2nd study period data are presented in **Table 3-2**.

In order to explore the impact of the change in rainfall data estimates used – CHIRPS instead of IMERG v04, univariable and multivariable logistic regressions were performed again on the 1st period dataset with the new rainy/dry variable definition. Independently of other variables, the relationship between weekly rain category and confirmation rate remains similar, with patients admitted on dry weeks having about 1.5 times the odds of testing positive than those admitted during rainy weeks. In a multivariable model however, this relationship now appears totally confounded (aOR 0.98; 95% CI 0.73 – 1.33) with the other variables associated with RDT-confirmation in the multivariable model. The change in rainfall data source affects very little however the levels of association observed initially with demographic groups, admissions incidence, duration and outcome of stay at the CTC and area of residence.

We finally fitted a multivariable model to both study periods (n=2,195) and including all the variables independently associated with RDT-confirmation in either dataset: demographic group, admissions incidence, duration and outcome of stay at the CTC, area of residence and weekly rainfall category. Results of this multivariable model are reported in **Table 3-3**. In comparison with a similarly formulated model fitted to the 1st period data only, the following changes can be noticed: 1) evidence of association of RDT-confirmation with each variable is the same, except for area of residence for which evidence is weaker ; 2) strength of association of RDT-confirmation with weekly admissions incidence and to a lesser degree with duration of stay at the CTC increases; 3) strength of association of RDT-confirmation with individual demographic groups and being transferred to another department decreases slightly; 4) the association of RDT-confirmation with rainfall category is inverted, with patients admitted during rainy weeks now more likely to test positive than those admitted during dry weeks.

Table 3-2 Predictors for RDT-confirmation of cholera amongst patients admitted to the CTC during the 2nd study period

	Enrolled	Positive	Univariable logistic regression		Multivariable logistic regression	
	n (% of total)	n (% of enrolled)	OR (95% CI)	*p-value**	aOR (95% CI)	*p-value**
Total	777 (100%)	335 (43.1%)				
Sex						
Male	380 (48.9%)	169 (44.5%)	1.11 (0.84 - 1.48)	*0.45*		
Female	397 (51.1%)	166 (41.8%)	reference			
Age group						
Under 2-year-old	14 (1.8%)	6 (42.9%)	0.88 (0.3 - 2.58)	*0.08*	0.8 (0.26 - 2.44)	*0.15*
2 to 4-year-old	60 (7.7%)	27 (45%)	0.96 (0.56 - 1.65)		0.83 (0.46 - 1.47)	
5 to 15-year-old	198 (25.5%)	70 (35.4%)	0.64 (0.46 - 0.9)		0.66 (0.46 - 0.94)	
16 and older	505 (65%)	232 (45.9%)	reference		reference	
Demographic group						
Girls - under 2	9 (1.2%)	4 (44.4%)	1.08 (0.28 - 4.11)	*0.19*		
Girls - 2 to 4 years	26 (3.3%)	12 (46.2%)	1.16 (0.51 - 2.59)			
Girls - 5 to 15 years	99 (12.7%)	38 (38.4%)	0.84 (0.52 - 1.35)			
Women - 16 and older	263 (33.8%)	112 (42.6%)	reference			
Boys - under 2	5 (0.6%)	2 (40%)	0.9 (0.15 - 5.47)			
Boys - 2 to 4 years	34 (4.4%)	15 (44.1%)	1.06 (0.52 - 2.19)			
Boys - 5 to 15 years	99 (12.7%)	32 (32.3%)	0.64 (0.4 - 1.05)			
Men - 15 and older	242 (31.1%)	120 (49.6%)	1.33 (0.93 - 1.88)			
Weekly admissions incidence						
Low (1 to 9 admissions)	114 (14.7%)	14 (12.3%)	reference	*p<0.001*	reference	*p<0.001*
Moderate (10 to 19 admissions)	191 (24.6%)	76 (39.8%)	4.72 (2.51 - 8.86)		4.54 (2.35 - 8.77)	
High (20 to 49 admissions)	317 (40.8%)	155 (48.9%)	6.83 (3.75 - 12.47)		6.21 (3.31 - 11.64)	
Very High (50 to 103 admissions)	155 (19.9%)	90 (58.1%)	9.89 (5.19 - 18.83)		8.24 (4.2 - 16.16)	
Antibiotics treatment before sampling						
No	738 (95%)	315 (42.7%)	0.71 (0.37 - 1.35)	*0.29*		
Yes	39 (5%)	20 (51.3%)	reference			
Duration of stay						
Less than 24h	53 (6.8%)	13 (24.5%)	reference	*p<0.001*	reference	*p<0.001*
1 to 2 nights	463 (59.6%)	183 (39.5%)	2.01 (1.05 - 3.86)		1.69 (0.85 - 3.36)	
3 to 4 nights	206 (26.5%)	113 (54.9%)	3.74 (1.89 - 7.4)		2.75 (1.34 - 5.64)	
5 nights or more	33 (4.2%)	23 (69.7%)	7.08 (2.68 - 18.69)		4.76 (1.76 - 12.91)	
Missing	22 (2.8%)	3 (13.6%)	excluded			
Outcome						
Discharged	721 (92.8%)	320 (44.4%)	reference	*0.83*		
Transferred	26 (3.3%)	11 (42.3%)	0.92 (0.42 - 2.03)			
Death	1 (0.1%)	1 (100%)	excluded			
Missing	29 (3.7%)	3 (10.3%)				
Area of residence						
North	375 (48.3%)	156 (41.6%)	0.68 (0.47 - 0.98)	*0.13*		
Centre	162 (20.8%)	83 (51.2%)	reference			
South	208 (26.8%)	83 (39.9%)	0.63 (0.42 - 0.96)			
Outside Uvira or missing	32 (4.1%)	13 (40.6%)	0.65 (0.3 - 1.41)			
Weekly rain category (CHIRPS data)						
Dry (<= 21.2 mm)	324 (41.7%)	131 (40.4%)	0.83 (0.62 - 1.11)	*0.2*		
Rainy (> 21.2 mm)	453 (58.3%)	204 (45%)	reference			

* $\chi 2$ significance test of likelihood ratio (LR test)

Confirmation rate estimates of non-enrolled patients

We based our estimation of confirmation amongst non-enrolled patients on the multi-variable model fitted to both study periods and reported in **Table 3-3**. This model predicts an overall confirmation rate over 3,055 admissions records of 42.2 % (43.2 % and 40.6 % for the 1st and 2nd study periods respectively). A total of 324 patients (9.6 % of all admissions) were not enrolled in the confirmation study and have an incomplete record.

Table 3-3 Predictors for RDT-confirmation of cholera amongst patients admitted to the CTC over combined periods – model used for RDT-confirmation prevalence estimates

	Enrolled n (% of total)	Positive n (% positive)	Multivariable logistic regression aOR (95% CI)	*p-value**
Total	2195 (100%)	897 (40.9%)		
Demographic group				
Girls - under 2	25 (1.1%)	10 (40%)	1.19 (0.51 - 2.77)	
Girls - 2 to 4 years	113 (5.1%)	47 (41.6%)	1.01 (0.64 - 1.61)	
Girls - 5 to 15 years	295 (13.4%)	137 (46.4%)	1.44 (1.06 - 1.95)	
Women - 16 and older	667 (30.4%)	245 (36.7%)	reference	*p<0.001*
Boys - under 2	26 (1.2%)	3 (11.5%)	0.23 (0.07 - 0.79)	
Boys - 2 to 4 years	129 (5.9%)	61 (47.3%)	1.39 (0.92 - 2.1)	
Boys - 5 to 15 years	313 (14.3%)	109 (34.8%)	0.85 (0.63 - 1.17)	
Men - 15 and older	625 (28.5%)	285 (45.6%)	1.57 (1.23 - 2.01)	
Weekly admissions incidence				
Low (1 to 9 admissions)	223 (10.2%)	28 (12.6%)	reference	
Moderate (10 to 19 admissions)	399 (18.2%)	120 (30.1%)	2.66 (1.63 - 4.36)	*p<0.001*
High (20 to 49 admissions)	745 (33.9%)	310 (41.6%)	4.07 (2.56 - 6.47)	
Very High (50 to 103 admissions)	828 (37.7%)	439 (53%)	6.36 (3.99 - 10.15)	
Duration of stay				
Less than 24h	137 (6.2%)	39 (28.5%)	reference	
1 to 2 nights	1233 (56.2%)	410 (33.3%)	1.15 (0.74 - 1.79)	
3 to 4 nights	563 (25.6%)	317 (56.3%)	2.66 (1.67 - 4.22)	*p<0.001*
5 nights or more	114 (5.2%)	76 (66.7%)	3.83 (2.12 - 6.93)	
Missing	148 (6.7%)	55 (37.2%)	excluded	
Outcome				
Discharged	1987 (90.5%)	813 (40.9%)	reference	
Transferred	73 (3.3%)	36 (49.3%)	1.98 (1.13 - 3.48)	*0.02*
Death	4 (0.2%)	4 (100%)	excluded	
Missing	131 (6%)	44 (33.6%)	excluded	
Area of residence				
North	1023 (46.6%)	431 (42.1%)	1.27 (0.99 - 1.62)	
Centre	479 (21.8%)	188 (39.2%)	reference	*0.07*
South	603 (27.5%)	240 (39.8%)	0.96 (0.73 - 1.26)	
Outside Uvira or missing	90 (4.1%)	38 (42.2%)	1.23 (0.75 - 2.02)	
Weekly rain category (CHIRPS data)				
Dry (<= 21.2 mm)	1308 (59.6%)	505 (38.6%)	0.81 (0.66 - 0.99)	*0.04*
Rainy (> 21.2 mm)	887 (40.4%)	392 (44.2%)	reference	

* $\chi 2$ significance test of likelihood ratio (LR test)

Time and space patterns of admissions and RDT-confirmation

1,262 patients admitted to the CTC during the 2nd study period were residing in one of the 16 AS of Uvira (excluding disputed avenues). Information on neighbourhood of residence was missing for 2 patients and 56 resided outside of Uvira (now including disputed avenues). This represents an annual incidence rate of 35.5 CTC admissions per 10,000 inhabitants for the entire city, somewhat lower than during the 1st analysis (41.8 per 10,000). With a slightly lower estimated confirmation rate than during the 1st period, the annual incidence rate of RDT-confirmed cholera was estimated to 15 cases per 10,000 for the 2nd period.

This lower annual incidence rate of CTC admissions is also reflected in the weekly incidence rate distribution, that ranged from 0 to 3.4 patients per 10,000 and a median of 0.4 per 10,000 (in comparison with 0.64 per 10,000 during the the 1st period). The northern and southern areas experienced similar annual incidence of both suspected and RDT-confirmed cases, with 47.4 and 20.6 cases per 10,000 respectively in the north, and 45.2 and 19.1 cases per 10,000 in the south. As for the 1st study period, the centre of town was noticeably less affected than the northern and southern areas, with 18.8 and 7.3 cases of suspected and RDT-confirmed cases per 10,000.

At the AS level, the highest annual incidence rate for both suspected and RDT-confirmed cases was observed in AS Kalundu Cepac in the south (71.5 and 31.5 per 10,000 respectively) closely followed by AS Kavimvira in the north (65.1 and 30.1 per 10,000 respectively). In a given week, AS Kalundu Cepac also experienced the highest incidence rate of suspected cases (18.2 per 10,000), followed again closely by AS Kavimvira (13.6 per 10,000). Annual incidence rates for suspected and RDT-confirmed cases were the lowest in AS Tanganyika, AS Kimanga and AS Rombe I in the central part of town.

At least one admission was recorded every week during the 2nd study period, although no resident of Uvira was admitted on week 14 of 2019 (early April). At least one suspected case from AS Kasenga Cepac, AS Kasenga Etat and AS Kavimvira (north) was admitted during more than two thirds of the weeks, and those AS were the only ones with a weekly median incidence above 0. Patients were enrolled in the study for 71 weeks out of 77, and at least one positive RDT result was recorded for 43 weeks (60.1%). Overall, the number of AS contributing admissions in any given week ranged from 0 to 16 (median 7; IQR 5-10). During very high admission incidence weeks, between 11 and 16 AS contribute to CTC admissions. **Figure 3-2** shows the distribution of cases over time and areas of Uvira for both study periods. **Table 3-4** and **Figure 3-3** summarise incidence rates and endemicity by neighbourhood for 2nd and both study periods.

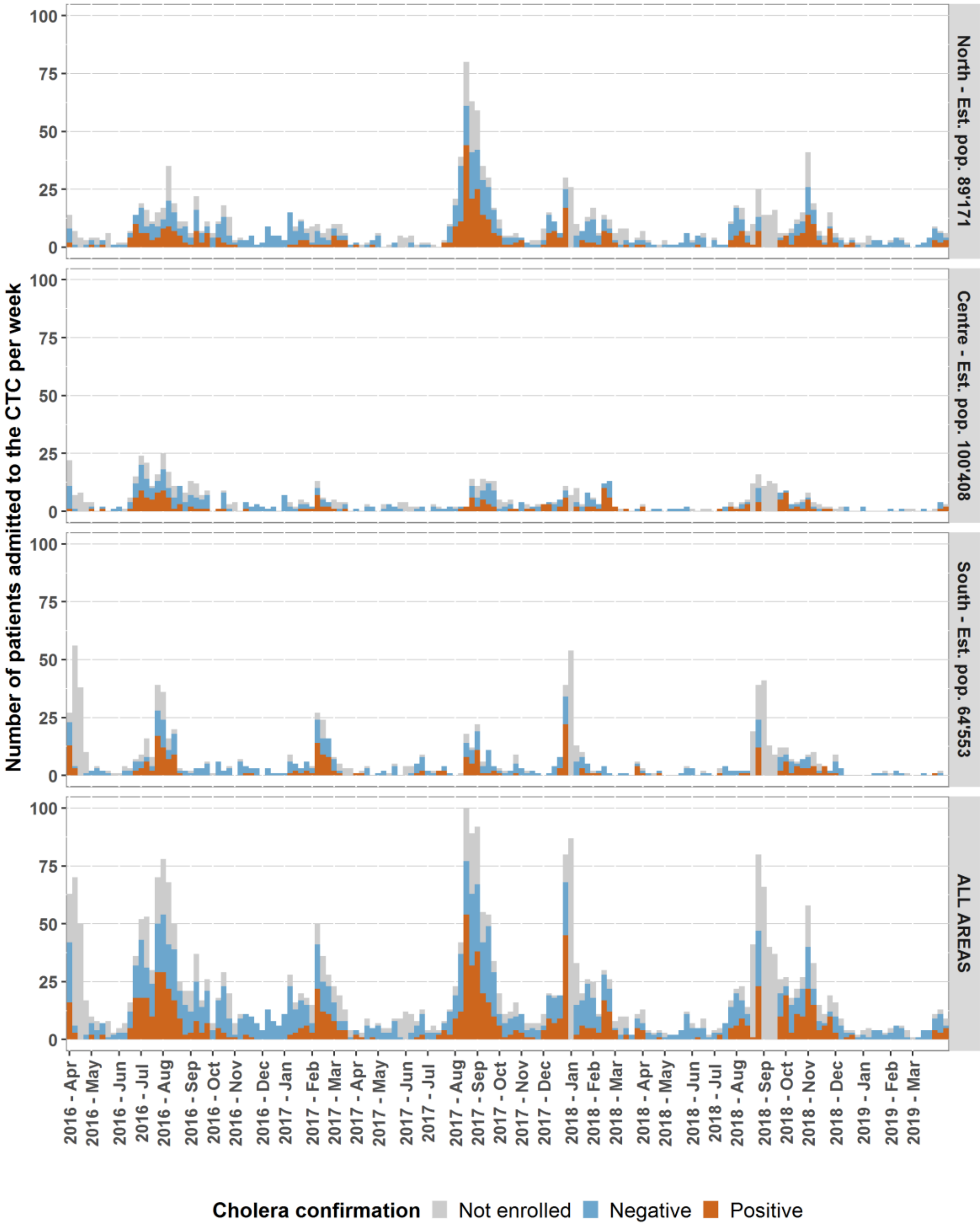


Figure 3-2 Distribution of admissions and RDT-confirmed cases over areas of Uvira

Table 3-4 CTC admissions and RDT-positive cholera incidence rates and endemicity by neighbourhood and area of residence for 2nd study period and overall

		Estimated population in November 2017	Number of admissions / patients enrolled / cholera confirmed	Cumulative CTC admissions incidence per 10'000	Estimated confirmation rate	Annual CTC admissions / confirmed cholera incidence* per 10'000	Distribution of weekly CTC admissions incidence per 10'000 over study period		Number of weeks with at least one admission	Number of weeks with patients enrolled / at least one confirmed case
							Median (IQR)	Maximum	n (% over study period)	
2nd study period (77 weeks)	AS Kasenga_Cepac	23210	188 / 126 / 48	81	42.2%	59.2 / 25	0.86 (0 - 1.29)	6.03	56 (73%)	47 / 24
	AS Kasenga_Etat	24697	127 / 91 / 43	51.4	45.5%	36.8 / 16.8	0.4 (0 - 0.81)	4.05	48 (62%)	41 / 27
	AS Kavimvira	13936	120 / 74 / 32	86.1	46.2%	65.1 / 30.1	0.72 (0 - 0.72)	13.63	51 (66%)	40 / 15
	AS Kilomoni	13231	95 / 40 / 12	71.8	38.4%	42.8 / 16.4	0 (0 - 1.51)	6.05	37 (48%)	24 / 6
	AS Kiyaya	14097	66 / 44 / 21	46.8	44.5%	33.3 / 14.8	0 (0 - 0.71)	5.67	31 (40%)	23 / 11
	NORTH	**89171**	**596 / 375 / 156**	**66.8**	**43.5%**	**47.4 / 20.6**	**0.67 (0.34 - 1.23)**	**4.6**	**74 (96%)**	**69 / 38**
	AS Mitumba	19276	74 / 45 / 22	38.4	38.0%	27.8 / 10.6	0 (0 - 0.52)	3.63	36 (47%)	25 / 13
	AS Mulongwe	13842	48 / 37 / 18	34.7	39.8%	25.1 / 10	0 (0 - 0.72)	2.89	32 (42%)	25 / 12
	AS Kimanga	14338	26 / 13 / 7	18.1	41.3%	12 / 5	0 (0 - 0)	2.09	17 (22%)	8 / 6
	AS Rombel	17190	32 / 26 / 15	18.6	38.0%	12.1 / 4.6	0 (0 - 0.58)	1.75	24 (31%)	18 / 13
	AS Saint_Paul	19566	60 / 28 / 12	30.7	40.7%	24.3 / 9.9	0 (0 - 0.51)	3.58	24 (31%)	16 / 10
	AS Tanganyika	16196	23 / 13 / 9	14.2	36.3%	9.2 / 3.3	0 (0 - 0)	1.23	19 (25%)	12 / 8
	CENTRE	**100408**	**263 / 162 / 83**	**26.2**	**39.0%**	**18.8 / 7.3**	**0.2 (0.1 - 0.4)**	**1.59**	**63 (82%)**	**49 / 32**
	AS Kabindula	15983	97 / 49 / 16	60.7	44.4%	41.3 / 18.3	0 (0 - 0.63)	11.89	32 (42%)	23 / 10
	AS Kalundu_Catholique	8522	43 / 25 / 12	50.5	36.6%	42.6 / 15.6	0 (0 - 1.17)	7.04	21 (27%)	15 / 9
	AS Kalundu_Cepac	11556	120 / 52 / 22	103.8	44.0%	71.5 / 31.5	0 (0 - 0.87)	18.17	31 (40%)	22 / 7
	AS Kalundu_Etat	14168	66 / 35 / 16	46.6	40.2%	35.7 / 14.4	0 (0 - 0.71)	4.94	31 (40%)	24 / 9
	AS Nyamianda	14324	77 / 47 / 17	53.8	41.6%	39.2 / 16.3	0 (0 - 0.7)	7.68	34 (44%)	24 / 8
	SOUTH	**64553**	**403 / 208 / 83**	**62.4**	**42.2%**	**45.2 / 19.1**	**0.31 (0.15 - 0.77)**	**8.37**	**60 (78%)**	**52 / 25**
	Total Uvira municipality	254132	1262 / 745 / 322	49.7	42.2%	35.5 / 15	0.39 (0.16 - 0.87)	3.42	76 (99%)	71 / 43
	Unknown residence location or outside Uvira		58 / 32 / 13							

** based on the mean number of admissions over 52 consecutive weeks and the confirmed/admission ratios estimated by the multivariable logistic regression model (total period) for each area or AS*

… continued next page

		Estimated population in November 2017	Number of admissions / patients enrolled / cholera confirmed	Cumulative CTC admissions incidence per 10'000	Estimated confirmation rate	Annual CTC admissions / confirmed cholera incidence* per 10'000	Distribution of weekly CTC admissions incidence per 10'000 over study period		Number of weeks with at least one admission	Number of weeks with patients enrolled / at least one confirmed case
							Median (IQR)	Maximum	n (% over study period)	
Both study periods (160 weeks)	AS Kasenga_Cepac	23210	429 / 284 / 108	184.8	42.2%	58.6 / 24.7	0.65 (0.32 - 1.29)	6.46	120 (75%)	93 / 49
	AS Kasenga_Etat	24697	324 / 242 / 109	131.2	45.5%	45.7 / 20.8	0.4 (0 - 1.21)	5.26	107 (67%)	94 / 57
	AS Kavimvira	13936	313 / 229 / 110	224.6	46.2%	81.9 / 37.8	0.72 (0 - 1.44)	24.4	102 (64%)	83 / 38
	AS Kilomoni	13231	259 / 157 / 50	195.8	38.4%	72.1 / 27.7	0.76 (0 - 1.51)	10.58	90 (56%)	64 / 26
	AS Kiyaya	14097	160 / 111 / 54	113.5	44.5%	40.8 / 18.1	0 (0 - 0.71)	12.77	68 (43%)	51 / 27
	NORTH	**89171**	**1485 / 1023 / 431**	**166.5**	**43.5%**	**57.8 / 25.1**	**0.67 (0.34 - 1.26)**	**8.97**	**155 (97%)**	**142 / 88**
	AS Mitumba	19276	211 / 135 / 54	109.5	38.0%	29.7 / 11.3	0.52 (0 - 1.04)	6.74	83 (52%)	67 / 32
	AS Mulongwe	13842	146 / 109 / 37	105.5	39.8%	31.7 / 12.6	0 (0 - 0.72)	5.06	70 (44%)	57 / 26
	AS Kimanga	14338	61 / 37 / 12	42.5	41.3%	13.7 / 5.7	0 (0 - 0)	2.79	39 (24%)	26 / 10
	AS Rombel	17190	85 / 61 / 29	49.4	38.0%	16.9 / 6.4	0 (0 - 0.58)	3.49	50 (31%)	39 / 25
	AS Saint_Paul	19566	151 / 85 / 32	77.2	40.7%	24.2 / 9.9	0 (0 - 0.51)	3.58	73 (46%)	50 / 22
	AS Tanganyika	16196	79 / 47 / 21	48.8	36.3%	16 / 5.8	0 (0 - 0.62)	2.47	55 (34%)	38 / 19
	CENTRE	**100408**	**733 / 474 / 185**	**73**	**39.0%**	**22.2 / 8.7**	**0.3 (0.1 - 0.6)**	**2.49**	**137 (86%)**	**117 / 69**
	AS Kabindula	15983	239 / 128 / 47	149.5	44.4%	40.4 / 17.9	0 (0 - 1.25)	12.51	70 (44%)	56 / 23
	AS Kalundu_Catholique	8522	93 / 66 / 26	109.1	36.6%	33.1 / 12.1	0 (0 - 1.17)	7.04	57 (36%)	46 / 21
	AS Kalundu_Cepac	11556	300 / 154 / 64	259.6	44.0%	75 / 33	0.87 (0 - 1.73)	20.77	82 (51%)	62 / 27
	AS Kalundu_Etat	14168	226 / 156 / 70	159.5	40.2%	52 / 20.9	0.71 (0 - 1.41)	7.76	84 (53%)	67 / 33
	AS Nyamianda	14324	149 / 94 / 31	104	41.6%	34.7 / 14.4	0 (0 - 0.7)	7.68	69 (43%)	53 / 19
	SOUTH	**64553**	**1007 / 598 / 238**	**156**	**42.2%**	**46.9 / 19.8**	**0.46 (0.15 - 1.08)**	**8.68**	**138 (86%)**	**121 / 62**
	Total Uvira municipality	254132	3225 / 2095 / 854	126.9	42.2%	41 / 17.3	0.45 (0.24 - 1.02)	3.93	159 (99%)	152 / 100
	Unknown residence location or outside Uvira		161 / 100 / 43							

** based on the mean number of admissions over 52 consecutive weeks and the confirmed/admission ratios estimated by the multivariable logistic regression model (total period) for each area or AS*

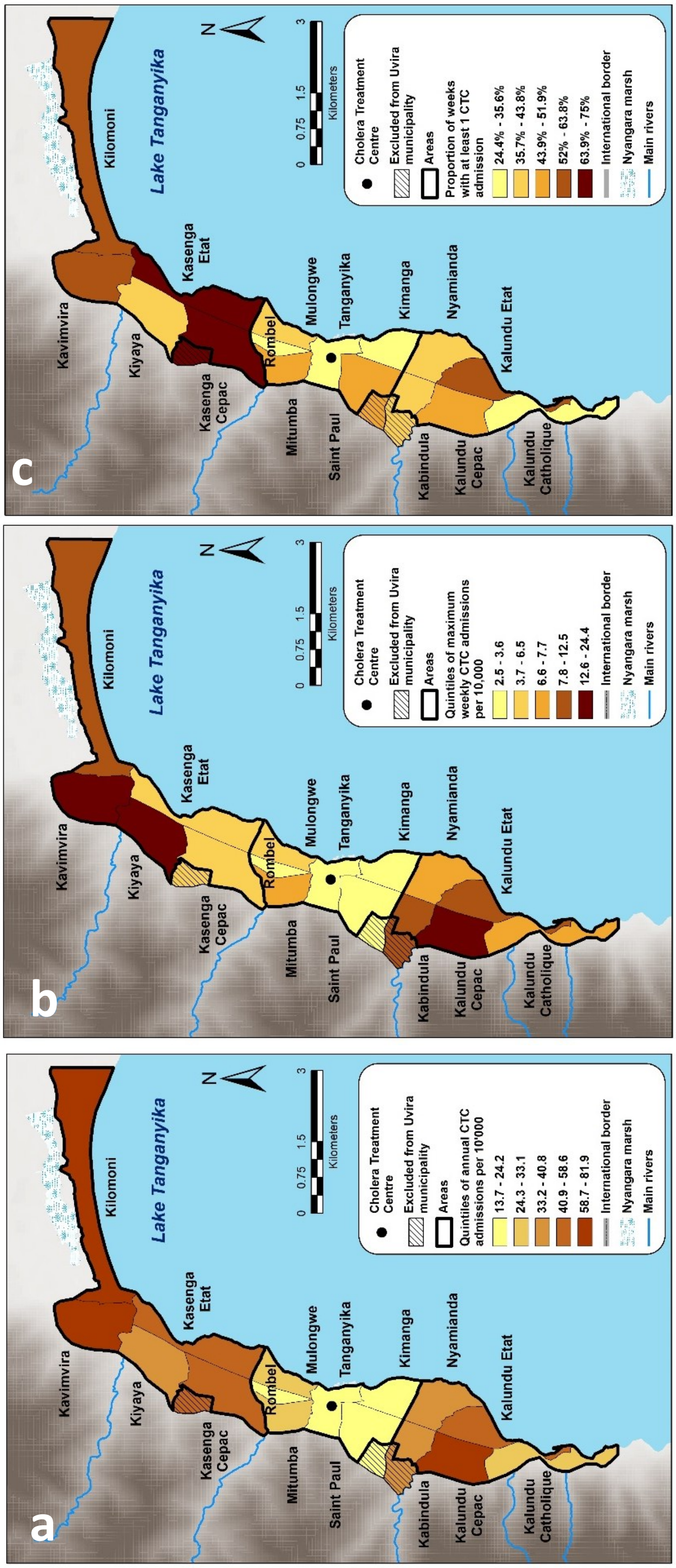


Figure 3-3 Incidence and endemicity by neighbourhood (AS) over combined study periods: (a) annual CTC admissions incidence per 10’000 (b) maximum weekly CTC admissions incidence per 10’000 (c) proportion of weeks with at least one CTC admission.

3.2.3 Prevalence and diversity of enteric infections amongst CTC patients: a summary

269 rectal swab samples collected from participants to the cholera RDT testing study between the 24th of September 2017 and the 9th of July 2018 were analysed by polymerase chain reaction (PCR) at LSHTM during the summer 2018 by Ms Camille Williams for her MSc Public Health for Development summer project and supervised by the PhD candidate A. Jeandron. The objectives of this analysis were to estimate the proportional distribution and describe the diversity of common enteric infections amongst suspected cholera cases.

The Luminex xTAG® GPP multiplex PCR kit was used on deoxyribonucleic acid (DNA) extracted from samples preserved on Whatman® FTA® Elute cards to detect simultaneously the presence of the following 15 enteropathogens: adenovirus, Campylobacter spp., *Clostridium difficile,* Cryptosporidium spp*, Entamoeba histolytica, Escherichia coli* O_{157}, Enterotoxigenic *Escherichia coli* (ETEC) LT/ST, Giardia spp., norovirus GI/GII, rotavirus A, Salmonella spp., shiga toxin-producing *Escherichia coli* (STEC), Shigella spp., toxigenic *Vibrio cholerae,* and *Yersinia enterocolitica.* The manufacturer issued a warning on the reliability of Giardia detection and those results were therefore discarded.

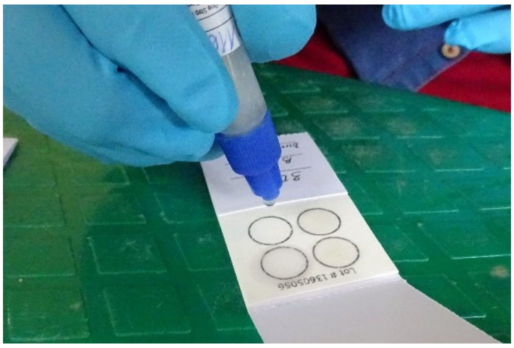


Figure 3-4 Preservation on Whatman® FTA® Elute cards of patient stool, sampled with a rectal swab and eluted in physiological water

Out of the 269 samples analysed, toxigenic *V. cholerae* was the most frequently detected pathogen (38 %), followed by ETEC (36 %), Cryptosporidium spp. (28 %), Campylobacter spp.(17 %) and Shigella spp. (16 %). *E. histolytica* and *Y. enterocolitica* were not detected in any samples. All other targeted pathogens - adenovirus*, C. difficile, E. coli* O157, norovirus GI/GII, rotavirus A, Salmonella and STEC - were detected in 10 participants or less. None of the targeted pathogens were detected in 44 (16.4 %) samples, whereas two or more pathogens were present in 45 % of them. 64 % of the samples positive for *V. cholerae* were also positive for one or more other pathogens. The mean number of pathogens per participant was 1.6, with patients under 15 years of age harboring approximately 1.5 times the number of pathogens detected in patients aged 16 or more.

Of the 269 participants, 84 (31 %) were positive for *V. cholerae* by both RDT and PCR; 149 (55 %) were negative for *V. cholerae* by both RDT and PCR, and 36 (13%) had discordant results. Results discordance could however be attributed to the different strains detected by RDT and PCR: RDTs target *V. cholerae* O_1 or O_{139} – independently of their toxigenicity – whilst Luminex xTag® GPP targets any toxigenic *V. cholerae*, including non-O_1 non O_{139}.

More details on this study are available in the article published in the journal BMC Infectious Diseases **[5]**.

3.2.4 Discussion

The results obtained by analysing admissions and RDT-confirmation data over a second study period are relatively consistent with the findings of the initial analysis. Annual incidence rates of suspected / RDT-positive cases observed in Uvira remain of similar magnitude to those reported elsewhere in cholera affected areas of Sub-Saharan Africa **[6-9]**. Cholera endemicity in Uvira is confirmed again, with at least one RDT-confirmed case identified in 50 % of the weeks during which at least one patient was enrolled.

Individual characteristics of patients and their stay at the CTC were consistently associated with their RDT-confirmation results over both study periods. Boys under 2 are five times less likely to test positive than adult women, while girls aged 5 to 15 years and adult men are 1.5 times more likely to be confirmed. Staying longer than two nights in the CTC and being transferred to another hospital department are also associated with higher rates of cholera confirmation by RDT.

Contextual characteristics of admissions appear less robustly associated with RDT-confirmation during the 2nd period. Although RDT-confirmation remains strongly predicted by the weekly admissions incidence category, evidence of a difference in confirmation rates across residence location weakens when considering both study periods. Even more strikingly, the association between dry/rainy weeks and confirmation rate is inverted between the two study periods, and overall, patients admitted during dry weeks are 20 % less likely to test positive than those admitted during rainy weeks. This is more in line with the published literature but differs from the results from the initial analysis, that are robust to a change in rain estimates data source **[8,10-13]**.

Higher RDT-confirmation rates during high and very high weekly admissions incidence weeks are consistent with the well-recognised outbreak potential of cholera. However, the variability observed between the two study periods in other time (rain) and space (residence) predictors of RDT-confirmation rates highlights the importance of time and space resolution of data analysis. Results from the analysis of a single incidence peak in CTC admissions may not be generalisable over longer time periods and inversely, trends observed over long time periods may not be relevant to single events.

The diversity of enteric pathogens detected by PCR on preserved stool samples from a subset of participants to the RDT-testing study confirms the range of entero-pathogens concomitantly infecting patients presenting to the CTC with acute diarrhoea. It is unsurprising considering that they all transmit via faecal-oral routes, undoubtedly encouraged by the environmental and behavioural risk factors present in Uvira.

Two thirds of those testing positive for cholera with RDT or by PCR were also infected by one or more other diarrhoea causing pathogens. No specific sign or symptom can unequivocally distinguish cholera from other infectious causes of severe watery diarrhoea. Milder forms of cholera disease especially are difficult to distinguish from infectious diarrhoea caused by other pathogens, especially those responsible mostly for non-inflammatory syndromes – no fever, abdominal pains or bloody stools – such as ETEC or Cryptosporidium in adults and rotavirus in children **[14-17]**. Our results therefore question the suitability of systematically attributing the aetiology of acute diarrhoea in CTC patients who test positive to cholera RDTs.

Knowledge of the exact aetiology of acute infectious diarrhoea is largely unnecessary to provide adequate treatment – our results however emphasize the need for sufficient infection prevention and control (IPC) measures at the CTC and other hospital departments to avoid various enteric infections spreading between patients. Our results also support a comprehensive approach to the prevention of diarrhoeal diseases in Uvira with a wide range of water, sanitation and hygiene improvements, along with other cholera specific prevention and management measures.

3.3 Chapter 3 - References

1. Jeandron A, Cumming O, Rumedeka BB, Saidi JM, Cousens S. *Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo.* **PLoS One**. 2018;13(8):e0201306.
2. GPM IMERG Late Precipitation L3 1 day 0.1 degree x 0.1 degree V04. Goddard Earth Sciences Data and Information Services Center (GES DISC); 2016. https://disc.gsfc.nasa.gov/datacollection/GPM_3IMERGDL_04.html. Accessed 12/12/2017.
3. Dinku T, Funk C, Peterson P, et al. *Validation of the CHIRPS satellite rainfall estimates over eastern Africa.* **Quarterly Journal of the Royal Meteorological Society**. 2018;144(S1):292-312.

4. Funk CC, Peterson PJ, Landsfeld MF, et al. *A quasi-global precipitation time series for drought monitoring*. **US Geological Survey data series**. 2014;832(4):1-12.

5. Williams C, Cumming O, Grignard L, et al. *Prevalence and diversity of enteric pathogens among cholera treatment centre patients with acute diarrhea in Uvira, Democratic Republic of Congo*. **BMC Infect Dis**. 2020;20(1):741.

6. Bwire G, Ali M, Sack DA, et al. *Identifying cholera "hotspots" in Uganda: An analysis of cholera surveillance data from 2011 to 2016*. **PLoS Negl Trop Dis**. 2017;11(12):e0006118.

7. Gujral L, Sema C, Rebaudet S, et al. *Cholera epidemiology in Mozambique using national surveillance data*. **J Infect Dis**. 2013;208 Suppl 1:S107-114.

8. Mengel MA, Delrieu I, Heyerdahl L, Gessner BD. *Cholera outbreaks in Africa*. **Curr Top Microbiol Immunol**. 2014;379:117-144.

9. Sauvageot D, Njanpop-Lafourcade BM, Akilimali L, et al. *Cholera Incidence and Mortality in Sub-Saharan African Sites during Multi-country Surveillance*. **PLoS Negl Trop Dis**. 2016;10(5):e0004679.

10. Bompangue D, Giraudoux P, Handschumacher P, et al. *Lakes as source of cholera outbreaks, Democratic Republic of Congo*. **Emerg Infect Dis**. 2008;14(5):798-800.

11. Bompangue D, Giraudoux P, Piarroux M, et al. *Cholera epidemics, war and disasters around Goma and Lake Kivu: an eight-year survey*. **PLoS Negl Trop Dis**. 2009;3(5):e436.

12. Emch M, Feldacker C, Islam MS, Ali M. *Seasonality of cholera from 1974 to 2005: a review of global patterns*. **Int J Health Geogr**. 2008;7:31.

13. Rebaudet S, Sudre B, Faucher B, Piarroux R. *Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes*. **J Infect Dis**. 2013;208 Suppl 1:S46-54.

14. LaRocque R, Harris JB. Cholera: Clinical features, diagnosis, treatment, and prevention. 2018; https://www.uptodate.com/contents/cholera-clinical-features-diagnosis-treatment-and-prevention#H938291586. Accessed 23/03/2020, 2020.

15. Learoyd TP, Gaut RM. *Cholera: under diagnosis and differentiation from other diarrhoeal diseases*. **Journal of travel medicine**. 2018;25(suppl_1):S46-s51.

16. Barr W, Smith A. *Acute diarrhea*. **American family physician**. 2014;89(3):180-189.

17. O'Ryan G M, Ashkenazi-Hoffnung L, O'Ryan-Soriano MA, Ashkenazi S. *Management of acute infectious diarrhea for children living in resource-limited settings*. **Expert Review of Anti-infective Therapy**. 2014;12(5):621-632.

4

Tap water service and households' water-related practices

Originally planned to establish and follow a cohort of dwellings and measure any change in water-related practices reported by households during tap water supply improvements, two household surveys were conducted in Uvira in October 2016 and November 2017. This chapter summarises the main results of these surveys pertaining to water source use, perceived quality and reliability of used sources, time spent to collect water, quantity of water used for domestic activities and quality of drinking water. These results were then used in **Research Paper 2: Predicting quality and quantity of water used by urban households based on tap water service**.

4.1 Introduction

Cholera and other diarrhoeal diseases are transmitted along faecal-oral routes, most of which are influenced by water-related practices in households, pertaining to both microbial quality of drinking water, and water quantity used for hygiene purposes. As improvements in tap water supply were being planned for Uvira, pathways for change were proposed for the impact this intervention may have on cholera and other diarrhoeal diseases (**Figure 4-1**). Improved access to tap water through private and public tap connections is expected to increase use of safe water for drinking and reduce the use of and contact with unsafe surface water **[1,2]**. Increased proximity of tap water is expected to reduce the time spent to collect water and increase the amount of water used at home, especially for hygiene practices **[3-5]**. Improved reliability of water supply at the taps is expected to further encourage sole use of tap water for drinking and reduce occasional use of unsafe water **[6,7]**, and also to increase the concentration of residual chlorine at the tap and limit the risk of pathogen ingress into the supply network at low water pressure, thereby reducing the risk of drinking water being contaminated with diarrhoeal pathogens **[8-10]**. The magnitude of change was also expected to depend on the level of tap water access and use at baseline.

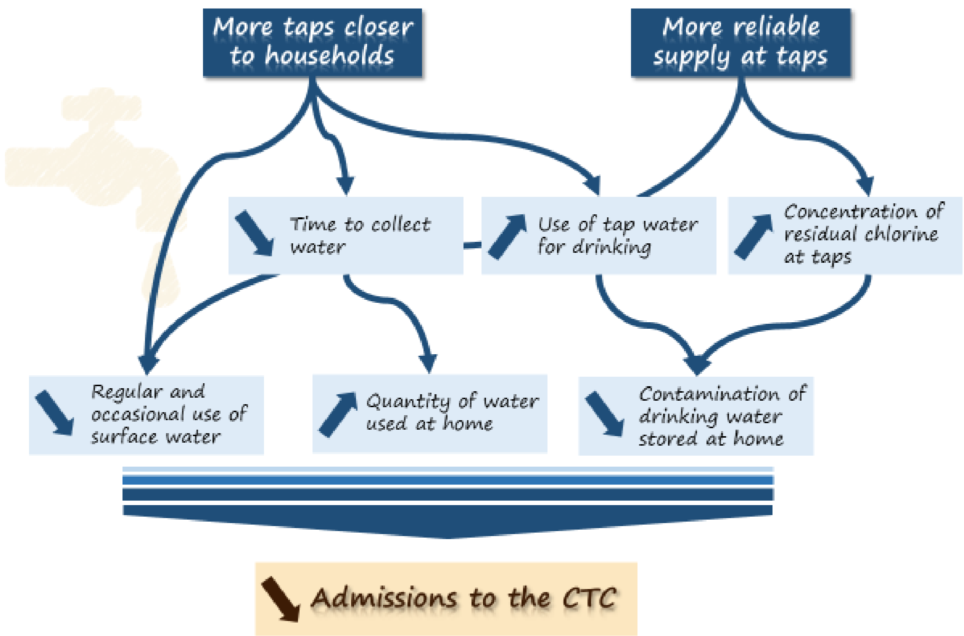


Figure 4-1 Proposed change pathways for the impact of tap water supply improvements over households' water-related practices and admissions to the CTC in Uvira

In addition to evaluating the impact of the improvements directly on admissions to the Cholera Treatment Centre (CTC), it was therefore decided to establish a cohort of approximately 500 dwellings in Uvira, sampled across various levels of tap water access and use, and to interview resident households on relevant practices: 1) water sources used for drinking and domestic activities; 2) perceived quality and reliability of used sources; 3) time spent to collect water at the main source; 4) quantity of water used for domestic activities and 5) microbial quality of drinking water.

4.2 Methods

The study setting is described in detail in chapter *Data and methods: an overview*, section *2.1*.

4.2.1 Households sampling and interviewing

In October 2016, comprehensive data on the location and functionality of Regideso taps were unavailable to use for household sampling. Therefore, to establish a cohort of households representing a wide range of access to tap water in Uvira, and in absence of a household sampling frame, a two-stage random spatial sampling method was used based on a piped water availability index.

The existing tap water distribution network was mapped with ArcGIS 10.3 (ESRI, Aylesbury, UK) to represent the pipe network including pipe diameter, the position of the 3,016 functional tap connections as geographically referenced in September 2012 by the Regideso, and the position of the water reservoir (**Figure 4-2**).

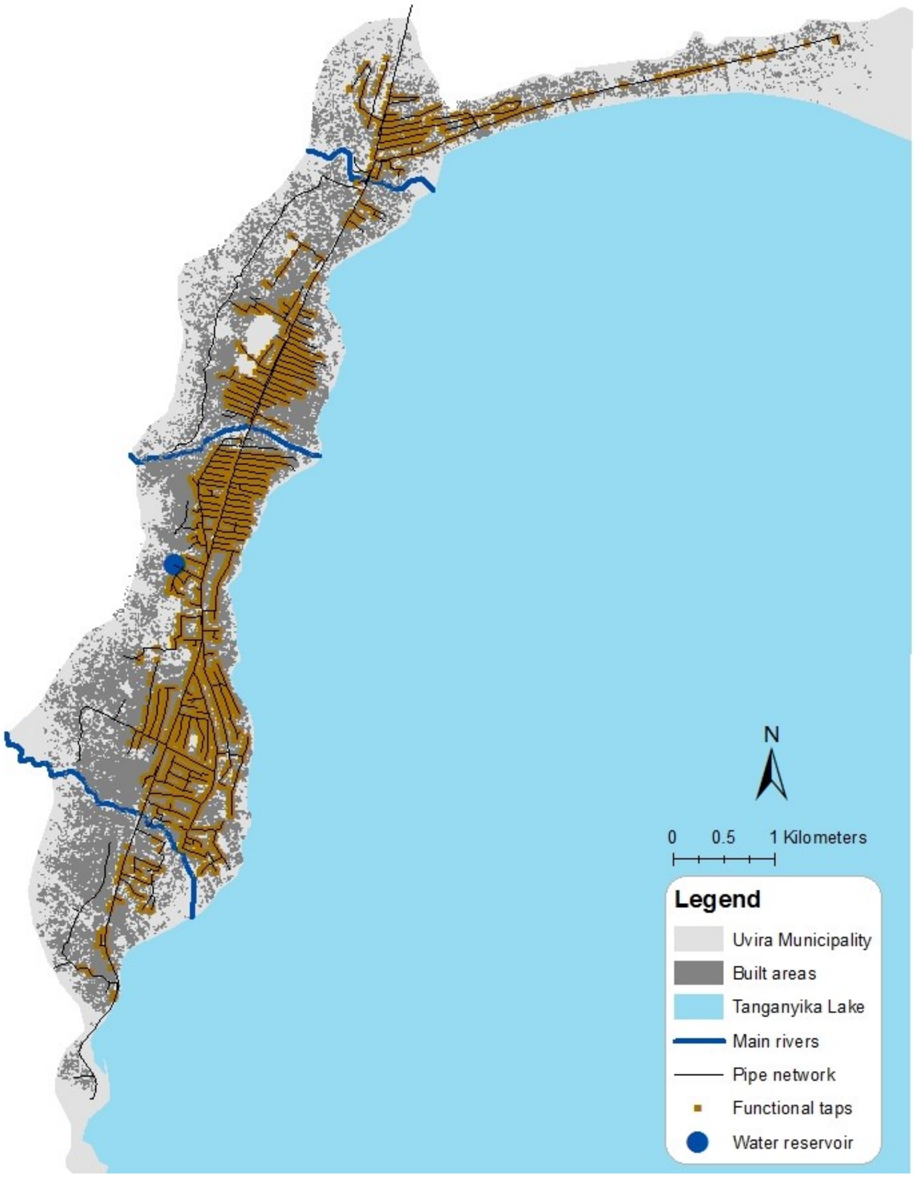


Figure 4-2 Map of Uvira showing the tap water supply network and functional taps in September 2012, used for sampling frame construction

To estimate water availability at each tap, it was assumed that this was inversely related to the friction head loss in the distribution network, itself mostly dependent on the ratio of pipe length: pipe diameter (L:D) between the water reservoir and each tap. An inverse friction head loss for each functional tap on the network was derived and ranged between approximately 0.01 and 10 m^{-1}. In order to construct an indicator of tap water availability for all built areas, the ArcGIS kernel density function was used. In brief, this function fits the inverse friction head loss of each functional taps to a smoothly tapered surface limited by the search radius with a quadratic function, with the highest value at the tap location and lowest at the search radius limit. This function therefore calculates the magnitude of tap water availability per km2 within a search area, with a spatial resolution of 5 m x 5 m units. The values were then classified into 5 quantiles of equal surface area. In order to avoid a threshold effect, the search area was set to range from 250 m to 750 m by 50 m increments, with areas classified into 5 quantiles for each radius, and only areas consistently classified within the same stratum of tap water availability per km^2 were included in the spatial sampling (**Figure 4-3**). Those areas included 54.1 % of the estimated total number of buildings in Uvira.

Assuming that 10 % of the georeferenced locations would not lead to an eligible household, and that 10 % of eligible households would not be willing to participate, 125 random locations were generated in each stratum for a target sample size of 100 households enrolled per stratum. In an adapted version of the "Mecca" method described by Himelein et al **[11]**, interviewers attempted to reach all sampled locations with the help of a Geographical Positioning System (GPS) handheld device, with locations considered ineligible if they could not be accessed within 20 m without trespassing on a non-authorized area (for example a military camp or school compound) or due to physical terrain (river, ditch, lake, steep slope, dense vegetation). Once at the location they faced north and walked in that direction until the first building. If no building could be seen or accessed in that direction, they were instructed to turn to the closest building. The building was considered eligible if it was inhabited on the day of the 1st visit by one or more households, either by the actual presence of its inhabitants or as reported by neighbours. If the building was ineligible, team members were instructed to proceed to the next building in the initially chosen direction until they found an eligible building with a limit of three ineligible buildings. In the case of a multi-household dwelling, a list of households was drawn up and one of them randomly selected from the list with a random number table.

Any member of the household aged 18 or more was considered eligible and selected households were revisited up to three times over the survey period until an adult was available for an interview. After explaining the purpose of the study and answering any questions, written informed consent was obtained from an adult member of the household willing to participate. If non-literate, another adult from the household or the neighbourhood witnessed the explanations and consent. Enrolment and interviews were conducted in Swahili by a team of trained interviewers and supervisors. The questionnaire (in French), a summary of questionnaire variables in English and consent form are available in **Appendix 4-1a, Appendix 4-1b** and **Appendix 2-6** respectively.

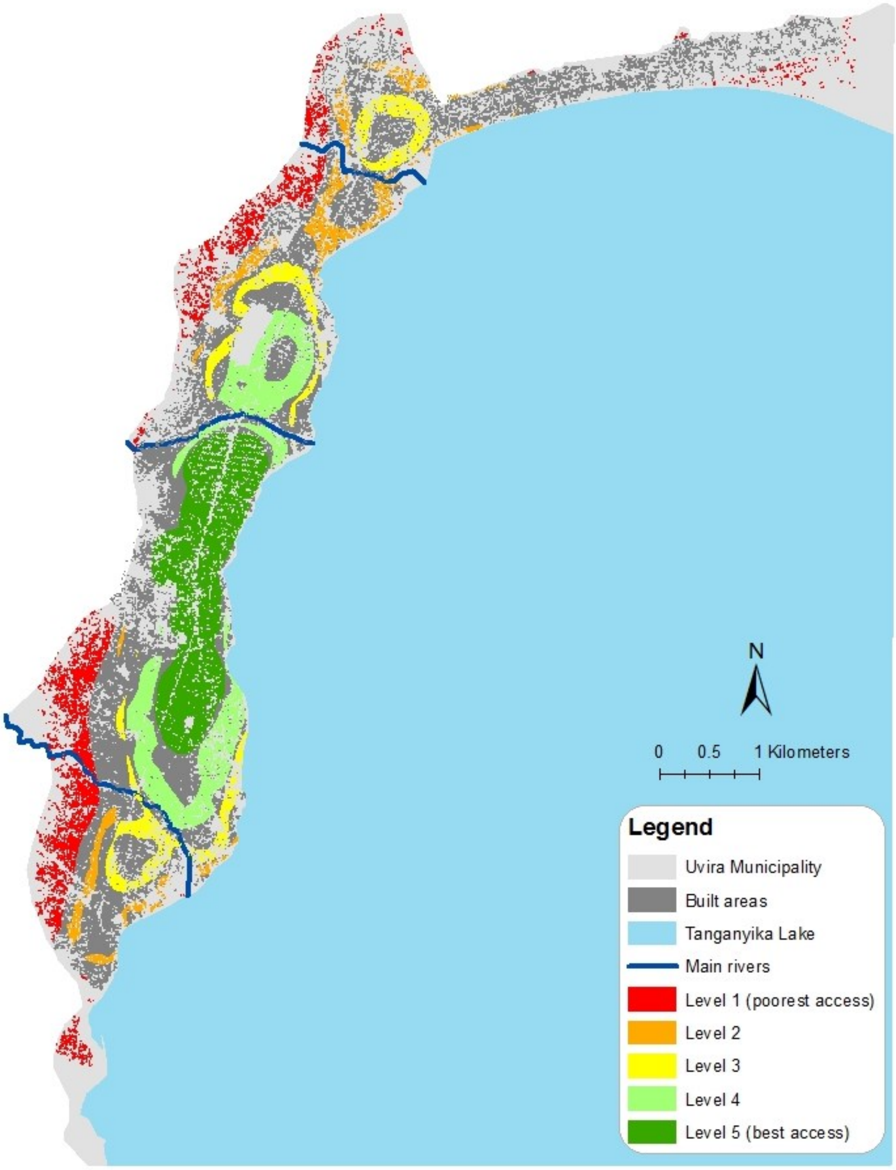


Figure 4-3 Map of Uvira showing the five quintiles of tap water access defined for geographical sampling of households

During the 2nd survey implemented in 2017, the dwellings sampled and georeferenced during the 1st survey were revisited. If the household inhabiting the building was different from the one interviewed in the first survey round, the same enrolment process was used with the new family. If the building was uninhabited, destroyed or the household did not wish to be re-interviewed, the next building directly to the right was visited.

For the 1st survey, data were collected every day between the 10th and the 25th of October 2016 after piloting and pre-testing of the questionnaire in 8 households. The 2nd survey took place between the 25th of October and the 11th of November 2017. A team of 10 interviewers and 3 supervisors was recruited and trained for the 1st survey, and 7 interviewers and two supervisors amongst them were available to perform the 2^{nd} survey after re-training.

For both surveys, geographical data were collected on hand-held GPS devices while interview data were collected on hand-held tablets or phones using the Open Data Kit (ODK) system.

Interviews for the 2016 survey were divided into 13 sections, including basic socio-economic and demographics information, water use the previous day, main and alternative sources of water used the preceding 2 weeks, water source accessibility and cost, and water treatment practices. In 2017, as less time and resources were available for the survey whilst security concerns were greater, the questionnaire was shortened and omitted detailed questions on alternative water sources and the cost of water.

4.2.2 Drinking and surface water samples collection and analysis

Detailed methods for drinking and surface water sampling and analysis are described in chapter *Data and methods: an overview*, section 2.6. Drinking water quality data collected during the 2016 survey suffered major limitations due to reagent and equipment failures as well as an analysis protocol that did not account for high turbidity (see chapter *Data and methods: an overview,* section *2.6.1*). Hence, only microbiological quality results from the 2017 survey were considered in the present chapter.

4.2.3 Data preparation

Results reported in this chapter focus on water sources used for drinking and domestic activities, perceived quality and reliability of used sources, time spent to collect water at the main source, quantity of water used for domestic activities and microbial quality of drinking water. Data from the 2016 survey were analysed whenever available and reliable. For excluded variables however (drinking water quality and time spent to collect water), data from the 2017 survey were used instead, assuming that no significant changes in households' water-related practices would have occurred between 2016 and 2017. For each analysis, dependent and independent variables used were from the same survey round, in a cross-sectional approach.

Time spent to collect water

Due to an error in questionnaire programming on ODK Collect, only a subset of households was asked about the time spent to collect water during the 2016 survey. Results reported in the present chapter are therefore those obtained from data collected during the 2017 survey.

When households reported that their main source of water was located outside their compound, households were asked the outbound and inbound times it took to travel to collect the water. Typical waiting time at the source (e.g. for a tap outside the compound) was also recorded. For main sources within the compound, return travel time to collect water was set to 1 minute with no waiting time. Surface water sources were assumed to have no waiting time either.

Quantity of water used at home for domestic activities

Households were interviewed about the amount of fresh water (as in not recycled from a previous activity) used at home the day preceding the interview for various domestic activities using a visual aid depicting commonly used water containers and utensils with their respective volumes (**Appendix 4-2**). The amounts reported for adult bathing, child bathing, laundry, dishwashing, food and produce rinsing, dwelling cleaning, handwashing and drinking were added up into a total amount of fresh water used at home for domestic activities. Water used at home to prepare food or items for sale, such as home brewed drinks or mud bricks, or to render a paid-for service, such as motorbike washing or laundry, were excluded from the total. Households also reported if they had performed any of the above activities directly at the point of collection, or if they had re-used water for another activity.

In order to be able to relate quantity of water used at home and time spent to collect water, results in the present chapter are based on the 2017 survey data.

Drinking and surface water microbiological contamination

Microbiological contamination was examined using the number of colony forming units (CFU) *Escherichia coli* per 100 ml obtained after analysis of each sample, using categories related to diarrhoeal disease risk as defined by WHO guidelines: safe (CFU<=0), low risk (CFU between 1 and 10 per 100 ml), intermediate risk (CFU between 11 and 100 per 100 ml), high risk (CFU between 101 and 1,000 per 100 ml) and very high risk (CFU >1,001 per 100 ml) **[12]**. Petri plates with "Too numerous to count" (TNTC) CFU were considered to have 151 CFU, as 150 colonies is the most conservative limit of countable colonies per plate usually taken in microbiology **[13]**. This number was then multiplied by the appropriate factor depending on sample dilution to obtain a CFU per 100 ml. Data for surface water was log10 transformed before analysis.

Wealth index

A total of 13 binary variables and 4 ordinal variables were used to derive a wealth index for 518 households enrolled during either 2016 or 2017 surveys. These variables included durable assets owned, sources of lighting, fuel for cooking, dwelling ownership, size and materials, and sanitation access and use. For sources of lighting, fuel for cooking, dwelling materials and sanitation use, multiple answers were allowed. In the case of multiple responses, the response expected a priori to be associated with a higher wealth status was retained.

The wealth index was constructed by means of polychoric principal component analysis (PCA). The first component of polychoric PCA explained 50.4 % of the variance with an eigenvalue of 8.57. Scores ranged from -2.17 to 0.34, with a distribution skewed towards lower scores. For further analysis, scores were normalised and unskewed using a log function with new _score = log(score+2.59). The new scores ranged from -0.86 to 0.13. The distribution of each variable and their contribution to the wealth index scores are summarized in **Appendix 4-3**. For the present analysis, wealth index was classified into quintiles.

4.2.4 Statistical methods

Data analyses were performed in the R environment **[14]**. Continuous variables were described using median and interquartile range (IQR) as measures of central tendency and spread. Amount of water used at home for domestic activity was modelled using linear regression, and outliers with high influence on model parameters (Cook's distance greater than three times mean Cook's distance) were identified and removed (n=4). Odds of contamination of stored drinking water were examined using logistic regression.

In both regression models above, households' wealth quintile was considered as an a priori confounder of any causal effect of tap water service on drinking water quality and water quantity used.

The spatial sampling strategy for dwellings used for the surveys was not designed to generate representative figures for the entire Uvira population, but rather aimed at describing households' water-related practices at a wide range of tap water service level. We however calculated approximate post-sampling weights for each survey in order to be able to provide estimates of water-related practices and risk factors more representative of the entire population than unweighted results from interviewed households only. Weights were constructed on the basis of the number of dwellings from which households were enrolled out of the total estimated number of buildings, for each tap water service level (**Table 4-1**).

Table 4-1 Post-sampling weights for 2016 and 2017 surveys, based on tap water service level

Tap water service level	Estimated number of buildings	Number of buildings included in survey 1	Number of buildings included in survey 2	Weights survey 1	Weights survey 2
1 (worst)	4452	83	78	0.00199	0.00212
2	3128	95	87	0.00122	0.00133
3	3852	92	88	0.00155	0.00162
4	6208	95	91	0.00242	0.00253
5 (best)	9342	106	102	0.00327	0.00339
All levels	26982	471	446		

4.3 Results

4.3.1 Characteristics of interviewed households

Dwellings location

471 households were enrolled and interviewed during the 1st survey in 2016. 397 (82.3 %) of these households were re-interviewed in 2017, and 49 new households were enrolled during the 2nd survey, for a total of 446 interviews. A total of 520 unique households were interviewed, during one or both surveys, in 481 dwellings (**Figure 4-4**).

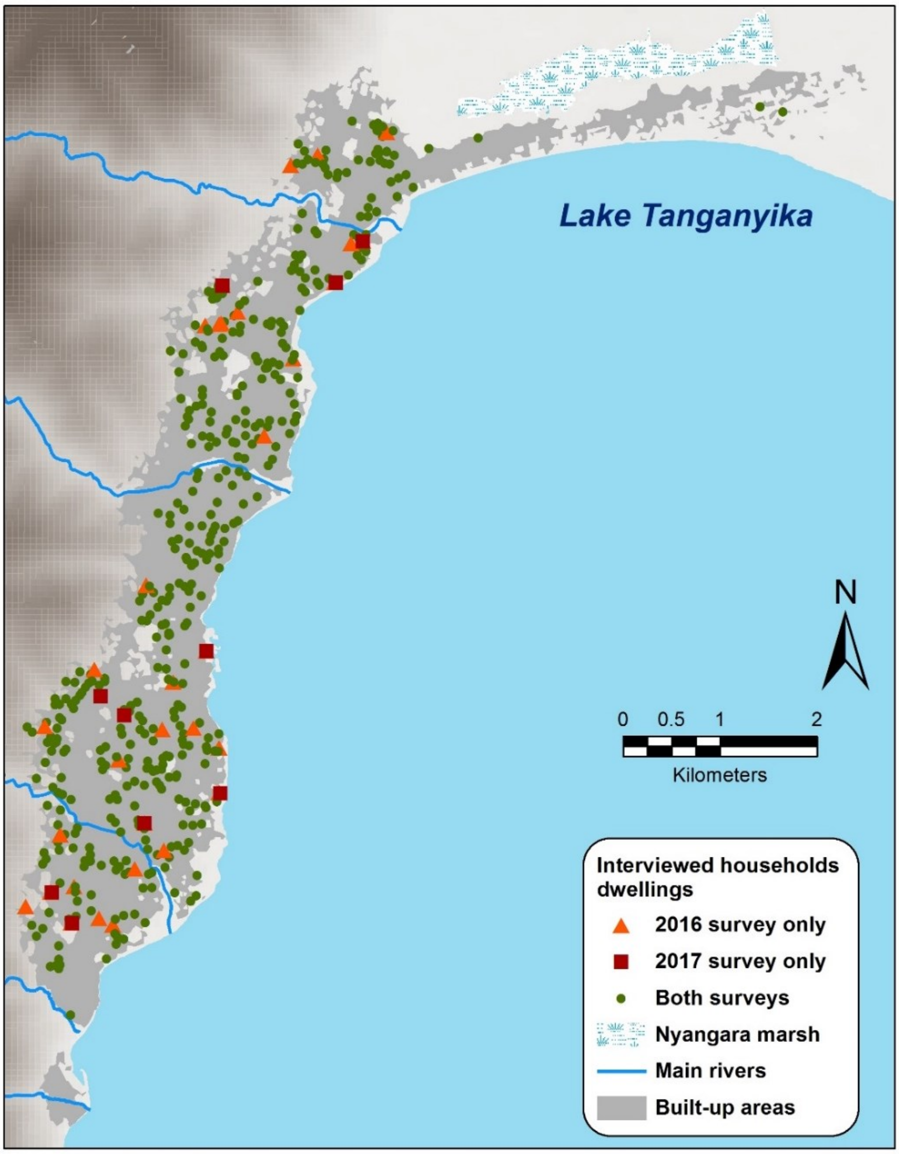


Figure 4-4 Location of interviewed households' dwellings for the 2016 and 2017 surveys

Tap water service level and wealth

Distribution across tap water service levels and wealth was very similar for both 2016 and 2017 datasets, with the highest and lowest levels of tap water service slightly over- and under-represented respectively. Expectedly, tap water service level and wealth index show a degree of correlation, with a Spearman's correlation coefficient of 0.50 over 520 unique households (**Table 4-2**).

Table 4-2 Distribution of households across tap water service levels and wealth quintiles for 520 unique households enrolled in either 2016 or 2017

	Wealth quintile						
	1 (poorest)	2	3	4	5 (wealthiest)	*missing*	All (100%)
Tap water service level							
1 (worst)	35 (36.8%)	33 (34.7%)	15 (15.8%)	9 (9.5%)	2 (2.1%)	*1 (1.1%)*	95
2	34 (31.5%)	26 (24.1%)	26 (24.1%)	13 (12%)	7 (6.5%)	*2 (1.9%)*	108
3	17 (17.3%)	22 (22.4%)	23 (23.5%)	20 (20.4%)	16 (16.3%)	*0 (0%)*	98
4	16 (15.2%)	12 (11.4%)	21 (20%)	22 (21%)	33 (31.4%)	*1 (1%)*	105
5 (best)	8 (7%)	7 (6.1%)	18 (15.8%)	35 (30.7%)	43 (37.7%)	*3 (2.6%)*	114

Households demographic composition and compound sharing

A typical household was composed of 7 members: one baby under 2 years of age (interquartile range IQR 0 – 1; range 0 – 4), one child between 2 and 5 (IQR 0 – 2; range 0 – 6), two children between 5 and 15 (IQR 1 - 3; range 0 – 9), and three adults above 15 (IQR 2 – 5; range 1 – 12). Nearly two thirds of the households interviewed did not share their compound with other households, while about 5 % reported sharing their compound with more than 4 other households.

4.3.2 Domestic and drinking water sources

All water sources reported for all domestic usages (2016 survey)

The 471 households interviewed reported having used an average of 1.8 sources per household in the two weeks preceding the survey (**Table 4-3**). The sources of water the most widely reported – as main or alternative sources - for all domestic use (including drinking) are the rivers, taps outside the compound, and Lake Tanganyika, used respectively by 62 %, 52.9 % and 32.9 % of interviewed households. A majority (65 %) of interviewed households reported having used two or more different sources of water in the two previous weeks. Nearly two thirds (62 %) of the interviewed households had used river water as main or alternative source in the previous two weeks.

Households were increasingly more likely to have used tap water as a source as the tap water service in their area improved: 95.3 % in best served locations vs 60.2 % in worst served areas. Conversely, households were increasingly likely to have used non-tap sources as their service worsened, although the trend is less striking (67 % of households in best served locations vs 81.7 % overall) (**Table 4-3**).

Table 4-3 Number and type of sources reported to have been used to collect water for all domestic use during the two weeks preceding interview, by household (survey 2016)

		Tap water service quintile n (% of households)				
	All	1 (worst)	2	3	4	5 (best)
Number of households	**471**	83	95	92	95	106
Number of households having used						
Lake Tanganyika	**155 (32.9%)**	3 (3.6%)	35 (36.8%)	41 (44.6%)	29 (30.5%)	47 (44.3%)
Rivers and streams	**292 (62%)**	64 (77.1%)	74 (77.9%)	54 (58.7%)	61 (64.2%)	39 (36.8%)
Tap outside of the compound	**249 (52.9%)**	47 (56.6%)	42 (44.2%)	60 (65.2%)	52 (54.7%)	48 (45.3%)
Tap in the compound	**117 (24.8%)**	4 (4.8%)	14 (14.7%)	13 (14.1%)	29 (30.5%)	57 (53.8%)
Rainwater	**3 (0.6%)**	1 (1.2%)	1 (1.1%)	0 (0%)	0 (0%)	1 (0.9%)
Leaking water supply pipe	**7 (1.5%)**	1 (1.2%)	0 (0%)	0 (0%)	6 (6.3%)	0 (0%)
Reported number of sources used						
1 source only	**164 (34.8%)**	43 (51.8%)	31 (32.6%)	26 (28.3%)	26 (27.4%)	38 (35.8%)
2 sources only	**226 (48%)**	28 (33.7%)	43 (45.3%)	55 (59.8%)	50 (52.6%)	50 (47.2%)
3 sources or more	**81 (17.2%)**	12 (14.5%)	21 (22.1%)	11 (12%)	19 (20%)	18 (17%)
Mean number of sources reported per household						
	1.82	1.63	1.89	1.84	1.93	1.81
Number of households having used non-tap source						
	385 (81.7%)	67 (80.7%)	89 (93.7%)	80 (87%)	78 (82.1%)	71 (67%)
Number of households having used tap water						
	360 (76.4%)	50 (60.2%)	56 (58.9%)	73 (79.3%)	80 (84.2%)	101 (95.3%)

Main and alternative sources (2016 survey)

The most frequently reported main water source in quantity for all domestic use was tap outside the compound (35.7 %) followed by rivers (32.7 %), tap inside the compound (19.3 %) and Lake Tanganyika (11.9 %) (**Table 4-4**).

78 % of the households having reported using a tap in their compound in the previous two weeks were using it as their main source. Similarly, 67 % of the households having reported using a tap outside of the compound were using it as their main source. Lake water is cited as main source of water by only 36 % of the households reporting using lake water, i.e. lake water appears to be most often used as an alternative to other sources.

The main source had been used by a large majority (84 %) on the day or preceding day of the interview, while the alternative sources appeared to be used in a less regular way. The main and alternative sources had mostly been adopted when the household moved to the dwelling, ie most often more than two years before the interview.

Table 4-4 Type of source reported to have been used as main (in quantity) source to collect water during the two weeks preceding interview, by household (survey 2016)

		Tap water service quintile n (% of households)				
	All	1 (worst)	2	3	4	5 (best)
Number of households	**471**	83	95	92	95	106
Number of households reporting as main source						
Lake Tanganyika	**56 (11.9%)**	2 (2.4%)	20 (21.1%)	20 (21.7%)	8 (8.4%)	6 (5.7%)
amongst users	***36.1%***	*66.7%*	*57.1%*	*48.8%*	*27.6%*	*12.8%*
Rivers and streams	**154 (32.7%)**	46 (55.4%)	50 (52.6%)	33 (35.9%)	20 (21.1%)	5 (4.7%)
amongst users	***52.7%***	*71.9%*	*67.6%*	*61.1%*	*32.8%*	*12.8%*
Tap outside the compound	**168 (35.7%)**	33 (39.8%)	24 (25.3%)	27 (29.3%)	41 (43.2%)	43 (40.6%)
amongst users	***67.5%***	*70.2%*	*57.1%*	*45%*	*78.8%*	*89.6%*
Tap in the compound	**91 (19.3%)**	2 (2.4%)	1 (1.1%)	12 (13%)	24 (25.3%)	52 (49.1%)
amongst users	***77.8%***	*50%*	*7.1%*	*92.3%*	*82.8%*	*91.2%*
Leaking water supply pipe	**2 (0.4%)**	0 (0%)	0 (0%)	0 (0%)	2 (2.1%)	0 (0%)
amongst users	***28.6%***	*0%*			*33.3%*	

Figure 4-5 shows the combinations of main and alternative sources across households as an alluvial diagram that represents the links in types of water sources used by individual households as main, 1st alternative and 2nd alternative water sources. .Nearly all households using a tap inside or outside their compound as main source report using surface water sources as alternative sources. Taps inside or outside the compound are however alternative sources to a small number of households that use a higher volume of water collected from surface sources.

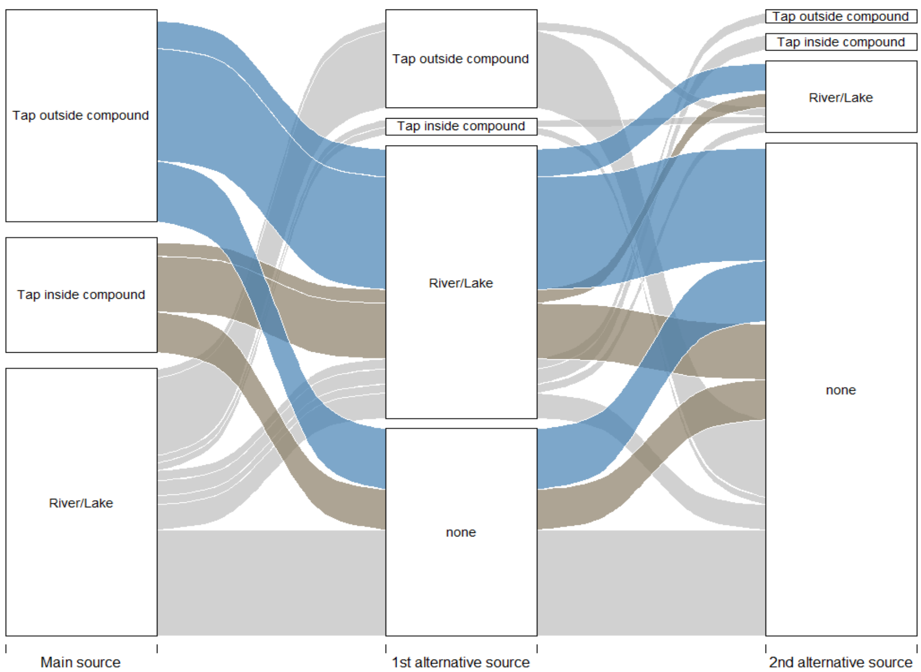


Figure 4-5 Main and alternative sources distribution amongst households interviewed in 2016. Blocks height represents the proportion of households using a water source as main or alternative source. Flows represent different combinations of main and alternative sources.
Note: the diagram does not show distributions encountered in less than 5 households – 8 combinations representing 14 households are not represented; N=455

Main and alternative sources of drinking water (2016 survey)

More than half the households interviewed (52 %) reported having drunk water collected at a tap outside their compound in the previous two weeks, and nearly 44 % had drunk water collected at a river. More than three quarters of the interviewed households (76.4 %) reported having drunk tap water during the previous two weeks, but exclusive tap water drinking was reported by 46.7 % (**Table 4-5**).

Table 4-5 Type and number of sources reported to have been used for drinking during the two weeks preceding interview, by household (survey 2016)

		Tap water service quintile n (% of households)				
	All	1 (worst)	2	3	4	5 (best)
Number of households	**471**	83	95	92	95	106
Number of households reporting using for drinking						
Lake Tanganyika	**29 (6.2%)**	3 (3.6%)	9 (9.5%)	9 (9.8%)	5 (5.3%)	3 (2.8%)
amongst users	***18.7%***	*100%*	*25.7%*	*22%*	*17.2%*	*6.4%*
Rivers and streams	**207 (43.9%)**	49 (59%)	67 (70.5%)	43 (46.7%)	31 (32.6%)	17 (16%)
amongst users	***70.9%***	*76.6%*	*90.5%*	*79.6%*	*50.8%*	*43.6%*
Tap outside the compound	**245 (52%)**	47 (56.6%)	40 (42.1%)	58 (63%)	52 (54.7%)	48 (45.3%)
amongst users	***98.4%***	*100%*	*95.2%*	*96.7%*	*100%*	*100%*
Tap in the compound	**115 (24.4%)**	4 (4.8%)	14 (14.7%)	13 (14.1%)	29 (30.5%)	55 (51.9%)
amongst users	***98.3%***	*100%*	*100%*	*100%*	*100%*	*96.5%*
Rainwater	**1 (0.2%)**	0 (0%)	1 (1.1%)	0 (0%)	0 (0%)	0 (0%)
amongst users	***33.3%***	*0%*	*100%*			*0%*
Leaking water supply pipe	**7 (1.5%)**	1 (1.2%)	0 (0%)	0 (0%)	6 (6.3%)	0 (0%)
amongst users	***100%***	*100%*			*100%*	
Reported number of drinking water sources used						
1 source only	**308 (65.4%)**	59 (71.1%)	55 (57.9%)	52 (56.5%)	61 (64.2%)	81 (76.4%)
2 sources only	**136 (28.9%)**	21 (25.3%)	37 (38.9%)	33 (35.9%)	24 (25.3%)	21 (19.8%)
3 sources or more	**13 (2.8%)**	2 (2.4%)	3 (3.2%)	2 (2.2%)	6 (6.3%)	0 (0%)
Unknown	*14 (3%)*	*1 (1.2%)*	*0 (0%)*	*5 (5.4%)*	*4 (4.2%)*	*4 (3.8%)*
Number of households reporting drinking tap water						
	360 (76.4%)	51 (61.4%)	54 (56.8%)	71 (77.2%)	81 (85.3%)	103 (97.2%)
Number of households reporting drinking exclusively tap water						
	220 (46.7%)	27 (32.5%)	14 (14.7%)	40 (43.5%)	53 (55.8%)	86 (81.1%)

Almost all households reporting having used a tap inside or outside of the compound as a source reported drinking the tap water (98 %). A much higher proportion of households using river water reported drinking it than households reporting using lake water (71 % vs 20 %). The diversity of sources is lower for water used for drinking than for all uses, with 32 % of households reporting that they had drunk water from two or more sources in the previous two weeks (**Table 4-5**).

77 households (16.3 %) reported not drinking their main source of water, either the lake or a river/stream for 74 of them.

Perceived quality of water sources (2016 survey)

The most commonly reported issue with all sources was coloured/turbid/dirty water (62.7 % of all sources), followed by bad taste and low pressure for tap sources (42.7 % and 43 % respectively). Colouring, smell or taste issues were more often reported with lake and river water, while low pressure affected nearly half the users of taps outside and inside the compound (**Table 4-6**).

Low pressure was reported as frequently for taps inside or outside the compound, at any source rank and any tap level service.

Table 4-6 Perceived quality issues reported by households about the sources of water used during the two weeks preceding the interview (survey 2016)

		The water is coloured or dirty/ turbid	The water has a bad taste	The water smells bad	Not enough pressure
		n (% of total source users)			
Source of the water					
Lake Tanganyika	155	141 (91%)	135 (87.1%)	81 (52.3%)	
Rivers and streams	320	285 (89.1%)	167 (52.2%)	128 (40%)	
Tap outside of the compound	257	64 (24.9%)	36 (14%)	29 (11.3%)	116 (45.1%)
Tap in the compound	117	43 (36.8%)	25 (21.4%)	21 (17.9%)	45 (38.5%)
Other	10	6 (60%)	4 (40%)	4 (40%)	

Water availability at taps (2016 survey)

259 households reported on the number of days and nights the week preceding the interview when water was available at the taps they used as their main source (**Table 4-7**).

More than a third of those reporting using a tap outside their compound as their main source of water reported there was no water at the used tap during either daytime or night-time the previous week. A quarter of them reported there was no water at the used tap during both day and night-time. The likelihood of reporting unavailability of water during day- and/or night-time for users of taps outside the compound tended to decrease as tap water service improved. A large majority (85.6 %) of users of taps outside the compound as main source also reported not having water at the used tap when they intended to use it, with little variation across tap water service quintiles. Nearly a third of users of a tap outside the compound reported collecting water during the night-time, more frequently as tap water service quintile increased.

Table 4-7 Reported water availability at taps used as main water sources, inside and outside the compound (survey 2016)

	N	All quintiles	N	1 (worst)	N	2	N	3	N	4	N	5 (best)
				Tap water service quintile								
Number of days water was available in the past week												
Tap outside of the compound (N)	168		33		24		27		41		43	
lo water available during daytime (n (%))		61 (36.3%)		16 (48.5%)		11 (45.8%)		12 (44.4%)		7 (17.1%)		15 (34.9%)
Median (Range; IQR)		2 (0 - 7; 0 - 4)		0 (0 - 7; 0 - 4.5)		1 (0 - 7; 0 - 3)		0 (0 - 7; 0 - 2)		2 (0 - 7; 1 - 5)		2.5 (0 - 7; 0 - 4)
missing (n (%))		*24 (14.3%)*		*7 (21.2%)*		*2 (8.3%)*		*6 (22.2%)*		*6 (14.6%)*		*3 (7%)*
Tap in the compound (N)	91		2		1		12		24		52	
lo water available during daytime (n (%))		25 (27.5%)		0 (0%)		1 (100%)		4 (33.3%)		8 (33.3%)		12 (23.1%)
Median (Range; IQR)		3 (0 - 7; 0 - 5)		2.5 (2 - 3; 2.25 - 2.75)		0 (0 - 0; 0 - 0)		2.5 (0 - 7; 0 - 5.25)		2 (0 - 7; 0 - 4)		3.5 (0 - 7; 0.75 - 5)
missing (n (%))		*5 (5.5%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*1 (4.2%)*		*4 (7.7%)*
Number of nights water was available in the past week												
Tap outside of the compound (N)	168		33		24		27		41		43	
water available during night-time (n (%))		61 (36.3%)		20 (60.6%)		13 (54.2%)		9 (33.3%)		8 (19.5%)		11 (25.6%)
Median (Range; IQR)		1 (0 - 7; 0 - 5)		0 (0 - 7; 0 - 0)		0 (0 - 7; 0 - 4)		0 (0 - 7; 0 - 4.25)		3 (0 - 7; 0.25 - 5)		4 (0 - 7; 0 - 5)
missing (n (%))		*45 (26.8%)*		*9 (27.3%)*		*3 (12.5%)*		*11 (40.7%)*		*11 (26.8%)*		*11 (25.6%)*
Tap in the compound (N)	91		2		1		12		24		52	
water available during night-time (n (%))		22 (24.2%)		1 (50%)		0 (0%)		5 (41.7%)		5 (20.8%)		11 (21.2%)
Median (Range; IQR)		4 (0 - 7; 0 - 6)		3.5 (0 - 7; 1.75 - 5.25)		7 (7 - 7; 7 - 7)		2 (0 - 7; 0 - 4.25)		4 (0 - 7; 1 - 5)		4 (0 - 7; 0 - 6)
missing (n (%))		*14 (15.4%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*3 (12.5%)*		*11 (21.2%)*
No water during day and night time in the past week												
Tap outside of the compound	168	43 (25.6%)	33	15 (45.5%)	24	9 (37.5%)	27	7 (25.9%)	41	5 (12.2%)	43	7 (16.3%)
missing (n (%))		*47 (28%)*		*10 (30.3%)*		*3 (12.5%)*		*11 (40.7%)*		*12 (29.3%)*		*11 (25.6%)*
Tap in the compound	91	14 (15.4%)	2	0 (0%)	1	0 (0%)	12	2 (16.7%)	24	4 (16.7%)	52	8 (15.4%)
missing (n (%))		*15 (16.5%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*3 (12.5%)*		*12 (23.1%)*
No water at the tap when going to collect water												
Tap outside of the compound	168	144 (85.7%)	33	26 (78.8%)	24	21 (87.5%)	27	20 (74.1%)	41	38 (92.7%)	43	39 (90.7%)
missing (n (%))		*1 (0.6%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*1 (2.4%)*		*0 (0%)*
Tap in the compound	91	79 (86.8%)	2	2 (100%)	1	1 (100%)	12	9 (75%)	24	23 (95.8%)	52	44 (84.6%)
missing (n (%))		*1 (1.1%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*1 (1.9%)*
Collecting water at the tap during nighttime												
Tap outside of the compound	168	58 (34.5%)	33	5 (15.2%)	24	6 (25%)	27	7 (25.9%)	41	25 (61%)	43	15 (34.9%)
missing (n (%))		*1 (0.6%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*1 (2.3%)*
Tap in the compound	91	54 (59.3%)	2	2 (100%)	1	0 (0%)	12	6 (50%)	24	17 (70.8%)	52	29 (55.8%)
missing (n (%))		*2 (2.2%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*0 (0%)*		*2 (3.8%)*

Approximately a quarter of those reporting using a tap inside the compound as their main source of water reported that there was no water at the tap during either daytime or night-time the previous week. 15 % of them reported no water availability during both day and night-time. Unavailability of water was slightly less often reported as tap water service increased. As for users of taps outside the compound, a large majority (86.8 %) of users of a tap inside the compound as main source experienced a lack of water at the tap when they intended to use it, with little variation over tap water service quintiles. A majority of them (59.3 %) had collected water at night, independently of tap water service.

Missing data was more frequent for taps outside of the compound, and for night-time.

Households' maximum storage capacity (2016 survey)

When asked about the maximum amount of water they could store in their house if filling up all the containers they had, half the households reported a capacity greater than 200 litres (median 205 litres; range 20 – 7'675; IQR 130 – 310), and greater than 29 litres per capita (median 29.2; range 2.8 – 959; IQR 18 – 44). Households reported a storage capacity ranging from 0.15 to 56 times the amount of water used the day before in their dwelling, with half of them between 0.9 and 2.1 times. Storage capacity expressed in number of days' consumption is significantly higher for households using taps as a main source of water (median 1.5 versus 1.2 for users of other sources; Kruskal-Wallis sum rank test p<0.001) and for the three wealthier categories (**Table 4-8**).

Table 4-8 Households' maximum water storage capacity (survey 2016)

	Household maximum water storage capacity (as ratio over quantity of water used at home the previous day)	
	n	Median (range; IQR)
Main source of water		
Lake Tanganyika	56	1.1 (0.3 - 4.2; 0.8 - 1.6)
Rivers and streams	154	1.2 (0.1 - 11.3; 0.7 - 1.9)
Tap outside the compound	168	1.5 (0.2 - 55.8; 1 - 2.5)
Tap in the compound	91	1.5 (0.2 - 40.8; 1.1 - 2.5)
Leaking water supply pipe	2	3.7 (2.3 - 5; 3 - 4.3)
Wealth quintile		
1	99	1 (0.1 - 5; 0.7 - 1.5)
2	89	1.2 (0.2 - 5.9; 0.9 - 1.8)
3	94	1.5 (0.2 - 11.8; 1 - 2.1)
4	92	1.4 (0.3 - 11.3; 0.9 - 2.5)
5	91	1.7 (0.5 - 55.8; 1.2 - 2.6)
missing	*6*	*1.9 (0.4 - 2.7; 1.5 - 1.9)*

4.3.3 Monetary cost of tap water

For households reporting using a tap they do not own

303 households (64.3 %) reported having used in the previous two weeks a tap that was not their own (as a Regideso subscriber) – overall this represented 314 tap sources used (main or alternative). Most of them (225 or 71.7 %) were paid for monthly, while 42 (13.4 %) were free for the households reporting using them. A further 35 (11.1 %) taps were paid for per 20L jerrycan, and 12 (3.8 %) in-kind. Of the 314 taps, nearly half (48.5 %) were owned by another household while 31 % were owned by an NGO/the community or considered public. 13 % were owned by the landlord of the dwelling, and 6% by a family member in another dwelling. Free use and monthly payment methods were broadly similarly distributed for taps inside or outside the compound but taps inside the compound are more likely to be paid for in-kind and less per jerrycan. The main source of water is twice as likely to be paid for monthly as the other sources and is more often owned by the landlord than a family member in another dwelling.

The price paid monthly ranged from 100 congolese francs (CDF) to 25,000 CDF (median 1,500 CDF; IQR 500 – 2,000 CDF)[5]. When estimated per capita, price paid monthly ranged from 11 CDF to 3,429 CDF (median 167 CDF; IQR 83 – 306 CDF). The price paid monthly per capita was higher for a tap inside the compound (median 300 CDF per capita per month) than for a tap outside of the compound (median 157 CDF per capita per month - p-value of Kruskal-Wallis test < 0.001). For sources paid by per jerrycan, a majority (20 out of 35) charged 100 CDF per 20L jerrycan, with prices for the others ranging from 50 CDF to 400 CDF.

Out of the 224 households reporting paying monthly for their water, 145 households (64.7 %) had experienced difficulties in paying for their water at some point in the previous year. 25 of the 35 (71.4 %) households paying per jerrycan reported having had to revert to a free source of water at least once in the previous month because they did not have enough money to pay for water.

For households owning the tap they use

57 interviewed households owned the tap they used as the main source of water. 37 of them (65 %) reported having a functional meter installed on it, of which 30 reported that the Regideso invoiced them based on meter readings. 27 households reported being invoiced a flat rate.

---

[5] In October 2016, the exchange rate for 1 USD was approximately 975 CDF and the daily minimum wage for an "ordinary worker" was fixed by law in 2009 at 1,680 CDF.

42 tap owners could remember how much they paid the Regideso the month preceding the interview. It ranged between 2,000 CDF to 37,000 CDF, with a median of 8,190 CDF (IQR 6,000 CDF – 10,990 CDF). The amount paid per capita was estimated for the 22 households reporting being sole users of the tap and remembering the amount of their last Regideso monthly invoice. The monthly price per capita ranged between 233 CDF and 7,400 CDF, with a median of 1,025 CDF (IQR 644 CDF – 1,044 CDF).

4.3.4 Water collection outside the compound

In 2016, 439 households (93.3 %) reported having used at least one source of water outside their compound in the two weeks preceding the interview. 380 households (80.7 %) reported that their main source of water had been located outside their compound. In 2017, only details about the main source were collected from interviewed households. 361 households out of 446 interviewed (80.9 %) reported having used a main source of water located outside of their compound during the previous two weeks.

Household members involved in water collection (2016 survey)

310 households could report who went to collect water at their main water source outside the compound the last time. The remaining 70 households which reported using a main source outside their compound could not remember or know who went to collect the water.

The female head of household was most often involved in water collection (54.5 % of households), followed by female children aged 10 to 15 years old (40.3 % of households) (**Table 4-9**).

Table 4-9 People involved in water collection outside the compound the last time the main source was used (survey 2016). Note: Many households reported multiple people having gone together.

	Main source outside the compound (N=310 - missing data 70 HH [18%])
	n (%)
Male head of household	19 (6.1%)
Female head of household	169 (54.5%)
Male adult of the household	44 (14.2%)
Female adult of the household	96 (31%)
Male child aged 10-15	94 (30.3%)
Female child aged 10-15	125 (40.3%)
Male child aged 10 and under	25 (8.1%)
Female child aged 10 and under	29 (9.4%)
Child or adult not member of the household	3 (1%)
Remunerated water collector	2 (0.6%)

The data suggest that male children are more involved in collection of water at alternative sources, as are non-members of the household, remunerated or not. More than

11 % of the households reported that a male or a female child under 10 was involved in water collection.

More than half the time when the female head of household was involved in water collection, she was accompanied by at least one other member of the family (**Figure 4-6**).

Figure 4-6 Mother and children collecting water from a leaking water pipe

Activities performed directly at the source (2017 survey)

112 households (25.6 % of all interviewed households) reported having performed some water-consuming activities directly at their main source of water outside the compound in the two weeks preceding the interview. The most frequently reported activity was laundry (88 households; 19.7 % of all interviewed households), followed by bathing by children under 15 years of age (48 households; 10.8 %). 29 (6.5 %) and 31 (4.7 %) households reported performing adults bathing and dishwashing respectively.

Households using a surface water source as their main source were nearly 10 times more likely to perform activities directly at the source than households using a tap outside their compound (82.6 % vs 8.5 %).

Time spent to collect water outside the compound (2017 survey)

Out of 361 households reporting using a source of water outside their compound as their main source in quantity in the two weeks preceding the interview, 303 were able to report how much time it took their members to collect water. The total time spent per trip to collect water outside the compound, including a return journey and waiting time for tap users, ranged between 2 and 280 minutes with a median of 40 minutes (IQR 15 – 75 minutes).

When stratified by the type of main source, users of lake water as main source spent noticeably less time to collect water than users of river water, and users of taps spent more time to collect water than river water users (**Table 4-10**).

Data from 118 households using a tap outside of their compound as main source, show a tendency for the time spent to collect water to decrease at the three best levels of tap water service.

For users of a tap outside their compound, waiting time represents between 17 and 98 % of the total time spent to collect water (median = 86 %; IQR 75 – 92 %).

Table 4-10 Total time spent to collect water outside the compound, including return journey and waiting time, in minutes, stratified by main source type and tap water service level (survey 2017)

	Total time spent to collect water		
	N	Median (Range; IQR)	n missing
Main source of water			
Lake Tanganyika	66	15 (2 - 60; 8 - 30)	10
Rivers and streams	155	30 (3 - 150; 10 - 50)	19
Tap outside the compound	140	85 (3 - 280; 57 - 130.5)	29
Tap in the compound	82	Set to 1 minute	0
Rain	3		3
Tap water service quintile			
1 (worst)	78	60 (3 - 280; 30 - 126.25)	6
2	87	42.5 (1 - 250; 9.5 - 62.5)	11
3	88	25 (1 - 151; 8 - 60)	17
4	91	12 (1 - 130; 1 - 48)	16
5 (best)	102	4 (1 - 184; 1 - 42.5)	11

4.3.5 Quantity of water used at home for domestic activities

Water reuse (2017 survey)

32 households out of 446 (7.2 %) reported recycling some of their water for selected domestic activities. A large majority of these (28 households) declared only recycling water for house cleaning purposes, while dishwashing, food/produce rinsing and/or laundry were cited by the other 4 households.

Total amount of water used at home (2017 survey)

444 households reported the amount of water they used for various domestic activities, at home, on the day preceding the interview. The total amount of water used per household ranged from 10 to 460 litres, with a median of 143 litres (IQR 99 – 200 litres). When divided by the number of household members reported to be present the previous day, the amount used per person ranged from 3.3 to 107 litres per capita per day (lcd) with a median of 21 lcd (IQR 16 – 27 lcd).

Distribution of water quantity used per capita, stratified by type of main water source, tap water service level, wealth quintile and time spent to collect water, is shown in **Table 4-11**. Data suggests that users of a tap inside their compound use more water at home than users of other sources. Amount of water used per capita tends to increase slightly as tap water service level increases, especially at the two highest levels. Expectedly, it tends to decrease as time spent to collect water increases, noticeably between households spending less than 2 minutes vs those spending more than 2 minutes. Finally, amount of water used per capita seems to be higher for the wealthiest households.

Table 4-11 Quantity of water per capita reported to have been used at home the day preceding the interview, in litres per capita per day, stratified by main source type, tap water service level, categorised time spent to collect water and wealth quintile (survey 2017)

	N	**Water quantity used in lcd** Median (Range; IQR)
Main source of water		
Lake Tanganyika	66	20.4 (3.3 - 54.5; 14.5 - 28.2)
Rivers and streams	153	19.3 (6.5 - 52.3; 15.3 - 25)
Tap outside of the compound	140	20.5 (5.6 - 79; 15.6 - 25.9)
Tap in the compound	82	24.9 (11.7 - 107; 20.4 - 33)
Rain	3	14.9 (9.9 - 30.9; 12.4 - 22.9)
Tap water service quintile		
1 (worst)	78	18.1 (8.6 - 43.5; 14 - 22.1)
2	87	20.3 (3.3 - 45.5; 14.8 - 27.1)
3	88	18.6 (6.5 - 107; 15 - 25.3)
4	91	22.9 (8.6 - 55; 18.1 - 30.8)
5 (best)	100	23.1 (5.6 - 75; 18.1 - 29.3)
Categorised time spent to collect water		
2min and less	83	25.3 (11.7 - 107; 20.5 - 34)
3 to 30min	126	19.2 (3.3 - 54.5; 15.3 - 26.4)
31 to 90min	115	20.4 (8.8 - 55; 15.4 - 26)
More than 91min	61	18.1 (8.6 - 79; 15 - 23.2)
missing	*59*	*20 (8.2 - 66.5; 14.9 - 26.1)*
Wealth quintile		
1	88	20 (8.2 - 54.5; 16.3 - 26.1)
2	86	18.5 (3.3 - 52.3; 14.2 - 24.8)
3	87	19.4 (5.6 - 43; 14.4 - 26.3)
4	89	21.3 (6.5 - 107; 16.4 - 28.6)
5	88	23 (12.2 - 57; 19.4 - 28.8)
missing	*6*	*25.7 (21.9 - 38.2; 23.9 - 26.3)*

Results of linear regression models of the total amount of water used at home by a household, as a function of its demographic composition and time spent to collect water, are presented in **Table 4-12**. In this analysis, households' wealth quintile, payment for water, water reuse or performing activities at the source were considered as potential confounders for the relationship between water quantity used and time spent to collect water.

Table 4-12 Results of linear regression models of the total amount of water used at home by households, as a function of its demographic composition, water reuse, activity performed at the source, payment for water, time spent to collect water and wealth quintile (survey 2017). Note: the final model was fitted to a dataset from which 4 data points were excluded, based on regression diagnostics performed on the full database model.

	Final model (N=381)		*Full dataset model (N=385)*	
	Coefficient (95% CI)	**p-value**	*Coefficient (95% CI)*	*p-value*
Intercept (household with 1 adult member)	99.4 (77.9; 120.9)	p<0.001	*101.4 (78.9; 124)*	*p<0.001*
Number of extra household members aged 15 or more	18 (15.4; 20.7)	p<0.001	*17.6 (15; 20.2)*	*p<0.001*
Number of extra household members aged between 5 and 14	15.2 (12.2; 18.2)	p<0.001	*14.4 (11.2; 17.5)*	*p<0.001*
Number of extra household members aged between 2 and 4	10.1 (4.3; 15.9)	p<0.001	*11.8 (5.7; 17.9)*	*p<0.001*
Number of extra household members aged 2 or less	2.5 (-5.4; 10.5)	0.53	*2 (-6.2; 10.3)*	*0.63*
Activity performed directly at the source	-29.1 (-41.9; -16.3)	p<0.001	*-30 (-43.3; -16.6)*	*p<0.001*
Reuse of water at home	-5.6 (-25.7; 14.5)	0.583	*-4.3 (-25; 16.4)*	*0.68*
Paying for water	-6.5 (-21.5; 8.4)	0.393	*-3.7 (-19.3; 11.9)*	*0.65*
Total time spent to collect water				
2 minutes and less	reference		*reference*	
Between 3 and 30 minutes	-23.4 (-43.2; -3.6)	0.02	*-16.8 (-37.4; 3.8)*	*0.11*
Between 31 and 90 minutes	-23.9 (-42.4; -5.5)	0.01	*-17.7 (-36.9; 1.5)*	*0.07*
More than 91 minutes	-34.3 (-53.1; -15.5)	p<0.001	*-30.9 (-50.6; -11.3)*	*0.002*
Wealth quintile				
1 (poorest)	-3.5 (-22.2; 15.3)	0.72	*-12.8 (-32.2; 6.6)*	*0.20*
2	-12.3 (-31; 6.4)	0.20	*-19.3 (-38.8; 0.3)*	*0.05*
3	-11.6 (-30; 6.8)	0.22	*-18 (-37.3; 1.2)*	*0.07*
4	-0.2 (-16.9; 16.6)	0.98	*-4.7 (-22.2; 12.8)*	*0.60*
5 (Wealthiest)	reference		*reference*	
Adjusted R-squared		0.57		*0.56*
Residual standard error		49.1 (365 df)		*51.6 (365 df)*

Model estimates suggest that a household composed of a single adult member, not performing any activity at the source or recycling water, not paying for water, spending less than 2 minutes to collect water and amongst the wealthiest category, uses 99.4 litres per day (95% CI 77.9 -120.9). Each extra member of household aged 15 or more is estimated to use an extra 18 litres per day (95% CI 15.4 – 20.7), while extra members aged 5 to 14 and 2 to 4 are estimated to use respectively 15.2 and 10.1 litres per day (95% CIs 12.2 – 18.2 ; 4.3 – 15.9). Extra members aged less than 2 appear to be associated with little increase in water consumption (2.5 litres; 95% CI -5.4 – 10.5).

Reporting water reuse in the home or paying for the water collected do not seem associated with a different amount of non-recycled water used but performing activities directly at the source is associated with a reduction of 29.1 litres (95% CI 16.3 – 41.9) in household daily consumption. Households using a main source located between 3 and 30 minutes and between 31 and 90 minutes from their home (return journey including waiting time) are estimated to use 23.4 and 23.9 litres less respectively (95% CIs 3.6 –

43.2 ; 5.5 – 42.4) than households spending 2 minutes or less (main source within the compound). Households reporting that they spend more than 90 minutes to collect water at the source are expected to use 34.3 litres less (95% CI 15.5 – 53.1) than those using a source within their compound. Independently of the time spent to collect water, there was little evidence of wealth quintile being associated with the quantity of water used at home by households.

The linear regression model was estimated to explain 57% of the total variance observed in total amount of water used by included households.

When the same linear model is run as a function of household composition, wealth quintile and tap water service only, the two highest level of service are associated with an increase of 18.1 and 20.8 litres respectively in comparison with the total amount of water used in households in the worst served areas, independently of wealth quintile.

Amount of water used for specific domestic activities

For the 444 households that reported their water use, the quantity of water used for adult bathing represents the highest proportion of total water use, with a median of 29.1 %, followed by children bathing (median 14.1 %) and laundry (median 13.3 %). The quantity reported to be used for handwashing represents the lowest proportion of the total at a median of 2.9 %. Proportions of the total amount of water used at home spent for specific activities are reported in **Table 4-13**.

Table 4-13 Proportions of the total amount of water used at home reported for specific activities by 444 households (survey 2017)

Activity	Proportion of the total amount of water used at home Median (IQR; max)
Adult bathing	29.1% (21.1 - 37.9%; 66.7%)
Children bathing	14.1% (7.7 - 18.9%; 41.7%)
Laundry	13.3% (0 - 28%; 63.7%)
Food / produce washing	11.4% (8 - 15.7%; 40%)
Dishwashing	9.3% (6.5 - 12.2%; 37.4%)
House cleaning	7.4% (4.8 - 10.5%; 38.3%)
Drinking	5.9% (4.1 - 8.7%; 40%)
Handwashing	2.9% (2 - 4.3%; 20%)

A similar regression model as for total quantity of water was applied to amounts of water reported to have been used for specific activities. Three activities appear to be affected by the time spent to collect water, independently of wealth and reuse or performing activities at the source: 1) adults bathing; 2) house cleaning and 3) dishwashing. The amount of water used for adult bathing decreases noticeably as time spent to collect water increases, with households spending between 30 and 90 minutes or more than 90 minutes using respectively 11.8 and 15.4 litres less per day on adult bathing, when the extra amount of water used for bathing by an extra household adult member is estimated at 11 litres. Similarly, the quantity of water used for house cleaning is reduced by approximately 5 litres per day once a household collects water outside of their compound. This decrease is however smaller than that observed as wealth quintile decreases. The quantity of water used for dishwashing is also reduced by about 2.5 litres per day when a household collects water outside the compound, a similar decrease to that observed as wealth quintile decreases or when household performs activities directly at the source. Detailed results of this analysis are available in **Appendix 4-4**.

4.3.6 Drinking water quality

Results of drinking water quality analyses are available for 441 households interviewed during the 2017 survey.

Turbidity levels

Drinking water samples turbidity ranged from 0.7 to 46.5 NTU, and 39.5 % of the samples had a low turbidity of 3 NTU or less. Turbidity greater than 10 NTU was measured in 7.4 % of the samples, mostly from samples collected from rivers (**Table 4-14**).

Table 4-14 Measured turbidity in drinking water samples collected from 441 households, stratified by reported origin of the sample (survey 2017)

		Turbidity n (% of samples)			
	N	**NTU<=3**	**3<NTU<=5**	**5<NTU<=10**	**NTU>10**
All samples	**441**	174 (39.5%)	133 (30.2%)	100 (22.7%)	34 (7.7%)
Source of the sample					
Lake	14	3 (21.4%)	5 (35.7%)	5 (35.7%)	1 (7.1%)
Rivers	115	11 (9.6%)	22 (19.1%)	55 (47.8%)	27 (23.5%)
Tap outside the compound	211	103 (48.8%)	78 (37%)	28 (13.3%)	2 (0.9%)
Tap inside the compound	94	55 (58.5%)	27 (28.7%)	10 (10.6%)	2 (2.1%)
Rain	7	2 (28.6%)	1 (14.3%)	2 (28.6%)	2 (28.6%)

Chlorine levels

39 samples out of 441 (8.8 %) showed presence of total chlorine, and 37 of them presence of residual chlorine. 23 had between 0.1 and 0.4 mg/l of residual chlorine, and 14 were measured between 0.5 and 1 mg/l.

Levels of residual chlorine measured in samples, by reported source, reported treatment with chlorine by the household and tap water service level of the household, are shown in **Table 4-15**.

Table 4-15 Number of households and levels of residual chlorine measured in their drinking water samples, by reported source of the sample, reported treatment with chlorine since collection and tap water service level of the household (N=441; survey 2017)

Reported source	Reported chlorination	Tap water service level	N	Residual chlorine levels 0 mg/l	0.1 - 0.4 mg/l	0.5 - 1 mg/l
Lake / River / Rain						
	No chlorination		133	132	0	1
River						
	Chlorination at source		3	0	3	0
Tap outside the compound						
	Chlorination at source		1	1	0	0
Tap outside the compound						
	No chlorination					
		1 & 2	74	74	0	0
		3	48	47	1	0
		4	41	38	3	0
		5	47	35	6	6
Tap inside the compound						
	No chlorination					
		1 & 2 & 3	13	13	0	0
		4	32	25	6	1
		5	49	39	4	6

Samples with residual chlorine were mostly reported to have been collected at a tap and had not been chlorinated since collection (n = 33). 3 samples originated from rivers and were reported to have been chlorinated at the point of collection. A sample supposedly originating from a river and not having been chlorinated showed residual chlorine levels of 0.5 mg/l or higher, suggesting a mistake in the information reported or the measurement.

Recommended levels of residual chlorine (0.5 mg/l or more) were more likely to be measured for drinking water originating from a tap inside or outside the compound, at the highest level of tap water service. Suboptimal levels of residual chlorine were likely to originate from taps inside or outside the compound, at lower levels of tap water service.

As expected, presence of residual chlorine in the sample is associated with lower levels of *E. coli* contamination, especially at the recommended concentration (**Table 4-16**).

Table 4-16 Results of *E. coli* analyses for drinking water samples, expressed in risk categories, according to residual chlorine levels measured (N=441; survey 2017)

	N	*E. coli* contamination level Safe	Low	Intermediate	High	Very High
Residual chlorine levels						
0 mg/l	404	118 (29.2%)	86 (21.3%)	75 (18.6%)	99 (24.5%)	26 (6.4%)
0.1 - 0.4 mg/l	23	14 (60.9%)	7 (30.4%)	2 (8.7%)	0 (0%)	0 (0%)
0.5 - 1 mg/l	14	12 (85.7%)	2 (14.3%)	0 (0%)	0 (0%)	0 (0%)

Risk factors for E. coli contamination

We first examined how drinking water sample characteristics were associated with *E. coli* contamination. Sample characteristics included reported origin of the drinking water, time since collection, reported treatment of the water to make it safer (including "perceived" treatments such as letting the water rest or filtering it through a cloth, for which efficacy to remove pathogens is likely low), means of serving the water and storage conditions. Safe storage was defined as drinking water stored in a closed container, stored above the floor and with an opening that did not allow a hand to be dipped into the water. The two last characteristics mentioned above were recorded when the interviewer was allowed to see where the drinking water was stored and how it was dispensed into the sampling bag. Levels of residual chlorine were considered to be on the causal pathway between sample characteristics and contamination levels, and therefore not adjusted for.

Table 4-17 *E. coli* contamination stratified by drinking water sample characteristics and results of univariable logistic regression (N=441, survey 2017)

		Unsafe (CFU>0)			High (CFU>10)		
	N	**n**	**Crude OR (95%CI)**	**p-value**	**n**	**Crude OR (95%CI)**	**p-value**
Source of the sample							
Lake	14	9	2.77 (0.86 - 8.92)	0.09	9	7.6 (2.27 - 25.43)	0.001
Rivers	115	105	16.18 (7.5 - 34.91)	p<0.001	96	21.33 (10.47 - 43.46)	p<0.001
Tap outside the compound	211	142	3.17 (1.92 - 5.25)	p<0.001	79	2.74 (1.53 - 4.91)	0.001
Tap inside the compound	94	37	reference		15	reference	
Rain	7	4	2.05 (0.43 - 9.71)	0.36	3	3.17 (0.65 - 15.42)	0.15
Safe storage							
No	385	266	reference		182	reference	
Yes	22	10	0.37 (0.16 - 0.89)	0.03	6	0.74 (0.31 - 1.77)	0.50
Not observed	34	21	0.72 (0.35 - 1.49)	0.38	14	0.75 (0.37 - 1.53)	0.43
Storage duration							
Today	178	127	reference		94	reference	
Yesterday	150	98	0.76 (0.47 - 1.21)	0.24	69	0.75 (0.49 - 1.16)	0.20
Day before yesterday	79	50	0.69 (0.4 - 1.21)	0.20	29	0.54 (0.31 - 0.92)	0.03
More than 2 days ago	33	21	0.7 (0.32 - 1.53)	0.38	9	0.31 (0.14 - 0.71)	0.01
Doesn't know	1	1	excluded		1	excluded	
Serving mean							
Container opening	356	236	reference		167	reference	
Utensil (Cup, laddle…)	55	42	1.64 (0.85 - 3.18)	0.14	24	0.81 (0.46 - 1.43)	0.47
Tap on the container	2	2	excluded		0	excluded	
Not observed	28	17	0.79 (0.36 - 1.73)	0.55	11	0.68 (0.31 - 1.49)	0.33
Reported treatment							
None	414	279	reference		188	reference	
Rest / Filter with a cloth	13	9	1.09 (0.33 - 3.6)	0.89	11	2.53 (0.77 - 8.34)	0.13
Boiling / Ceramic filter / chlorination at source or PoU	3	1	1.09 (0.33 - 3.6)	0.89	1	0.7 (0.23 - 2.18)	0.54
Doesn't know	1	0	excluded		0	excluded	

In univariable analyses, the odds of drinking water being unsafe (>0 CFU *E. coli*/100 ml) were associated with the reported origin of the water and with not being safely stored. The odds of being more heavily contaminated (>10 CFU *E. coli*/ 100ml) were associated with the reported origin of the water, and time since collection (**Table 4-17**).

After adjusting for potential confounding by safe storage practices and time since collection, the odds of the drinking water stored at home being unsafe were strongly associated with the origin of the water (**Table 4-18**). Drinking water collected at a tap outside the compound or at a river had 3.2- and 16.7-times greater odds respectively of being unsafe than water collected at a tap inside the compound. The sample size was too small to draw conclusions on contamination of drinking water collected at the lake or in a rain collector.

Adjusted for time since collection and safe storage, the odds of the drinking water being contaminated with more than 10 CFU *E. coli* per 100 ml were also strongly associated with the origin of the water (**Table 4-18**). Drinking water collected at a tap outside the compound, at the lake or at a river had 2.8-, 8.5- and 23.2-times greater odds respectively of having an intermediate or higher level of *E. coli* contamination than water collected at a tap inside the compound.

Table 4-18 Results of multivariable logistic regressions for two thresholds of *E. coli* contamination levels, as a function of origin of the sample, storage conditions and duration (N=441, survey 2017)

		Unsafe (CFU>0)			High (CFU>10)		
	N	**n**	**Adjusted OR (95%CI)**	**p-value**	**n**	**Adjusted OR (95%CI)**	**p-value**
Source of the sample							
Lake	14	9	2.73 (0.82 - 9.12)	0.102	9	8.48 (2.43 - 29.56)	p<0.001
Rivers	115	105	16.65 (7.49 - 36.98)	p<0.001	96	23.21 (10.97 - 49.1)	p<0.001
Tap outside the compound	211	142	3.15 (1.89 - 5.26)	p<0.001	79	2.79 (1.54 - 5.05)	p<0.001
Tap inside the compound	94	37	reference		15	reference	
Rain	7	4	1.93 (0.41 - 9.21)	0.409	3	3.46 (0.7 - 17.1)	0.13
Safe storage							
No	385	266	reference		182	reference	
Yes	22	10	0.55 (0.22 - 1.42)	0.22	6	1.42 (0.54 - 3.69)	0.48
Not observed	34	21	0.67 (0.3 - 1.48)	0.317	14	0.81 (0.35 - 1.83)	0.61
Storage duration							
Today	178	127	reference		94	reference	
Yesterday	150	98	0.99 (0.58 - 1.68)	0.969	69	1.26 (0.75 - 2.12)	0.39
Day before yesterday	79	50	1.29 (0.69 - 2.43)	0.429	29	1.27 (0.68 - 2.38)	0.46
More than 2 days ago	33	21	1.24 (0.53 - 2.88)	0.621	9	0.71 (0.29 - 1.77)	0.47
Doesn't know	1	1			1		

In a second stage, we explored the direct association between households' tap water service level, and *E. coli* contamination levels, adjusting for wealth quintile but not for other factors that were considered to be on the causal pathway for such an association. Results of this analysis are shown in **Table 4-19** and suggest that independently of wealth, household tap water service was strongly associated with the probability of drinking water being unsafe. Households residing in areas classified within the three better tap water service areas had two-, four- and six-times lower odds respectively of storing unsafe drinking water than households from the worst served areas. Wealth quintile did not appear independently associated with the odds of drinking water being unsafe. The odds of drinking water being contaminated at intermediate or higher levels were also associated with tap water service level, at the two highest levels only. Drinking water in households from levels 4 and 5 areas had three- and 10-times lesser odds of being contaminated with more than 10 CFU *E. coli* per 100 ml than that in households from worst served areas. Only households of the wealthiest quintile were less likely to store drinking water with more than 10 CFU *E. coli* per 100 ml independently from tap water service level.

Table 4-19 Results of multivariable logistic regressions for two thresholds of *E. coli* contamination levels, as a function of household wealth quintile and tap water service (N=441, survey 2017)

		Unsafe (CFU>0)			High (CFU>10)		
	N	n	Adjusted OR (95%CI)	p-value	n	Adjusted OR (95%CI)	p-value
Tap water service level							
1 (worst)	77	67	reference		55	reference	
2	85	75	1.21 (0.47 - 3.13)	0.69	60	1.07 (0.53 - 2.13)	0.85
3	87	63	0.47 (0.2 - 1.08)	0.08	49	0.65 (0.33 - 1.29)	0.22
4	90	50	0.25 (0.11 - 0.57)	p<0.001	31	0.3 (0.15 - 0.6)	p<0.001
5 (best)	102	42	0.16 (0.07 - 0.36)	p<0.001	14	0.1 (0.05 - 0.23)	p<0.001
Wealth quintile							
1 (poorest)	87	69	reference		56	reference	
2	86	70	1.13 (0.52 - 2.49)	0.75	54	0.89 (0.46 - 1.73)	0.74
3	86	61	0.8 (0.38 - 1.68)	0.55	45	0.74 (0.38 - 1.43)	0.37
4	88	55	0.75 (0.36 - 1.56)	0.44	34	0.59 (0.3 - 1.17)	0.13
5 (wealthiest)	88	40	0.46 (0.22 - 0.96)	0.04	18	0.3 (0.14 - 0.64)	0.002
missing	6						

4.3.7 Surface water contamination with *E. coli*

120 samples collected at 10 time points between the 26th of June 2017 and the 8th of November 2017 from 12 locations (5 lake shores and 7 river points) were analysed for *E. coli* contamination levels (**Table 4-20**). Log10 CFU *E. coli* per 100 ml ranged from 2.4 to 5.9, corresponding to 250 CFU/100 ml to 755,000 CFU/100 ml. Only 5 samples, all collected from the lake, were classified as "high risk" (between 101 and 1,000 CFU/100 ml), while all the others had a level contamination considered as "very high risk". Lake water was consistently less contaminated than river water with means of 3.8 vs 5.2 log10 CFU/100 ml respectively (repeated measures ANOVA p-value<0.001).

Table 4-20 Surface water contamination with *E. coli*, over 10 time points in 2017

		Log_{10} CFU *E. coli*/100ml	
Sampling location	**N**	**Mean (sd)**	**Range**
River Kalimabenge - upstream	10	5.05 (0.72)	3.7 - 5.88
River Kalimabenge - downstream	10	5.15 (0.61)	4.18 - 5.88
River Karigo	10	5.03 (0.68)	4 - 5.88
River Kavimvira - upstream	10	5.25 (0.45)	4.54 - 5.88
River Kavimvira - downstream	10	5.44 (0.36)	4.78 - 5.88
River Mulongwe - upstream	10	4.91 (0.5)	4.4 - 5.88
River Mulongwe - downstream	10	5.39 (0.5)	4.4 - 5.88
Tanganyika lake - Kalundu	10	4.12 (0.49)	3.44 - 4.58
Tanganyika lake - Kasenga	10	4.14 (0.33)	3.7 - 4.58
Tanganyika lake - Kilomoni	10	3.83 (0.64)	2.7 - 4.58
Tanganyika lake - Kimanga	10	3.35 (0.6)	2.4 - 4.58
Tanganyika lake - Mulongwe	10	3.76 (0.85)	2.4 - 4.58

4.3.8 *V. cholerae* contamination of drinking and surface water

485 samples of drinking water stored in households were analysed for cholera presence during the 2016 survey. None of the samples tested positive for *V. cholerae* O_1 by the field method employed.

None of the 120 surface water samples collected in 2017 tested positive for presence of *V. cholerae* O_1 either. However, amongst 107 samples collected at similar locations between the 11th and 26th of October 2016, 4 samples, collected from Lake Tanganyika in the Kilomoni neighbourhood, tested positive for *V. cholerae* O_1. No results for *E. coli* contamination are available however for drinking or surface water samples collected in 2016.

4.3.9 Weighted selected results from both surveys

Table 4-21 presents a selection of water-related practice indicators estimated from both survey datasets with post-sampling weighing of households.

Table 4-21 Weighted selected results from both surveys

	Data	**% (95% CI)**	**Median (95% CI)**
Proportion of households using rivers or lake as their MAIN water source	*survey 2016*	37.2% (33.4-41%)	
Proportion of households using a tap outside their compound as their MAIN water source	*survey 2016*	37.7% (33.1-42%)	
Proportion of households using a tap on premises as their MAIN water source	*survey 2016*	25.2% (21.3-29%)	
Proportion of households using tap water as ONE of their drinking water sources	*survey 2016*	79.6% (76-83%)	
Proportion of households using tap water EXCLUSIVELY as drinking water source	*survey 2016*	48.2% (44-52.6%)	
Time spent to collect water for users of a main source outside of the compound - including journey and waiting time - in minutes	*survey 2017*		40 (35-50)
Quantity of water used at home per capita per day, in lcd	*survey 2017*		21.3 (20.4-22.1)
Proportion of households storing SAFE drinking water (<1 CFU *E. coli*/100ml)	*survey 2017*	38.2% (34-42.8%)	
Proportion of households storing drinking water considered as intermediate, high or very high risk (>10 CFU *E. coli* /100ml)	*survey 2017*	40.5% (36.5-45%)	

The proportion of households using surface water or taps outside their compound as main source of water is similar at approximately 37 %, and higher than the proportion of households using mainly a tap within their compound (25 %). A large majority of households use tap water as one of their sources of drinking water (nearly 80 %) but drinking exclusively tap water is only expected for 48 % of households. The median time spent to collect water outside the compound for the 75 % of households not using a tap on premises is 40 minutes, including return journey and waiting time. The median amount of water used at home for domestic activities is 21.3 litres per capita per day. About 38 % of households are storing safe drinking water with no *E. coli* contamination, but 40 % are storing drinking water with *E. coli* contamination levels considered as intermediate, high or very high risk for diarrhoeal diseases.

4.4 Results summary

For all domestic use, including drinking, households in Uvira report using several water sources on a regular basis, as has been observed in other settings, especially urban **[15]**. Tap water service influences how likely a tap is to be amongst those sources, and how likely a household is to use surface water, either habitually or occasionally.

A substantial proportion of the households reporting using tap water – a quarter of users of a tap inside the compound and a third of users of a tap outside the compound - do not use it as their main source of water, and this proportion increases at the two (inside compound tap) or three (outside compound) lowest levels of tap water service. This highlights that having access to a tap, either in or outside the compound, does not automatically translate into using it most often, especially at low service levels. This may be a consequence of reduced availability of water at the taps at lower service levels (*see below*).

A majority of households drink the water collected at their main source, although about a third of those using surface water as their main source by quantity report not drinking it but using a tap instead. This implies a preference for tap water for drinking, which is further supported by the fact that a majority of households using surface water as main source perceive it as either dirty, tasting or smelling bad - much more frequently than users of tap water. This preference translated into most households having used tap water for drinking during the two weeks preceding the interview (80%). It is however insufficient to ensure taps are the only source of drinking water for notable proportion of them: only 50% of the population of Uvira is estimated to use exclusively tap water for drinking, more frequently where tap water is more regularly available, at the two best tap water service levels.

Lack of water at the tap is widely reported by users, whether inside or outside their compound. Total absence of water throughout the week preceding the interview was reported by a quarter of households using a tap outside their compound as main source and was increasingly frequently reported as tap water service worsened. Many households report collecting water from taps at night outside their compound, despite the absence of street lighting, criminality during the night-time or even curfews, substantiating previous findings of increased risk of exposure to violence during water collection **[16]**.

As in many other settings **[17]**, the female head of household appears to be the most frequent water collector outside the compound. Children under 15, and especially girls, are also commonly involved in water collection, including girls aged less than 10. For a quarter of interviewed households, water collection is associated with performing some water consuming activities, mostly laundry and child bathing, at surface water sources. Overall, half the households relying on a source outside their compound spend more than 40 min for every return journey made to collect water. More time is invested by households to collect water from taps than from rivers or lake, up to 4 hours a trip, although waiting commonly represents the majority of the time spent.

Figure 4-7 Washing dishes and collecting water in Lake Tanganyika

Half the population of Uvira is estimated to use less than 21 litres of water per capita per day at home for domestic activities, approximately the minimum recommended amount necessary to maintain basic hygiene according to United Nations High Commissioner for Refugees (UNHCR) **[18]**. The quantity of water used daily by a household is strongly associated with the time spent to collect water, independently of wealth, cost of water and performance of activities directly at the source. The "water use plateau" **[19]** can be observed in the data collected in Uvira: the quantity of water consumed daily decreases sharply as soon as the water is collected outside the compound, but remains stable as time needed to collect water increases further, until a threshold (here 90 min return journey) where the quantity of water used decreases again as time to collect water increases. Increased water usage due to collection time savings seems to be allocated in particular to adult bathing, household cleaning and dishwashing.

Analyses of drinking water stored at home first showed that very few households reported performing any treatment of their drinking water, with only 4 reporting their water having been chlorinated at the collection point. Recommended residual chlorine levels (0.5 mg/l or more) were measured in less than 3 % of the interviewed households, all located in neighbourhoods with the two highest tap water service levels. 60 % of the drinking water samples were also more turbid than optimal for chlorination treatment (>3 NTU) and 7 % would not be suitable for chlorination only (>10 NTU) **[20]**. Contamination of stored drinking water with *E. coli* is estimated to be found in more than 60 % of households in Uvira, with a greatly increased risk in households using river water. Drinking water collected at a tap outside the compound has three times greater odds of being contaminated than that collected inside the compound, suggesting that post collection contamination, documented in many other contexts, is highly prevalent **[21,22]**. Our data cannot however rule out contamination already at the tap, especially in areas with a lower level of service, due to the intermittency of water distribution **[8,9]**.

V. cholerae O_1 was not detected in stored drinking water sampled and was only detected in a single location of lake water sampling over 4 weeks in October 2016. Although the field method employed may lack sensitivity and this finding should be interpreted with caution, it implies that *V. cholerae* O_1 in a non-VBNC form is not ubiquitously present in either surface water or stored drinking water in households. Therefore, the importance of exposure to cholera directly from surface water sources in the disease transmission in Uvira cannot be confirmed from our findings, despite what some literature points to **[23]**.

The results described above clearly suffer from some limitations, inherent to the cross-sectional nature of the surveys and the use of interviews to collect information. In brief, analysis of data for a single point in time (one survey or the other in the present case) restricts the possibility of establishing causality in relationships between variables and challenges the representativity of the findings over time. In particular, conclusions cannot be drawn on households' water-related behaviour changes in the short-term, in response to tap water supply interruptions.In addition, using interviews to collect information on current and past behaviours is liable to information bias, particularly social desirability or recall biases. Finally, interviewing any adult household member may have led to further information bias, depending on the respondent's involvement in water-related practices.

4.5 Preamble for research paper 2

Results from the two households' surveys in Uvira demonstrated a clear association between our constructed indicator of tap water service and water-related practices in households highly relevant to diarrhoeal diseases transmission routes. Living in an area better served with tap water is associated with an increased probability of using tap water as main source of water for domestic use and as source of drinking water. It decreases the probability as well of using surface water for drinking, even occasionally. Households in better served areas also spend less time collecting water from a tap, which translates into an increased amount of water use at home for domestic activities, especially adult bathing, home cleaning and dishwashing. They are also less likely to store drinking water contaminated with *E. coli*, and at the best service level, they are more likely to have the recommended concentration of residual chlorine to prevent re-contamination in the short-term.

The above results were further documented in *Research Paper 2*, as an improved, continuous indicator for tap water service could be constructed based on the actual amount of water invoiced at georeferenced taps in March 2018. Correspondence between the tap water service levels used above, and the improved tap water service indicator is described in **Table 4-22**.

Research Paper 2 focuses on developing and assessing spatially explicit predictive models for estimating the probability of households storing contaminated drinking water and for estimating the quantity of water used at home for domestic activities, based on piped water access and distance from surface water sources.

Table 4-22 Correspondence between the tap water service level and the continuous tap water service indicator used in research paper 2 (N = 420)

		Continuous tap service indicator Median (IQR; Range)		
	n	**Radius 250m**	**Radius 500m**	**Radius 750m**
Tap water service level				
1 (worst)	52*	2.6 (1 - 5.6; 0.2 - 12.2)	22.1 (14.4 - 29.1; 4.3 - 65.9)	41.1 (30.2 - 66.8; 5.6 - 107.8)
2	87	13.8 (9.8 - 23.1; 2.2 - 124.7)	22 (15.3 - 32.5; 7.7 - 91.4)	28.3 (21.3 - 38.4; 11.3 - 104.2)
3	88	37 (20 - 49; 8 - 669)	42 (28 - 50; 15 - 450)	41 (30 - 56; 18 - 355)
4	91	67 (49 - 121; 19 - 668)	64 (49 - 122; 28 - 514)	61 (49 - 101; 25 - 425)
5 (best)	102	141 (101 - 262; 17 - 526)	125 (81 - 168; 19 - 414)	109 (73 - 132; 16 - 337)

* 26 households located in areas disputed between Uvira municipality and Uvira territory were excluded from the analysis and were all located within tap water service level 1.

4.6 Research Paper 2

Title: Predicting quality and quantity of water used by urban households based on tap water service

Authors: Aurelie Jeandron[1*], Oliver Cumming[1], Lumami Kapepula[2] and Simon Cousens[3]

[1] Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[2] Centre de Recherche en Hydrobiologie (CRH), Département d'Hydrologie, Section Hydrochimie, Uvira, République Démocratique du Congo

[3] Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

Status: Published in npj Clean Water in December 2019

Citation: Jeandron, A., Cumming, O., Kapepula, L. et al. Predicting quality and quantity of water used by urban households based on tap water service. npj Clean Water 2, 23 (2019).

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

<u>SECTION A – Student Details</u>

Student ID Number	173940	**Title**	Ms
First Name(s)	Aurélie		
Surname/Family Name	Jeandron		
Thesis Title	Tap water access and its relationship with cholera and other diarrhoeal diseases in an urban, cholera-endemic setting in the Democratic Republic of the Congo		
Primary Supervisor	Prof. Simon Cousens		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?	npj Clean Water		
When was the work published?	DECEMBER 2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

<u>SECTION C – Prepared for publication, but not yet published</u>

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Conceptualised the household survey methodology; designed survey tools and curated the data; managed laboratory resources and performed the water quality analyses; developed the model and performed the formal analyses; prepared the first draft; reviewed and revised towards final version.

SECTION E

Student Signature	Aurelie Jeandron
Date	21/05/2020

Supervisor Signature	Simon Cousens
Date	22/05/2020

npj | Clean Water

www.nature.com/npjcleanwater

ARTICLE OPEN

Predicting quality and quantity of water used by urban households based on tap water service

Aurelie Jeandron[1*], Oliver Cumming[1], Lumami Kapepula[2] and Simon Cousens[3]

Despite significant progress in improving access to safe water globally, inadequate access remains a major public health concern in low- and middle-income countries. We collected data on the bacterial quality of stored drinking water and the quantity of water used domestically from 416 households in Uvira, Democratic Republic of the Congo. An indicator of tap water availability was constructed using invoices from 3685 georeferenced piped water connections. We examined how well this indicator predicts the probability that a household's stored drinking water is contaminated with *Escherichia coli*, and the total amount of water used at home daily, accounting for distance from alternative surface water sources. Probability of drinking water contamination is predicted with good discrimination overall, and very good discrimination for poorer households. More than 80% of the households are predicted to store contaminated drinking water in areas closest to the rivers and with the worst tap water service, where river water is also the most likely reported source of drinking water. A model including household composition predicts nearly two-thirds of the variability in the reported quantity of water used daily at home. Households located near surface water and with a poor tap water service indicator are more likely to use water directly at the source. Our results provide valuable information that supports an ongoing large-scale investment in water supply infrastructure in Uvira designed to reduce the high burden of cholera and other diarrhoeal diseases. This approach may be useful in other urban settings with limited water supply access.

npj Clean Water (2019)2:23 ; https://doi.org/10.1038/s41545-019-0047-9

INTRODUCTION

The public health importance of both drinking water quality and the quantity of water available for domestic consumption has been long recognised.[1] The quality of drinking water influences "waterborne" disease transmission, that is the ingestion of infectious agents via contaminated drinking water, while the quantity of water available domestically influences "water-washed" disease transmission, that is where insufficient water is available to allow adequate hygiene practices.[2,3] Many efforts have been made to quantify the health risks associated with both, with an emphasis on diarrhoeal diseases.[4–13] In turn, the estimated levels of health risk have been used to classify households' water sources, combining water quality characteristics, such as protection from microbiological contamination, and accessibility and availability criteria, such as distance to the source, time needed to fetch water and continuity of water availability.

In 2015, the international community set the ambitious target of "safely managed water for all" by 2030 (Sustainable Development Goals (SDGs)—target 6.1) in an ongoing effort to address the detrimental health and social impacts of poor access to safe water for those without these services.[14] "Safely managed water" is defined by three criteria: (1) water free of faecal and priority contamination; (2) water accessible directly on premises; (3) water available when needed. This definition replaces the dichotomy between unimproved/improved drinking water source previously used for the Millennium Development Goals (MDGs) between 1990 and 2015 that did not capture adequately the levels of accessibility and availability offered by improved sources.[15]

Estimates of the health benefits of improved accessibility and continuously available water sources are rather heterogeneous across studies, but a recent review estimates that moving from an unimproved drinking water source to a continuous piped water supply on the premises could reduce the risk of diarrhoeal diseases by up to 75%.[12] Using quantitative microbial risk assessment methods (QMRA), Bivins et al. estimated that 13,700 disability-adjusted life years (DALY) are attributable to intermittency of piped water supplies in Sub-Saharan Africa (SSA) with 109,000 annual DALY worldwide.[16] The causal pathway for the above health impacts involves a change in both quality and quantity of water used by households for drinking and domestic purposes, but establishing the respective importance of each is challenging. A continuous piped water supply on the premises suggests that households do not require water handling and storage, which are known risk factors for microbial contamination.[17] Continuous tap water service and pressure also reduce the risk of contaminant ingress into the distribution network and water quality deterioration.[18] Access to water on the premises implies minimal time and effort needed from household members to collect water, and doubles or more water consumption in comparison with households using a source outside their premises.[1,8] Water quantity consumed by households was indeed shown to be stable when using a source outside the compound located as far as ~30 min of return collection journey, to decrease as the journey time became longer than 30 min, but to increase sharply when the source was located on the premises. An uninterrupted supply that users trust to perform consistently well also reduces the probability of a household occasionally reverting to other water sources of lesser quality or more demanding to fetch, and the probability of some hygiene practices being temporarily abandoned.[19,20]

[1]Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT London, UK. [2]Centre de Recherche en Hydrobiologie (CRH), Département d'Hydrologie, Section Hydrochimie, Uvira, République Démocratique du Congo. [3]Department of Infectious Disease Epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, WC1E 7HT London, UK.
*email: aurelie.jeandron@lshtm.ac.uk

To enhance consistency and comparability of SDG estimates globally, emphasis is largely placed on households' main source of water, which overlooks the widespread use of multiple sources of water by households.[21] In addition, SDG estimates are mostly based on standardised national cross-sectional household surveys and do not capture geographical inequalities at a smaller scale, although methodological innovations, such as using spatial and remote-sensing tools and data, have been proposed to provide better coverage estimates at the district level.[22]

The Democratic Republic of the Congo (DRC) is one of a few countries in which access to piped water on premises has markedly declined over the MDG period, by nearly two-thirds for the urban population, from 48 to 17% coverage.[23] In addition to increasing urbanisation, this is most likely a consequence of the deterioration in the operational capacity of the National Water Agency, Regideso, which operates piped water systems in 94 cities and secondary urban centres across the country.[24] None of the piped water supply systems in DRC are considered uninterrupted, and none of the country's population has access to safely managed water.[23] The piped water infrastructure in Uvira, the second largest town of South-Kivu province, is no exception, and is currently being rehabilitated and extended to improve access to piped water services for its ~250,000 inhabitants, in an attempt to better control cholera and other diarrhoeal diseases. Uvira is an identified hotspot for endemic cholera in Eastern DRC, from which the disease regularly spreads to less frequently affected areas of the country.[25–27]

The present analysis draws on data collected as part of an evaluation of the impact of these improvements on households' water-related practices, and the incidence of suspected cholera and other diarrhoeal diseases. Its aim is to develop and assess spatially explicit predictive models for estimating the probability of households storing contaminated drinking water and for estimating the quantity of water used at home for domestic activities, based on piped water access and distance from surface water sources. This research has the potential to provide estimates of water service level coverage at a high spatial resolution, to enable better targeted water infrastructure investments and detection of geographical inequities.

RESULTS

Built-up areas in Uvira cover ~12 km^2, with an estimated population of 254,438 inhabitants in November 2017. Population density estimated at the street level ($n = 205$) ranged from 5248 to 261,342 inhabitants per km^2 (median: 23,835; interquartile range —IQR: 17,565–33,659). In March 2018, 3685 taps were invoiced for a total of 60,925 m^3, with invoices ranging from 1 to 1800 m^3 (Supplementary Fig. S3). This represented 80.6% of the water treated by the Regideso over the same period. The total volume of water treated at the Regideso plant during the 31 days preceding the last survey day (October 12th–November 11th, 2017) amounted to 76,163 m^3, leading to an estimated volume distributed of 61,387 m^3 and a daily average of 1980 m^3. Over all built-up areas of Uvira, the 250-m tap water service indicator ranged from 5×10^{-4} to 99.3 LCPD (Supplementary Fig. S4). The maximum distances from the closest river and from the lake shore, adjusted for slope, were 5429 meters and 1812 meters, respectively, for all built-up areas.

Data from 416 households were included in the data analyses, of which 371 (89.2%) were recruited during the 2016 survey. The tap water service indicator ind250 for interviewed households ranged from 0.03 to 85 LCPD. Maximum distance from the closest river and from the lake shore were 1880 meters and 1800 meters, respectively, for interviewed households (Supplementary Fig. S4). Drinking water contamination with *E. coli* was detected in 273 samples of the 411 analysed (66.4%). In total, 301 (72.4%) households reported having collected their stored drinking water from a tap. In total, 174 (58.6%) drinking water samples collected at a tap and analysed were contaminated with *E. coli*, in comparison with the 99 (87.6%) collected from surface sources, mostly from rivers. The reported total amount of water used by households at home on the day preceding interview ranged from 10 to 460 l (median 145 l, IQR 100–205), and the number of household members present the day preceding the interview ranged from 1 to 19 (median 7, IQR 5–10). In total, 115 (27.6%) households reported having performed water-consuming activities directly at the source.

The training data set contained information on 235 households, and the test data set 181 households (Table 1).

Household drinking water contamination

Drinking water contamination is strongly associated with the reported source of collection, with drinking water collected at a surface water source having five times the odds of contamination (95% CI 2.7–9.1) of drinking water collected at a tap. The odds of a household having collected their drinking water at a tap rather than a surface source are also strongly associated with distance from the nearest river and the tap water service indicator (Supplementary Table S2).

The best-fitting model to predict drinking water contamination included tap water service indicator, distance from the nearest river and a linear interaction between distance from the nearest river and tap water service indicator. There was only weak evidence of miscalibration of the model fit to the training and testing data, and no evidence suggesting that assumption of linear relationships on the logit scale was inappropriate. Figure 1a

Table 1. Characteristics of households included in training and testing data sets.

	Training ($N = 235$) *n* (%) or median (IQR/range)	Testing ($N = 181$) *n* (%) or median (IQR/range)
Drinking water contaminated with *E. coli*	151 (64.3%)	122 (67.4%)
Missing	2 (0.9%)	3 (1.7%)
Total amount of water used in the household the previous day (in l)	145 (100–210/10–460)	144 (102–200/20–385)
Missing	1 (0.4%)	1 (0.6%)
Number of household members present the previous day		
Total	7 (5–9/1–18)	7 (5–10/1–19)
Adults and children aged 15 and older	3 (2–5/1–12)	3 (2–5/1–12)
Children under 15	4 (2–5/0–11)	4 (2–5/0–10)
Tap water service indicator (ind250) in LCPD[a]	4.5 (1.9–12.8/0–85.1)	5.1 (2.2–12.9/0–62.9)
Distance to the nearest river in m[a]	672 (355–1091/0–1727)	630 (293–1148/30–1880)
Distance to the lake in m[a]	539 (272–879/0–1680)	607 (345–896/30–1799)
Wealth quintile (index range)		
#0 (−2.23 to −1.29)	43 (18.3%)	38 (21%)
−1.26 to −0.7	49 (20.9%)	34 (18.8%)
−0.7 to −0.16	41 (17.4%)	42 (23.2%)
−0.15 to −0.79	51 (21.7%)	33 (18.2%)
#4 (0.8–5.38)	51 (21.7%)	32 (17.7%)
Missing	0 (0%)	2 (1.1%)

[a]At a spatial resolution of 34 m × 42 m

shows the observed versus predicted probability for drinking water contamination by deciles of predicted probability, for model fits to training and testing data.

Discrimination performance of the selected model on the training and test data was fair with AUCs of 0.73 and 0.75, respectively (95% CI 0.66–0.8 and 0.67–0.83, respectively) (Fig. 1b). In the selected model fitted to the training data set, the predicted probability of drinking water contamination decreases as tap water service improves, although this relationship is modified by distance from the nearest river. The largest marginal effect of increased tap water access—i.e. the predicted change in contamination probability for each LCPD unit increment in tap water indicator—is predicted to occur at the shortest distance from the nearest river (−1.6%, 95% CI −0.8% to −2.4%). Model coefficients are reported in Table 2.

When comparing AUC estimates for the entire data set based on training model fit and stratified by household wealth quintile, the discrimination ability of the model decreases noticeably between the two lowest (AUC 0.80; 95% CI 0.71–0.89) and the three highest wealth quintiles (AUC 0.66; 95% CI 0.59–0.74). However, wealth quintile inclusion in the model as an independent variable only improves very slightly its discrimination performance on the training data set (AUC 0.76; 95% CI 0.7–0.82), while reducing it noticeably on the testing data set (AUC 0.68; 95% CI 0.6–0.77).

Unsurprisingly considering the difference in odds of contamination in samples coming from tap and surface sources, the predicted probability of drinking water contamination also discriminates well between the use of a surface water or a tap as drinking water source (AUC 0.86; 95% CI 0.82–0.90).

Figure 2 shows the geographic distribution of predicted probabilities of households storing contaminated drinking water across Uvira. The map highlights the areas surrounding rivers as those at the highest risk of contamination, and the central areas of town, south of the water reservoir, as at the lowest risk.

Quantity of water used at home for domestic activities

The best-fitting model on the training data set includes household composition, tap water service indicator multiplied by the number of household members, distances from the lake and nearest river and linear interaction terms between tap water service and distance from the nearest river, and between distance to the lake and distance from the nearest river. The selected model explains >60% of the variability in reported household water consumption in the training data (adjusted R^2 0.61). The mean difference between reported and predicted values (RMSE) is nonetheless high, and is nearly equivalent to the cumulated daily consumption of two adults and one child (50 litres). There was no evidence that using linear functions for the predictors was inappropriate. Reported and fitted values for both training and testing data sets are shown in Fig. 3.

The model predicts additional reported consumption of 21.8 and 10.6 litres for each extra adult and child present in the household, respectively. Water consumption at home is predicted to increase with improved tap water access, up to a distance of ~1300 m from the nearest river, with the highest marginal effect of 0.32 litres (95% CI 0.16–0.47 litres) increased consumption per capita for households closest to the rivers. Beyond 1300 m from the nearest river, there is little evidence of an effect of improved tap water access over water consumption at home, with wide confidence intervals, including no effect at all. Household water consumption is predicted to decrease, as distance from the lake increases for households >500 m from the nearest river, at a rate of up to −12.8 litres per 100-m distance increment. The predicted effect of increasing distance from the nearest river varies with both distance to the lake and tap water service. The marginal effect is positive for households close to the lake (<500 m) and up to good access to tap water (<20 LPCD, 85th

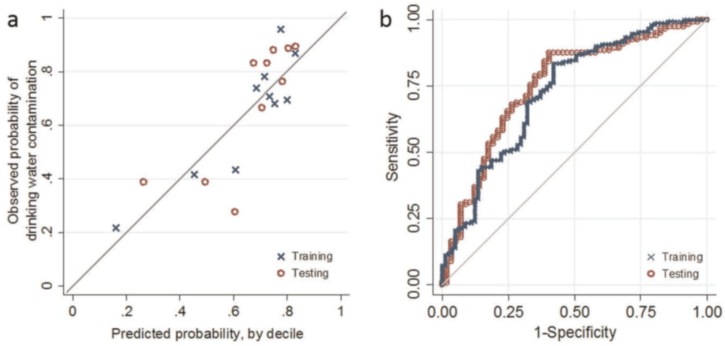


Fig. 1 Calibration and performance of the training and testing models for drinking water contamination with *E. coli*. Observed versus predicted probability for drinking water contamination by deciles of predicted probability, for model fits to training and testing data sets (**a**) and ROC for model fits to training and testing data sets (**b**).

Table 2. Selected logistic regression model for drinking water contamination with *E. coli*.

	Logit coefficient (SE)	OR (95% CI)	*p*-value[a]
Intercept	1.84 (0.41)	–	<0.001
Tap water service indicator (ind250) in LCPD	−0.1 (0.03)	0.9 (0.84–0.96)	0.002
Distance from the nearest river in m	−6.93 (4.83)[b]	0.9993 (0.9984–1.0003)	0.151
Dist. River X ind250	0.48 (0.29)[b]	1.00005 (0.99999–1.00011)	0.097

[a] *p*-value for chi-square of likelihood ratio test
[b] $\times 10^{-4}$
Pregibon's goodness-of-link test: $p = 0.36$; Hosmer–Lemeshow statistic for training and testing data: $p = 0.16$ and $p = 0.09$, respectively

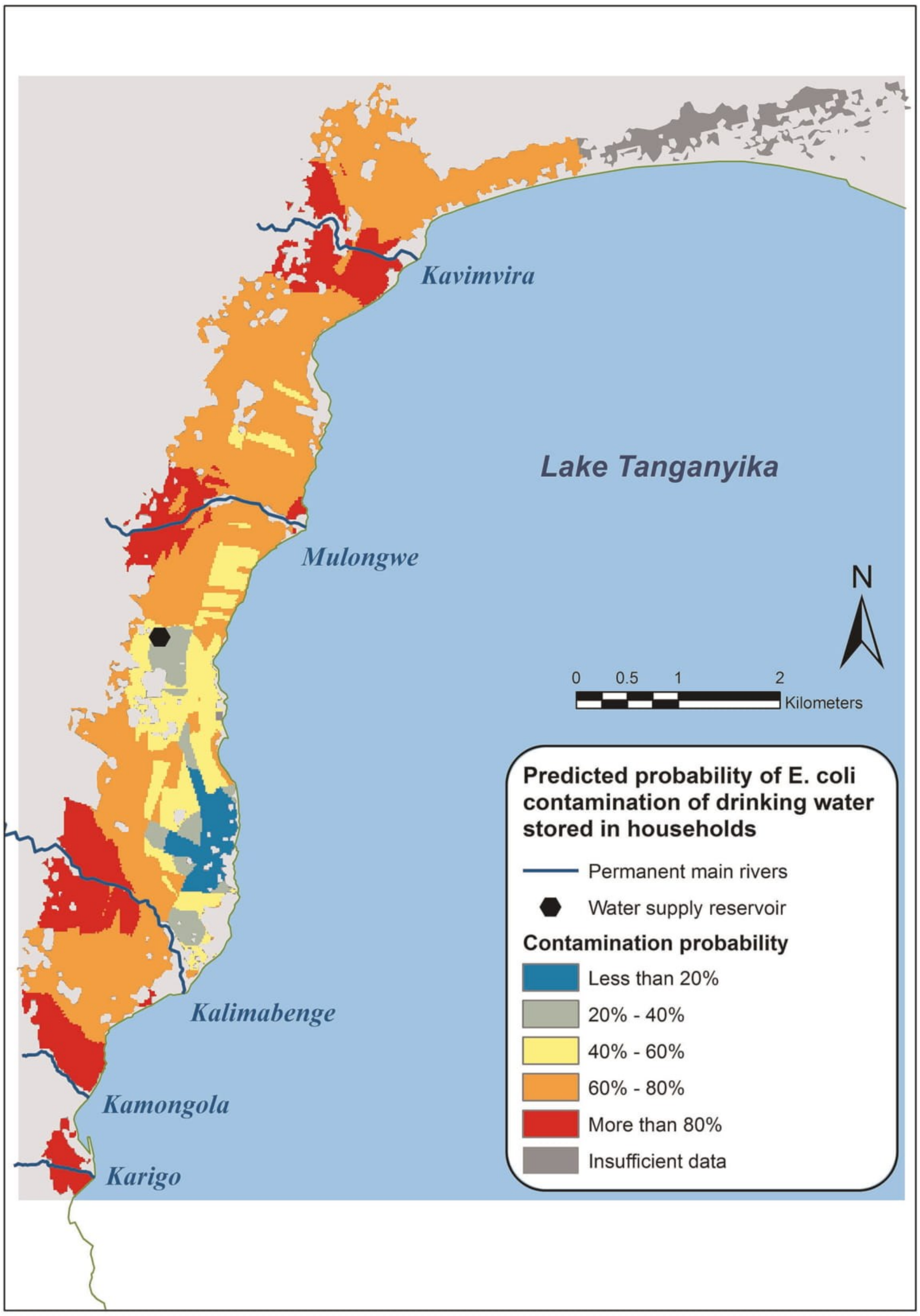


Fig. 2 Geographic distribution of predicted probabilities of households to store contaminated drinking water across Uvira.

percentile over the full data set). In other situations, household water consumption is predicted to decrease as distance from the lake increases, by up to −10.6 litres per 100-m distance increment at the median access to tap water (4.75 LPCD). Regression coefficients are presented in Table 3.

When adding wealth quintile to the selected model, its performance remains unchanged, with no evidence that wealth quintile independently affects the predicted household water consumption. When adding water use at the source, however, the model performance improves slightly (adjusted R^2 0.64; RMSE 48.7) and the model predicts a reduction in the quantity of water used at home of 34.9 litres (95% CI 19.4–50.3 litres; $p < 0.001$) by households performing water-consuming activities at the source, with little change to other model coefficients. A logistic regression model of household water use at the source suggests that tap water service, distances from the nearest river and from the lake and linear interaction between tap water service and distance from the nearest river are important predictors for such practice, with good discrimination (AUC 0.76). Over all observations in the entire data set, the predicted probability of household water use at source decreases on average by 1%, 2.9% and 2.8% as tap water service and distances from the nearest river and from the lake increase (for 1 LPCD increment and 100-m increments, respectively).

Figure 4a shows the predicted amount of water used for domestic activities performed at home, by a household composed of three adults or children over 15, and four children under 15, across Uvira. Households in areas poorly served by the existing water supply network and distant from both the lake and rivers are predicted to report using <140 litres per day at home. Areas near the rivers and the lake are also highlighted as areas with low predicted water consumption at home. These areas are, however, overlapping with those with a high probability of water use at the source (Fig. 4b).

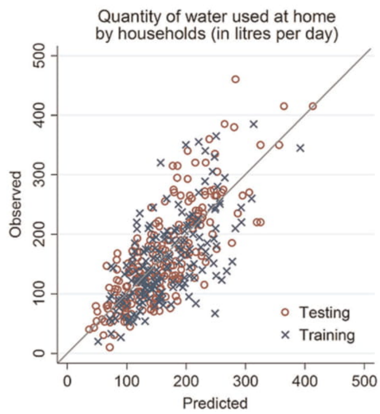


Fig. 3 Reported and predicted daily quantity of water used at home by households in litres.

Sensitivity analysis

When fitting similarly formulated models to the training data with indicators constructed over 500-m or 750-m radii and indicators based on tap location and density without invoicing data at radii 250 m, 500 m and 750 m, predictions of drinking water contamination with *E. coli* and predictions of the quantity of water used by a household show similar performance in comparison with ind250 (Supplementary Table S3).

DISCUSSION

Our study used household location relative to surface water sources and a measure of their access to tap water to predict the probability of microbial contamination of household drinking water and the quantity of water consumed domestically. These predictions were developed at the scale of Uvira, the second largest town of South-Kivu province in DRC and summarised into high-resolution maps of this secondary urban centre. The models were developed using a subset of households' survey data collected in October 2017, and their predictive performance assessed against the remaining portion of survey data.

The drinking water model predicted whether household-stored contaminated drinking water with fair discrimination performed better for the poorest households. Probability of contamination was strongly predicted by tap water service and distance from the nearest river, with the lowest probabilities predicted to occur for households with a better tap water service and further away from the rivers. The same improvement in tap water service was predicted to reduce the probability of contamination in areas <250 m from the nearest river by twice as much as in areas >1250 m from the nearest river. The same model was also shown to predict reasonably well the type of source used by the household for drinking water (surface water or tap), which was associated with very different risks of contamination.

The total quantity of water reported to be used at home by households was predicted by its demographic composition, tap water service and distances from both the lake and the nearest river. An adult member was predicted to report more than double the increase in daily household water consumption than a child, with 21.8 litres versus 10.6 litres, respectively. The marginal effect of tap water service improvement was predicted to be small, and decreased as distance from the nearest river increased. On average over the study households, an increase of one LCPD in tap water service represented only 1.1-litre increase in the reported total household consumption, corresponding to an average of 0.15 litre per household member. The lower amount of water consumed at home by households located close to the rivers and

Table 3. Selected linear regression model for predicting the amount of water used at home by households for domestic activities.

	Coefficient (SE)	95% CI	p-value[a]
Intercept	−12.68 (15.26)	−42.74–17.38	0.407
The number of household members aged 15 or more	21.79 (1.92)	18.01–25.58	<0.001
The number of household members under 15	10.62 (1.35)	7.96–13.28	<0.001
Tap water service indicator (ind250) × number of household members (in litres per day)	0.32 (0.08)	0.16–0.47	<0.001
Distance from the nearest river (in m)	0.08 (0.02)	0.04–0.11	<0.001
Distance from the lake (in m)	0.04 (0.02)	0.01–0.08	0.006
Dist. River × Dist. Lake	−0.87 (0.21)[b]	−1.28 to −0.45[b]	<0.001
Dist. River × (ind250 × number of household members)	−2.37 (0.66)[b]	−3.68 to −1.07[b]	<0.001

[a] p-value for χ^2 of likelihood ratio test
[b] $\times 10^{-4}$
Pregibon's goodness-of-link test for training and testing data: $p = 0.58$ and $p = 0.74$, respectively. RMSE for training and testing data: 57.8 and 50.6, respectively

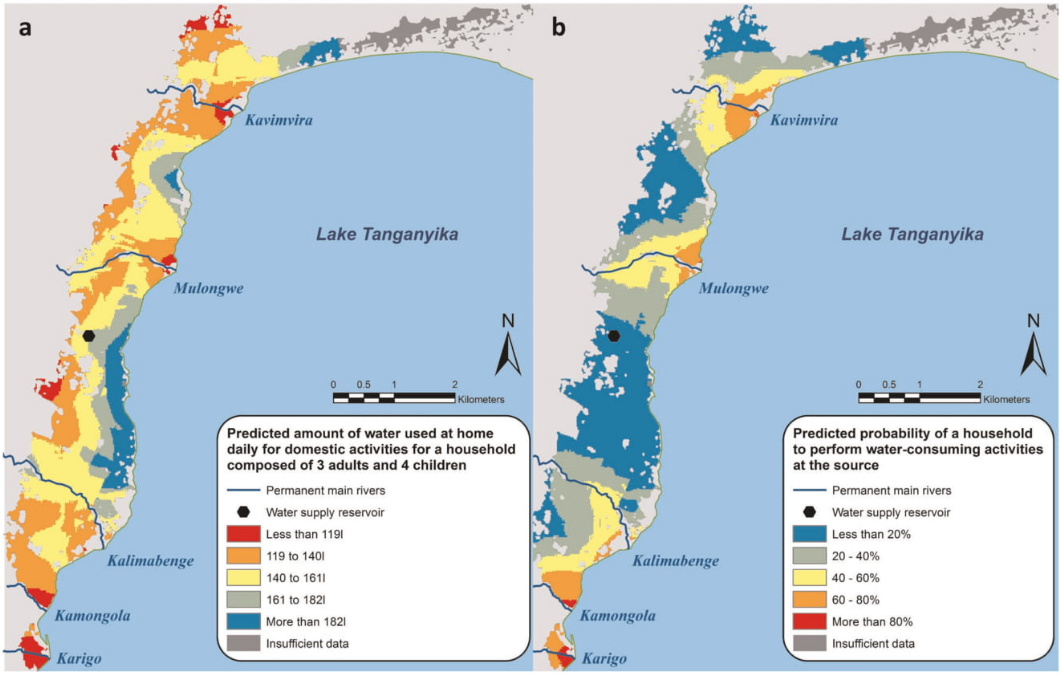


Fig. 4 Geographic distribution of predicted quantity of water used at home for domestic activities and of the predicted probability of performing water-consumingactivities at the water source. Geographic distribution of predicted quantity of water used at home for domestic activities for a standardised household composed of three adults and four children (**a**) and geographic distribution of the predicted probability of performing water-consuming activities at the water source (**b**).

lake appears associated with a higher probability of performing activities directly at the source.

The tap water service indicator used in the present analysis was built upon the assumption that a higher geographical density of water taps in the vicinity of a household represents a higher probability for that household of accessing and using it. By weighing the spatial density with the water volume invoiced and population density, the proxy indicator was then adjusted for a potential limited supply of water at the taps, because of unreliability, intermittency or because of "competition" with other households for access. Within these assumptions, the sensitivity analysis results of the indicator estimated over 500-m and 750-m radii first suggest that household behaviours related to water sources are influenced by spatial density of taps at least up to 750 m. The results obtained with the indicators ignoring the actual water volume invoiced for each tap then suggest that either reliability or intermittency of water supply at the tap has little bearing over these behaviours, or that the information on reliability or intermittency of a tap is already included in the measure of tap density. Taps are indeed more likely to be subscribed to and active in areas where the supply is reliable, which results in a spatial clustering captured by the tap-density measure. The lack of influence of wealth quintile in model predictions may also be explained by the strong correlation between wealth index and tap water service indicator (Pearson's correlation coefficient $r = 0.49$).

The good discrimination ability of the drinking water model indicates that factors exogenous to households play an important role in stored drinking water contamination. Assuming that households do have a preference for tap water as a drinking water source due to quality/health concerns, distance from the nearest river influences how access to a tap actually translates into tap use as a drinking water source, especially when tap access is relatively poor. The contamination OR between drinking water collected at a tap or at surface sources by households in Uvira (0.2; 95% CI 0.11–0.37) relates well to that reported by Bain et al. in their review of faecal contamination of drinking water in low- and middle-income countries.[28] Beyond encouraging its use as drinking water source, access to a closer and more reliable tap could plausibly reduce the risk of post-collection contamination linked to transportation and storage of tap water. Another review indeed determined the odds of contamination for stored household water more than double those of contamination of source water, even for piped water supplies.[29] The better performance of the drinking water model for poorer households implies that these households are even more dependent on exogenous factors, with few coping mechanisms. Even in the case of poor tap water service, wealthier households may have a stronger preference for piped water as a result of higher education achievements or social pressure, and use coping mechanisms—better storage containers, point-of-use water treatment and payment for their water to be collected for them at a tap for example.[20]

Predicting household water consumption at home for domestic activities provides weaker evidence of the relationship between proxy measures for access to tap water or surface water and the quantity of water used. The nonlinear relationship between access to water and water use has been highlighted before, especially outside of rural contexts with few remote sources.[30] When multiple sources are available, water used for different purposes is indeed a heterogeneous good, with characteristics—perceived quality, convenience of access and monetary and opportunity costs—influencing households collection and use behaviours depending on their education and capacity to pay.[31] Even in a setting like Uvira, where the diversity of sources is limited—tap, rivers or lake—simple measures of access are unlikely to capture enough of the preferences guiding households' choices in source use, and how access to the sources chosen impacts the quantity of water used at home. By attempting to replace in the model the two distance variables by a single one representing distance to the nearest surface water source, the predictive performance of the model decreased sufficiently to highlight different households' behaviours towards the two surface water sources. This could reflect a preference for one or the other surface source at equal distance, either related to the distance measure (e.g. distance not capturing the difficulties to access the lake shore in parts of town) or related to the water characteristics. The predictive model reached here is nevertheless compatible with the concept of a water consumption plateau put forward by Cairncross et al., according to which water consumption of households remains stable when the source is located between 5 and 30 min from the house, and decreases only when the location is further away than 30 min, while increasing massively as the source is brought closer than 5 min (source within premises).[2] Our model suggests that water consumption only decreases with increasing distance when the nearest river or the lake are further away than 500 meters, which roughly corresponds to 30 min (20 min there and back, with 10 min to fill up containers). The results, however, do not reflect the sharp increase in water consumption observed when a source is accessible on household premises, as the tap water indicator is unlikely to distinguish between households with an active tap within their compound and those with an active tap at a short distance but outside their compound, with less convenient access.

Our study has several limitations. The data used in the analysis were collected at a single point in time, during the rainy season. Although seasonality has a limited impact on the availability of surface and tap water, it was previously reported that tap water supply interruptions are more frequent in Uvira during the dry season versus the rainy season due to more frequent power cuts that affect the water treatment plant operations.[32] Heavy rains on the steep slopes overlooking Uvira also cause rivers and lake along the shores to become extremely muddy for several hours, deterring people from using the surface water. In addition, our data do not support the inclusion of areas of Uvira located more than 2 km from a river. A single measure of the outcomes will also not capture daily or weekly differences in activities, which may give rise to important variations in daily water consumption, with some activities not carried out daily. Reported water quantity use is also likely to suffer from a substantial random error. The observational nature of the relationships established by the two models should also be highlighted, and a causal relationship between the predictors and the outcomes should not be assumed, even though some degree of causality is plausible, especially for drinking water. Finally, the predictive nature of our methodology is unsuitable for untangling further the multiple factors affecting water use by households, especially when not only considering drinking water.

Geographical predictions of improved drinking water or improved sanitation coverage within SSA countries were previously performed at the district level or rural community level.[33,34] Multi-country meta-analyses also provide estimates of contamination levels for drinking water collected at different sources or different levels of tap water service.[35] In addition, several methods were developed to estimate and predict water demand for urban piped water systems in low- or middle-income countries, with a focus on engineering and financial planning.[31] However, we could not identify previous attempts to spatially predict drinking water quality or domestic water quantity used at the household level that provides valuable information for water supply improvements targeting health benefits. We believe that our approach and the results obtained here in Uvira warrant a further exploration of their value in other contexts. Should it be generalisable, our approach could allow identification of priority intervention areas for water supply improvements in urban settings, without implementing costly household surveys, by geolocalising taps, possibly along with the volume of water consumed at a single point in time, geolocalising alternative sources of water and population data at a relatively small scale. The present results also support their use to investigate the possible relationship between tap water access and observed time and space patterns of health outcomes incidence, such as suspected and confirmed cholera.

METHODS

Study area

Uvira extends ~10 km along the northernmost shores of Lake Tanganyika, <2 km inland and is crossed by five permanent rivers.

The population, estimated at 254,000 inhabitants in November 2017, relies on both surface water sources and the tap water system managed by the national water agency Regideso. Water from the river Mulongwe is drawn upstream of inhabited areas and treated at the Regideso water treatment plant before being fed into a single 1600 m^3 reservoir, from which it is distributed to private and shared taps by gravity. The current distribution system fails to serve adequately the taps located further away from the reservoir or higher in altitude, and the daily amount of water distributed is irregular. There are no wells or boreholes due to unfavourable geological terrain.

Uvira's water supply infrastructure is undergoing refurbishment and expansion since September 2018 through a project funded by the European Union (EU), the French Development Agency (AFD) and the Veolia Foundation in partnership with the Regideso.

Household data

The household data were collected as part of two surveys of household water-related practices, conducted in October 2016 and October–November 2017.

Recent, reliable data on the location and functionality of Regideso taps were unavailable at the time households were sampled in October 2016. Therefore, to establish a cohort of households representing a wide range of access to tap water in Uvira, and in the absence of a household-sampling frame, a two-stage random spatial sampling method was used based on a piped water availability index. Details of this sampling method are given in Supplementary Information, along with households' selection and enrolment methods. During the second survey implemented in 2017, the buildings sampled and georeferenced during the first survey were revisited. If the household inhabiting the building was different, the same enrolment process was used with the new family.

During both surveys, households were interviewed about water-related practices at home. This included the amount of fresh (as in not recycled from a previous activity) water used the previous day for various domestic activities using a visual aid, the number of adults and children present that day and water use at the source. The amounts of fresh water used at home for bathing, laundry, dishwashing, food and produce rinsing, dwelling cleaning, handwashing and drinking were added up into a total amount of fresh water used at home for domestic activities. Water used at home to prepare food or items for sale or to render a paid-for service was excluded from the total.

During each interview in 2017, a 150-ml sample of stored drinking water was collected in a sterile sample Whirl-pak® ThioBag® (Nasco, Fort Atkinson, WI) containing 30 g of sodium thiosulfate. The participant was requested to provide the water that would be used for drinking at the time of the interview, with the utensils usually used for serving such drinking water and reported where the stored drinking water had been collected.

Samples were brought back daily in cool bags to the Centre de Recherche Hydrobiologique (CRH) in Uvira and analysed within 6 h of collection. Turbidity was measured with a digital turbidimeter, and the volume of water filtered adjusted to aim for a turbidity of 3NTU once diluted with sterile water to reach a volume of 100 ml. Between 5 and 100 ml of each sample was filtered in sterile conditions through a 0.47-µm filter, and the filter was then aerobically incubated on sterile pads saturated with mColiblue24 broth (Hach Co, Frederick, MD) for 24 h at 35 °C in a portable incubator. The number of blue colonies grown on the filter was then counted and multiplied by the appropriate factor to obtain the number of colony-forming units (CFU) *Escherichia coli* per 100 ml. Although there is still debate about the relationship between CFU *E. coli* per 100 ml in drinking water and health risks, we used a single cut-off of one CFU *E. coli* per 100 ml to define a binary outcome of contaminated/non-contaminated drinking water.[36]

A household wealth index based on ownership of durable items and dwelling characteristics was constructed and classified into quintiles (details in Supplementary Information).

For each household interviewed, the shortest distances from one of the town's five main rivers and from Lake Tanganyika were computed with a slope adjustment based on Tobler's hiking function.[37]

Tap water service indicator

Data on the daily volume of water treated were collected from the register held at the Regideso water treatment plant. Unique identifiers of active tap connections were retrieved from the Regideso customers database, along with the volume invoiced in March 2018, and each of them geolocalised.

Based on these data, an index of tap water availability was constructed. In brief, a kernel density function with a radius of 250 m around each tap was used to combine data on the water volume invoiced for each functional tap with information on population density. This produced a smooth "surface" with a resolution of ~34 m × 42 m of tap water availability across Uvira in litres per capita per day (LCPD). This indicator was extracted at interviewed households' location and used as a continuous variable.

Population data sources, delineation methods for built-up areas and the construction of the tap water availability index are detailed in Supplementary Information.

Statistical methods

Logistic regression was used to assess whether the tap water service indicator, distance to the closest river and distance to the lake shore were predictive of the probability of stored household drinking water being contaminated. The model was developed on a random sample of 60% of the data set records (training), and tested on the remaining 40%. Model selection followed a hierarchical backward strategy, starting with a full model, including the three variables as continuous and all possible two- and three-way linear interactions.[38] Model terms were eliminated starting with the least significant, to identify the model with the lowest value of Akaike's information criterion (AIC).

The Hosmer–Lemeshow statistic and a plot of predicted against observed probabilities by decile of predicted probability were used to assess the model fit and calibration to the training and testing data sets. Lowess plots of standardised residuals and Pregibon's goodness-of-link test were used to assess whether assuming a linear relationship was appropriate.[39] The discrimination of the model was assessed using receiver–operator curves and derived area under the curve (ROC and AUC). The model's discriminative ability was examined by household wealth quintile, to assess whether model performance varied with household wealth. The importance of reported drinking water source as a possible explanatory variable for contamination was also explored. The predicted probability of household drinking water contamination was then mapped using the coefficients of the selected model derived from the training data set.

A similar approach was used to model the total quantity of water used within a household using linear rather than logistic regression and including demographic composition of the household (the number of children, number of adults) as additional covariates. Tap water service being expressed as a quantity of tap water per capita, we used the indicator multiplied by the number of household members present the previous day. Lowess plots of standardised residuals and Pregibon's goodness-of-link test were used to assess whether assuming a linear relationship was appropriate, while adjusted R^2 and root-mean-square error (RMSE) described the model fit to the data. Wealth quintile and reporting of having performed water-consuming activities (laundry, dishwashing and bathing) at the source were added separately as independent variables to the selected model in order to investigate whether they improved the model predictions. We also investigated to what extent water use at the source was predicted by tap water service, distance from the nearest river and distance from the lake shore. The predicted quantity of water used per household, for households having the median household composition, was then mapped.

To avoid undue influence of extreme outliers on model parameters, four records were excluded from the analysis: two households located at more than 4000 meters from the closest river and two households for which more than 20 members were reported present the day preceding the interview. To avoid extrapolating model estimates, these exclusions were taken into account by limiting the mapping to areas less than 2000 meters from the nearest river.

A sensitivity analysis was performed by replacing the tap water service indicator constructed at a 250-m radius (ind250) with indicators constructed at radius 500 m or 750 m (ind500 and ind750). A tap-density indicator at 250 m, 500 m and 750 m (dens250, dens500 and dens750) was also used. Expressed in number of taps per 1000 people, this indicator was constructed with a constant weight applied to all taps, ignoring individual taps invoicing and the possible variations in water availability and tap reliability invoicing data may represent.

Data preparation and tap water service indicator construction were performed with ArcGIS ArcMap 10.3 (ESRI, Aylesbury, UK) and R,[40] in particular the R package "sparr".[41] Data were analysed with STATA 14.2 (StataCorp, College Station, TX).

Ethical considerations

Household interviews were only performed after written consent to participate was obtained, in accordance with study approvals from the ethics committees of the School of Public Health at the University of Kinshasa, Democratic Republic of the Congo (ESP/CE/088c/2017), and of the London School of Hygiene and Tropical Medicine, United Kingdom (No. 10603). The study is part of a broader evaluation of the impacts of tap water supply improvements on cholera and other diarrhoeal diseases in Uvira registered at clinicaltrials.gov (Reference: NCT02928341). All the data were anonymised before analysis.

DATA AVAILABILITY

Anonymised data are available from the corresponding author upon request.

Received: 26 June 2019; Accepted: 14 November 2019;
Published online: 16 December 2019

REFERENCES

1. White, G. F., Bradley, D. J. & White, A. U. *Drawers of Water; Domestic Water Use in East Africa* (University of Chicago Press, 1972).
2. Cairncross, S. & Feachem, R. G. *Environmental Health Engineering in the Tropics: Water, Sanitation and Disease Control* 3rd edn (Taylor & Francis Group, Routledge, 2019).
3. Feachem, R., Bradley, D., Garelick, H. & Mara, D. D. *Sanitation and Disease: Health Aspects of Excreta and Wastewater Management*. (John Wiley & Sons, 1983).
4. Clasen, T., Schmidt, W. P., Rabie, T., Roberts, I. & Cairncross, S. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *Br. Med. J.* **334**, 782 (2007).
5. Esrey, S. A., Feachem, R. G. & Hughes, J. M. Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities. *Bull. World Health Organ* **63**, 757–772 (1985).
6. Esrey, S. A., Potash, J. B., Roberts, L. & Shiff, C. Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. *Bull. World Health Organ* **69**, 609–621 (1991).
7. Fewtrell, L. et al. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect. Dis.* **5**, 42–52 (2005).
8. Overbo, A., Williams, A. R., Evans, B., Hunter, P. R. & Bartram, J. On-plot drinking water supplies and health: A systematic review. *Int. J. Hyg. Environ. Health* **219**, 317–330 (2016).
9. Pruss-Ustun, A. et al. Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: an updated analysis with a focus on low- and middle-income countries. *Int J. Hyg. Environ. Health*. https://doi.org/10.1016/j.ijheh.2019.05.004 (2019).
10. Stelmach, R. D. & Clasen, T. Household water quantity and health: a systematic review. *Int J. Environ. Res. Public Health* **12**, 5954–5974 (2015).
11. Waddington, H., Snilstveit, B., White, H. & Fewtrell, L. *Water, Sanitation and Hygiene Interventions to Combat Childhood Diarrhoea in Developing Countries* (New Delhi: International Initiative for Impact Evaluation, 2009).
12. Wolf, J. et al. Impact of drinking water, sanitation and handwashing with soap on childhood diarrhoeal disease: updated meta-analysis and meta-regression. *Tropical Med. Int. Health* **23**, 508–525 (2018).

13. Wolf, J. et al. Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression. *Trop. Med. Int Health* **19**, 928–942 (2014).
14. World Health Organization & UNICEF. *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. (World Health Organization, Geneva, 2017).
15. Clasen, T. F. Millennium development goals water target claim exaggerates achievement. *Tropical Med. Int. health* **17**, 1178–1180 (2012).
16. Bivins, A. W. et al. Estimating infection risks and the global burden of diarrheal disease attributable to intermittent water supply using QMRA. *Environ. Sci. Technol.* **51**, 7542–7551 (2017).
17. Wright, J., Gundry, S. & Conroy, R. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. *Trop. Med. Int. Health* **9**, 106–117 (2004).
18. Kumpel, E. & Nelson, K. L. Comparing microbial water quality in an intermittent and continuous piped water supply. *Water Res.* **47**, 5176–5188 (2013).
19. Kumpel, E. & Nelson, K. L. Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality. *Environ. Sci. Technol.* **50**, 542–553 (2016).
20. Majuru, B., Suhrcke, M. & Hunter, P. R. How do households respond to unreliable water supplies? A Systematic Review. *Int. J. Environ. Res. Public Health* **13**, https://doi.org/10.3390/ijerph13121222 (2016).
21. Elliott, M. et al. Addressing how multiple household water sources and uses build water resilience and support sustainable development. *npj Clean. Water* **2**, 6 (2019).
22. Thomas, E., Andrés, L. A., Borja-Vega, C. & Sturzenegger, G. *Innovations in WASH impact measures: water and sanitation measurement technologies and practices to inform the Sustainable Development Goals*. (The World Bank, 2018).
23. World Health Organization & UNICEF. *Progress on drinking water, sanitation and hygiene: 2015 update and MDG assessment*. (World Health Organization, Geneva, 2015).
24. African Ministers' Council on Water (AMCOW). *Water Supply and Sanitation in the Democratic Republic of the Congo: Turning Finance into Services for 2015 and beyond* (Water and Sanitation Program, 2011).
25. Bompangue, D. et al. Lakes as source of cholera outbreaks, Democratic Republic of Congo. *Emerg. Infect. Dis.* **14**, 798–800 (2008).
26. Brecht, I. et al. Recurrent cholera outbreaks, Democratic Republic of the Congo, 2008–2017. *Emerg. Infect. Dis. J.* **25**, 856 (2019).
27. Jeandron, A., Cumming, O., Rumedeka, B. B., Saidi, J. M. & Cousens, S. Confirmation of cholera by rapid diagnostic test amongst patients admitted to the cholera treatment centre in Uvira, Democratic Republic of the Congo. *PLoS ONE* **13**, e0201306 (2018).
28. Bain, R. et al. Fecal contamination of drinking-water in low- and middle-income countries: a systematic review and meta-analysis. *PLOS Med.* **11**, e1001644 (2014).
29. Shields, K. F., Bain, R. E. S., Cronk, R., Wright, J. A. & Bartram, J. Association of supply type with fecal contamination of source water and household stored drinking water in developing countries: a bivariate meta-analysis. *Environ. Health Perspect.* **123**, 1222–1231 (2015).
30. Evans, B. E. et al. *Public Health and Social Benefits of at-house Water Supplies* Vol. 53. (Department for International Development, 2013).
31. Nauges, C. & Whittington, D. Estimation of water demand in developing countries: an overview. *World Bank Res. Obser* **25**, 263–294 (2010).
32. Jeandron, A. et al. Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo. *PLoS Med.* **12**, e1001893 (2015).
33. Oswald, W. E. et al. Prediction of low community sanitation coverage using environmental and sociodemographic factors in Amhara Region, Ethiopia. *Am. J. Trop. Med. Hyg.* **95**, 709–719 (2016).
34. Pullan, R. L., Freeman, M. C., Gething, P. W. & Brooker, S. J. Geographical inequalities in use of improved drinking water supply and sanitation across Sub-Saharan Africa: mapping and spatial analysis of cross-sectional survey data. *PLOS Med.* **11**, e1001626 (2014).
35. Bain, R., Johnston, R., Mitis, F., Chatterley, C. & Slaymaker, T. Establishing sustainable development goal baselines for household drinking water, sanitation and hygiene services. *Water* **10**, 1711 (2018).
36. Gruber, J. S., Ercumen, A. & Colford, J. M. Jr. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PloS ONE* **9**, e107429–e107429 (2014).
37. Tobler, W. *Three Presentations on Geographical Analysis and Modeling: Non- Isotropic Geographic Modeling; Speculations on the Geometry of Geography; and Global Spatial Analysis (93-1)*, https://escholarship.org/uc/item/05r820mz (1993).
38. Kleinbaum, D. G., Klein, M. & Pryor, E. R. *Logistic Regression: A Self-learning Text* 3rd edn (Springer, 2010).
39. Pregibon, D. Goodness of link tests for generalized linear models. *J. R. Stat. Soc. Ser. C. (Appl. Stat.)* **29**, 14–15 (1980).
40. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2013).
41. Davies, T. M., Hazelton, M. L. & Marshall, J. C. sparr: analyzing spatial relative risk using fixed and adaptive kernel density estimation in R. *J. Stat. Softw.* **1**. https://doi.org/10.18637/jss.v039.i01 (2011).

ACKNOWLEDGEMENTS

The study was conducted with the funding of the Veolia Foundation (PF9233) and the French Development Agency (ERS-EVA 364-2015). We are grateful to the funders for their valuable comments on the final paper. We would like to thank all study participants for their valuable time answering our questions, and to the entire team of interviewers and supervisors in Uvira for their hard work. We are also grateful to the staff of the Centre de Recherche en Hydrobiologie for a fruitful collaboration and of OXFAM GB in DRC for their continuing support in the field.

AUTHOR CONTRIBUTIONS

A.J., O.C. and S.C. conceptualised the household survey methodology. A.J. designed survey tools and curated the data. A.J. and L.K. managed laboratory resources and performed the water quality analyses. A.J. and S.C. developed the model and performed the formal analyses. The paper first draft was prepared by A.J. and substantially reviewed and revised by all authors.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at https://doi.org/10.1038/s41545-019-0047-9.

Correspondence and requests for materials should be addressed to A.J.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

4.7 Chapter 4 - References

1. Esrey SA, Potash JB, Roberts L, Shiff C. *Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma*. **Bull World Health Organ**. 1991;69(5):609-621.
2. Shields KF, Bain RES, Cronk R, Wright JA, Bartram J. *Association of Supply Type with Fecal Contamination of Source Water and Household Stored Drinking Water in Developing Countries: A Bivariate Meta-analysis*. **Environ Health Perspect**. 2015;123(12):1222-1231.
3. Cassivi A, Guilherme S, Bain R, Tilley E, Waygood EOD, Dorea C. *Drinking water accessibility and quantity in low and middle-income countries: A systematic review*. **Int J Hyg Environ Health**. 2019;222(7):1011-1020.
4. Overbo A, Williams AR, Evans B, Hunter PR, Bartram J. *On-plot drinking water supplies and health: A systematic review*. **International Journal of Hygiene and Environmental Health**. 2016;219(4):317-330.
5. Stelmach RD, Clasen T. *Household water quantity and health: a systematic review*. **Int J Environ Res Public Health**. 2015;12(6):5954-5974.
6. Achore M, Bisung E, Kuusaana ED. *Coping with water insecurity at the household level: A synthesis of qualitative evidence*. **Int J Hyg Environ Health**. 2020;230:113598.
7. Majuru B, Suhrcke M, Hunter PR. *How Do Households Respond to Unreliable Water Supplies? A Systematic Review*. **Int J Environ Res Public Health**. 2016;13(12).
8. Bivins AW, Sumner T, Kumpel E, et al. *Estimating Infection Risks and the Global Burden of Diarrheal Disease Attributable to Intermittent Water Supply Using QMRA*. **Environmental Science & Technology**. 2017;51(13):7542-7551.
9. Kumpel E, Nelson KL. *Mechanisms affecting water quality in an intermittent piped water supply*. **Environ Sci Technol**. 2014;48(5):2766-2775.
10. Kumpel E, Nelson KL. *Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality*. **Environmental Science & Technology**. 2016;50(2):542-553.
11. Himelein K, Eckman S, Murray S, Bauer J. *Alternatives to full listing for second stage sampling: Methods and implications*. **Statistical Journal of the IAOS**. 2017;33:701-718.
12. World Health Organization. Water S, Health T. Emerging issues in water and infectious disease. Geneva: World Health Organization; 2002.
13. Lightfoot NF, Maier E. *Microbiological analysis of food and water: guidelines for quality assurance.* Elsevier; 1998.
14. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2013.
15. Elliott M, Foster T, MacDonald MC, Harris AR, Schwab KJ, Hadwen WL. *Addressing how multiple household water sources and uses build water resilience and support sustainable development.* **npj Clean Water**. 2019;2(1):6.
16. Sommer M, Ferron S, Cavill S, House S. *Violence, gender and WASH: spurring action on a complex, under-documented and sensitive topic*. **Environment and Urbanization**. 2014;27(1):105-116.
17. Graham JP, Hirai M, Kim S-S. *An Analysis of Water Collection Labor among Women and Children in 24 Sub-Saharan African Countries*. **PLOS ONE**. 2016;11(6):e0155981.
18. Refugees UNHCf. *Handbook for emergencies.* United Nations High Commissioner for Refugees; 2007.

19. Cairncross S, Cliff JL. *Water use and health in Mueda, Mozambique*. **Transactions of the Royal Society of Tropical Medicine and Hygiene**. 1987;81(1):51-54.

20. LeChevallier MW, Evans TM, Seidler RJ. *Effect of turbidity on chlorination efficiency and bacterial persistence in drinking water*. **Applied and environmental microbiology**. 1981;42(1):159-167.

21. Jensen PK, Ensink JH, Jayasinghe G, Van Der Hoek W, Cairncross S, Dalsgaard A. *Domestic transmission routes of pathogens: the problem of in-house contamination of drinking water during storage in developing countries*. **Tropical Medicine & International Health**. 2002;7(7):604-609.

22. Wright J, Gundry S, Conroy R. *Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use*. **Tropical medicine & international health**. 2004;9(1):106-117.

23. Bompangue D, Giraudoux P, Handschumacher P, et al. *Lakes as source of cholera outbreaks, Democratic Republic of Congo*. **Emerg Infect Dis**. 2008;14(5):798-800.

5

Water supply interruptions and suspected cholera incidence: a time-series regression

This analysis was performed at the very beginning of the present research, as a number of cholera control stakeholders in the eastern provinces of the Democratic Republic of the Congo (DRC) hypothesized that cholera outbreaks in the endemic towns of Uvira or Kalemie were triggered by interruptions in tap water supply. Contributing to the growing evidence of an increased risk of diarrhoeal diseases amongst populations using intermittent tap water supplies was also relevant to the simultaneous definition of sustainable development goals (SDG) for water access. Using a lagged approach previously largely used for environmental exposure to weather or pollution over non-infectious outcomes, the objectives were to assess whether daily variations in tap water produced and distributed in Uvira were associated with an increase in admissions at the cholera treatment centre (CTC). This analysis was based on data retrospectively collected and compiled since 2009 from the CTC and the Regideso water treatment station.

Research paper 3

Title: Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo

Authors: Aurelie Jeandron[1], Jaime Mufitini Saidi[2], Alois Kapama[2], Manu Burhole[3], Freddy Birembano[3], Thierry Vandevelde[4], Antonio Gasparrini[5], Ben Armstrong[6], Sandy Cairncross[1], Jeroen H. J. Ensink[1]

[1] Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[2] Ministère de la Santé Publique, Division Provinciale de la Santé Publique, District Sanitaire d'Uvira, Uvira, Sud-Kivu, République Démocratique du Congo

[3] Ministère de la Santé Publique, Division Provinciale de la Santé Publique, Bukavu, Sud-Kivu, République Démocratique du Congo

[4] Fondation Veolia, Nanterre, France

[5] Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[6] Department of Social and Environmental Research, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, United Kingdom

Status: Published in PLoS Medicine in November 2015

Citation: Jeandron A, Saidi JM, Kapama A, Burhole M, Birembano F, Vandevelde T, et al. Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo. PLoS Med. 2015;12(10):e1001893.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	173940	**Title**	Ms
First Name(s)	Aurélie		
Surname/Family Name	Jeandron		
Thesis Title	Tap water access and its relationship with cholera and other diarrhoeal diseases in an urban, cholera-endemic setting in the Democratic Republic of the Congo		
Primary Supervisor	Prof. Simon Cousens		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	PLOS MEDICINE		
When was the work published?	OCTOBER 2015		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	NA		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	Conceived and designed the experiments; analyzed the data; wrote the first draft of the manuscript; agree with the manuscript's results and conclusions; collected, checked and managed the data.

SECTION E

Student Signature	Aurelie Jeandron
Date	21/05/2020

Supervisor Signature	Simon Cousens
Date	22/05/2020

PLOS | MEDICINE

RESEARCH ARTICLE

Water Supply Interruptions and Suspected Cholera Incidence: A Time-Series Regression in the Democratic Republic of the Congo

Aurélie Jeandron[1]*, Jaime Mufitini Saidi[2], Alois Kapama[2], Manu Burhole[3], Freddy Birembano[3], Thierry Vandevelde[4], Antonio Gasparrini[5], Ben Armstrong[6], Sandy Cairncross[1], Jeroen H. J. Ensink[1]

1 Environmental Health Group, Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom, **2** Ministère de la Santé Publique, Division Provinciale de la Santé Publique, District Sanitaire d'Uvira, Uvira, Sud-Kivu, République Démocratique du Congo, **3** Ministère de la Santé Publique, Division Provinciale de la Santé Publique, Bukavu, Sud-Kivu, République Démocratique du Congo, **4** Fondation Veolia, Nanterre, France, **5** Department of Medical Statistics, Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, United Kingdom, **6** Department of Social and Environmental Research, Faculty of Public Health and Policy, London School of Hygiene & Tropical Medicine, London, United Kingdom

* aurelie.jeandron@lshtm.ac.uk

Citation: Jeandron A, Saidi JM, Kapama A, Burhole M, Birembano F, Vandevelde T, et al. (2015) Water Supply Interruptions and Suspected Cholera Incidence: A Time-Series Regression in the Democratic Republic of the Congo. PLoS Med 12 (10): e1001893. doi:10.1371/journal.pmed.1001893

Academic Editor: Clarissa Brocklehurst, University of North Carolina at Chapel Hill, UNITED STATES

Received: January 21, 2015

Accepted: September 17, 2015

Published: October 27, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: AJ received funding from the Agence Française de Développement (grant RCH 310-2013) for this work. AG was supported by a Methodology Research fellowship from UK Medical Research Council (Grant ID: G1002296).The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Abstract

Background

The eastern provinces of the Democratic Republic of the Congo have been identified as endemic areas for cholera transmission, and despite continuous control efforts, they continue to experience regular cholera outbreaks that occasionally spread to the rest of the country. In a region where access to improved water sources is particularly poor, the question of which improvements in water access should be prioritized to address cholera transmission remains unresolved. This study aimed at investigating the temporal association between water supply interruptions and Cholera Treatment Centre (CTC) admissions in a medium-sized town.

Methods and Findings

Time-series patterns of daily incidence of suspected cholera cases admitted to the Cholera Treatment Centre in Uvira in South Kivu Province between 2009 and 2014 were examined in relation to the daily variations in volume of water supplied by the town water treatment plant. Quasi-poisson regression and distributed lag nonlinear models up to 12 d were used, adjusting for daily precipitation rates, day of the week, and seasonal variations. A total of 5,745 patients over 5 y of age with acute watery diarrhoea symptoms were admitted to the CTC over the study period of 1,946 d. Following a day without tap water supply, the suspected cholera incidence rate increased on average by 155% over the next 12 d, corresponding to a rate ratio of 2.55 (95% CI: 1.54–4.24), compared to the incidence experienced after a day with optimal production (defined as the 95th percentile—4,794 m^3).

Competing Interests: The authors have declared that no competing interests exist.

Abbreviations: CI, confidence interval; CTC, cholera treatment centre; df, degree of freedom; DLNM, distributed lag nonlinear models; DPD, N,N diethyl-p-phenylene diamine; DRC, Democratic Republic of the Congo; GPS, global positioning system; JMP, Joint Monitoring Programme; NASA GES DISC, NASA Goddard Earth Sciences Data and Information Services Center; NGO, non-governmental organisation; MDG, Millennium Development Goals; QMRA, Quantitative Microbiological Risk Assessment; RR, relative risk; WHO, World Health Organization.

Suspected cholera cases attributable to a suboptimal tap water supply reached 23.2% of total admissions (95% CI 11.4%–33.2%). Although generally reporting less admissions to the CTC, neighbourhoods with a higher consumption of tap water were more affected by water supply interruptions, with a rate ratio of 3.71 (95% CI: 1.91–7.20) and an attributable fraction of cases of 31.4% (95% CI: 17.3%–42.5%). The analysis did not suggest any association between levels of residual chlorine in the water fed to the distribution network and suspected cholera incidence. Laboratory confirmation of cholera was not available for this analysis.

Conclusions

A clear association is observed between reduced availability of tap water and increased incidence of suspected cholera in the entire town of Uvira in Eastern Democratic Republic of the Congo. Even though access to piped water supplies is low in Uvira, improving the reliability of tap water supply may substantially reduce the incidence of suspected cholera, in particular in neighbourhoods having a higher access to tap water. These results argue in favour of water supply investments that focus on the delivery of a reliable and sustainable water supply, and not only on point-of-use water quality improvements, as is often seen during cholera outbreaks.

Introduction

In 2012, the Democratic Republic of the Congo (DRC) reported more than 28% of all reported cholera cases in Africa, and 27% of cholera-related deaths globally [1]. The Great Lakes region and particularly Eastern DRC have been identified as a stable transmission focus for cholera, where cases of cholera have been reported every year since 1978 [2,3]. South-Kivu province, in particular, is considered as an endemic area, reporting cases nearly continuously since 2000 [4]. Although detailed data for this particular area is scarce, access to safe water in DRC in 2015 is generally poor, with 52% of the population using improved water sources, and only 8% of the population having access to piped water on premises [5].

Cholera has been predominantly linked with contaminated water ever since John Snow removed the pump handle from the London Broad Street pump in 1854. However, more recent research confirmed the role of direct human-to-human transmission already suggested by John Snow as an important route in the 1850s [6,7]. This more direct pathway has been suggested as an explanation for the explosive nature of cholera outbreaks, along with hyperinfectivity of cholera organisms when freshly shed by an infected individual [8].

Cholera control strategies, especially during outbreaks, generally focus on the provision of clean drinking water and the removal of potential contamination of that water by means of water treatment at source, or at point of use [4,9]. They also commonly include activities that promote personal, food, and household hygiene but more rarely address the issue of the amount of water available to a household and the reliability of the water source, even though those factors will impact water collection, storage, and hygiene practices and, as a result, the microbial quality of the water at point of use [10,11].

Particularly in endemic areas, an unreliable water supply that provides an unpredictable amount of tap water to households is therefore likely to increase cholera incidence through the occasional use of unsafe sources of water, unsafe water storage, and restricted personal and household hygiene behaviours. Water supply interruptions may also lead to contamination of the water in the piped network due to low or negative pressure and ingress of pathogens. There is a growing body of evidence on the impact of inconsistent access and use of improved water sources on the microbial quality of water at point of use [12]. However, when searching for published evidence of the impact of water, sanitation, and hygiene interventions during cholera epidemics, we found only a single study in which piped water supply was implemented in a cholera-affected community; it showed a 65% reduction in endemic cholera incidence, though the study suffered from major design flaws [13]. A recent review on water supply interventions impact on diarrhoeal diseases found only limited evidence on the impact of continuous and safe quality piped water supply on diarrhoeal diseases; a single study reported an estimated 73% reduction in the risk of diarrhoeal diseases compared to noncontinuous piped water supply [14]. The degree to which the unreliability of clean water supply affects cholera incidence, especially in endemic areas, has not been previously investigated.

Using data collected in a middle-sized town in DRC, our study explores the association between the daily volume of chlorinated tap water distributed to households and admissions of suspected cholera cases to the local Cholera Treatment Centre (CTC) with a time-series regression analysis.

Methods

Study Area

The city of Uvira is located in South Kivu Province in Eastern DRC on the shores of Lake Tanganyika, and had an estimated population of 205,000 people in 2012. The town is spread over nearly 10 km along Lake Tanganyika at an altitude of 800 m above sea level and is crossed by three rivers. Uvira has a moderate tropical climate with an average temperature of 26°C throughout the year.

The population relies on surface water sources (the lake and rivers) and on the water supply managed by the national water agency (Regideso), which provides chlorinated water through a sparse network of approximately 2,800 private and shared taps across town. It is estimated that 80% of the population regularly or occasionally uses this municipal water supply, and of that group, only a quarter have access to a tap in their yard or dwelling. Unreliable power supply from the local grid, limited resources for generator use, irregular supply of chemicals for water treatment (i.e., chlorine, aluminium sulphate, lime), frequent downtime for routine maintenance, and occasional equipment failure all result in an intermittent and unreliable supply of tap water.

Between 2004 and 2013, at least one case of suspected cholera has been reported weekly by the Uvira district health office, and a Cholera Treatment Centre (CTC) in Uvira district hospital has been set up to admit all patients with severe acute diarrhoea. Patients are treated for dehydration and administered broad spectrum antibiotics. However, in the absence of adequate local laboratory facilities, regular confirmation of cholera is not conducted, except occasionally at the onset of suspected cholera outbreaks. Admission and treatment at the CTC is free of charge for all patients.

Suspected Cholera Case and Water Supply Data

Daily admissions to the CTC during the period of January 1, 2009, to April 30, 2014, were extracted from the database held by the district health office since 2009, which is updated weekly from the paper register used by clinical staff at the CTC. Patients admitted at the CTC are recorded along with their neighbourhood of residence, and patients residing outside of

Uvira municipality were excluded from the analysis. In order to conform to the WHO suspected cholera case definition in endemic areas, patients aged 5 y or under were excluded [9].

Daily volume of water supplied for the same time period was collected from the register held at the Regideso water treatment plant in Uvira, in which river water drawn upstream of town is treated by sedimentation/flocculation and chlorination before being fed to a single 1,600 m^3 reservoir supplying the town gravity distribution system. Volumes of water supplied over 24 h were measured with a flowmeter placed at the only water treatment station output, where the levels of residual chlorine in the supplied water were also measured by means of N,N diethyl-p-phenylene diamine (DPD) tablets and the daily average (of over two or more measurements) recorded in the same register.

Although the piped water system serves, in theory, all areas of Uvira municipality, the 184 neighbourhoods of Uvira were classified as having a higher or lower tap water consumption based on an estimated monthly average volume of tap water available per person. This was calculated for each neighbourhood on the basis of the volume of water billed in February 2012 to 2,630 Regideso taps (each of which is geographically referenced and allocated to a particular neighbourhood by global positioning system [GPS]) divided by the estimated neighbourhood population. Based on actual tap metering, this estimate accounts for the large variability between taps in use (shared or private) and water availability that is dependent on the tap location on the distribution network.

Daily precipitation rates in mm/h estimates averaged for Uvira area at a 0.25 degree by 0.25 degree spatial resolution were obtained from NASA Tropical Rainfall Measuring Mission (TRMM) 3B42-v7, through the Giovanni online data system, developed and maintained by the NASA Goddard Earth Sciences Data and Information Services Center (NASA GES DISC). Daily variations of temperature are small in Uvira, and no complete daily temperature record for the region was identified. Temperature was therefore not included as a confounder in the analysis.

Statistical Methods

The relationship between the daily number of admissions at the CTC and the daily volume of water supplied by the Regideso was examined using generalized linear Poisson regression models allowing for overdispersion [15]. To account for seasonality and long-term trends in potential unmeasured confounders, a cubic spline of date (12 degrees of freedom [df] per year) was included in the model, as a more flexible alternative than including month-in-year terms. Terms for the day of the week (indicators) and daily precipitation rates (linear) were also added. An autoregressive term was introduced to the model in order to control for a significant temporal auto-correlation at days 1 and 2.

The relationship between water supply level and suspected cholera incidence was modelled using the framework of distributed lag nonlinear models (DLNMs). These models allow the net effect of an exposure to be computed as the sum of contributions at different lags, through the definition of a lag-response curve in addition to the exposure-response relationship [16]. DLNMs were originally developed for investigating temperature-health associations [17].

Informed by a recent review of incubation periods for cholera and delays due to storage capacity in the water production system, we considered lags up to 12 d and assumed that effects, if any, rose from none at lag 0 [18].

A linear exposure-response function scaled against the 95th percentile of volume (subsequently referred to as optimal production) was considered for association between volume and admissions. The delay in effect of water volume on admissions (modelled as a lag-response curve) was constrained to follow a smooth quadratic curve.

The overall cumulative effect of volume produced on CTC admissions over 12 d was computed by summing effects over the whole lag period. The association is reported as relative risk (RR).

The number and proportion of suspected cholera cases attributable to water production lower than the 95th percentile of volume were estimated from the model, as defined by Gasparrini and Leone [19]. These calculations assume that a fraction (RR-1)/RR of suspected cholera cases on any given day is attributable to the exposure to a suboptimal water supply during the 12 previous days modelled to give a relative risk, and are computed for the entire study duration. The figures are interpreted as the number of suspected cholera cases attributable to suboptimal water production during the study period, which are potentially preventable if the volume of water supplied was maintained above the optimal production level.

We explored the potential heterogeneity of effect in neighbourhoods with a higher or a lower tap water consumption. The median value of tap water invoiced across neighbourhood in February 2012 was 2.8 l per person and per day and this was taken as cut-off value for stratification of neighbourhoods into higher and lower tap water consumption. A regression model was fitted separately to each series of higher and lower consumption, based on the same specifications used in the main analysis. These models also included the number of admissions in the other area as an additional term, in order to account for potential inter-neighbourhoods transmission.

The sensitivity of the estimates to the modelling assumptions outlined above was tested by the comparison with models specified with different choices. The model parameters tested included the number of df of the spline function of date, the shape of the lag-response function, inclusion of an impact on day 0 and inclusion of a term for chlorine residual levels.

All analyses were performed in the R environment using the package dlnm [20]. The R code for reproducing the main results is included in S3 Text. The full code is available from the corresponding author on request.

Study Design and Ethics

Local authorities and international non-governmental organisations (NGO) involved in cholera control activities in Uvira observed that major cholera epidemics often seemed to follow water supply interruptions. In order to investigate this possible association, the authors retrospectively collected relevant data from the local water agency Regideso and from Uvira CTC, and the reliability, size, and completeness of the dataset was deemed sufficient to perform a time-regression analysis. No formal prospective analysis plan was submitted.

This study was approved by the Observational Committee of the London School of Hygiene & Tropical Medicine Research Ethics committee (reference 7745). The Committee waived informed consent from participants as the data were anonymously collected and analysed.

Results

Between January 1, 2009, and April 30, 2014, a total of 5,745 patients aged over 5 y were admitted to the CTC, with daily admissions ranging between 0 and 32. Daily volume of water supplied to the city ranged from 0 to 5,748 m^3 with a 95th percentile ("optimal" production) of 4,794 m^3, while the level of residual chlorine ranged from 0.1 to 1.7 mg/l (Table 1 and Fig 1).

Association of Volume of Water Supplied with CTC Admissions

A crude group analysis shows a decreasing mean number of admissions at lag 6 d with an increasing volume of water supplied (Fig 2). Fig 3 displays the overall cumulative exposure-response relationship between volume of water supplied and incidence of suspected cholera as

Table 1. Distribution of admissions to CTC, volume of tap water supplied, and levels of residual chlorine in the water between January 1, 2009, and April 30, 2014, in Uvira, DRC.

	Range	Median	Interquartile range	95th percentile	Number of records (% missing)
Daily admissions to CTC (all town)	0 to 32 patients	2	1–4	8	1,946 (0%)
Daily admissions to CTC (neighbourhoods with tap water consumption ≥2.8 l/day/person)	0 to 18 patients	1	0–2	5	1,946 (0%)
Daily admissions to CTC (neighbourhoods with tap water consumption <2.8 l/day/person)	0 to 18 patients	1	0–2	6	1,946 (0%)
Daily volume of water supplied	0 to 5,748 m^3	3,741	3,072–4,264	4,794	1,944 (0.1%)
Daily average of residual chlorine in water supplied	0.1 to 1.7 mg/l	0.60	0.50–0.70	0.80	1,839 (5.5%)
Daily precipitation rate	0 to 70.7 mm/h	0.4	0–2.6	14.2	1,944 (0.1%)

doi:10.1371/journal.pmed.1001893.t001

modelled with the DLNM time-series regression. When the water treatment plant produces no water, the model predicts an increase of 155% in the number of suspected cholera cases within the next 12 d, corresponding to an RR of 2.55 (95%CI: 1.54–4.24), compared to the optimal production level (4,794 m^3).

The temporal distribution of the effects is summarized in Fig 4. This curve indicates a maximum effect 5 to 7 d after the plant fails to produce any water, with the increased risk then vanishing after 12 d.

Admissions to the CTC attributable to a suboptimal volume supplied were 23.2% (95% CI 11.4%–33.2%) with 1,332 attributed cases out of 5,745.

Stratification by Neighbourhood Access to Tap Water

In the 92 neighbourhoods with a higher tap water consumption, 2,528 CTC admissions were recorded for the study period for an estimated population of 98,000 people (average yearly incidence of 4.8 suspected cholera cases per 1,000), while in the 92 neighbourhoods with a lower tap water consumption, 3,217 suspected cholera cases were admitted for an estimated population of 106,000 inhabitants (average yearly incidence of 5.7 suspected cholera cases per 1,000). Association between predicted incidence rates for suspected cholera in both areas and volume of water supplied by the treatment plant, at set values for other model covariates are shown in Fig 5. Although the incidence is predicted to be generally lower in areas with higher tap water consumption, this incidence increases markedly with a reduction in water availability, whereas the incidence in neighbourhoods with a lower consumption is not influenced by the volume of water produced. The risk of no water produced relative to an optimal production was significantly higher in areas with a higher tap water consumption (RR = 3.71, 95% CI 1.91–7.20) compared to neighbourhoods with a lower consumption (RR = 1.17, 95% CI 0.63–2.14; p (z-test) = 0.01).

In the neighbourhoods consuming more tap water, the number of suspected cholera cases attributable to suboptimal water production was estimated to be 793 out of 2,528 (31.4%, 95% CI 17.3–42.5).

Sensitivity Analysis

Variations of the assumptions on the lag-response function and the flexibility of the cubic spline of date in the main model lead to a 12-day overall cumulative RR of a comparable

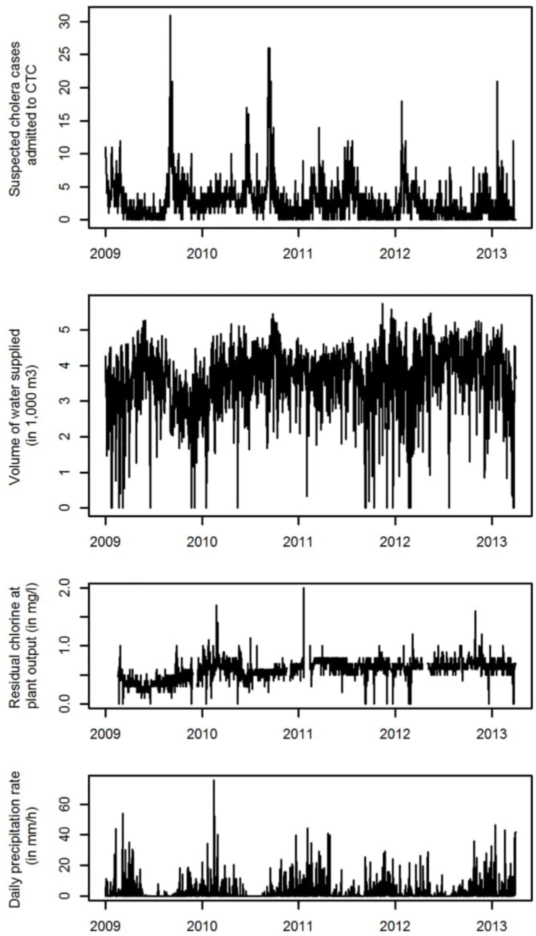


Fig 1. Daily time series between January 1, 2009, and April 30, 2014, for admissions at the CTC, volume of water supplied, level of residual chlorine in produced water and precipitation rate in Uvira, DRC.

doi:10.1371/journal.pmed.1001893.g001

magnitude, which falls between 2.44 and 6.66 (Table 2). This suggests that the reported results for the cumulative RR are robust to model parameters variation.

We found no evidence for confounding of the volume effect by residual chlorine concentration (Model H), which was not surprising given its low correlation with volume (r = -0.03). In fact, there was no evidence in these data for an association between residual chlorine levels in the water supplied and suspected cholera. The RR for low residual chlorine (0.4 mg/l) compared to high (0.8 mg/l) was 0.7 (0.43–1.14), when chlorine was entered into the model with the same 12-day lag structure as used for volume.

Discussion

Using time-series regression methods to investigate the relationship between chlorinated tap water availability and suspected cholera incidence, our results showed that the unreliability of

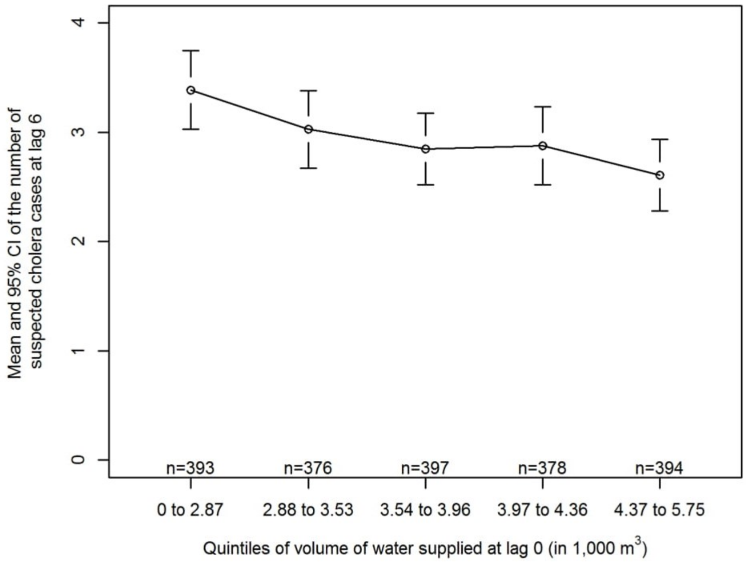


Fig 2. Mean incidence of suspected cholera at lag 6 d by quintile of volume of water supplied at lag 0 day. n: number of days.

doi:10.1371/journal.pmed.1001893.g002

tap water supply in a cholera endemic setting was associated with a more than 2-fold increase in suspected cholera incidence at city level, and a nearly 4-fold increase in areas with higher tap water consumption. The results also showed that 23% of the cases admitted at the CTC in Uvira between 2009 and spring 2014 may be attributed to irregular tap water supply.

Our study suggests that cholera incidence in Uvira could possibly be reduced by nearly a quarter by ensuring that the existing water treatment infrastructure performed better and supplied water daily in regular and sufficient quantity to its existing customers, even in a town

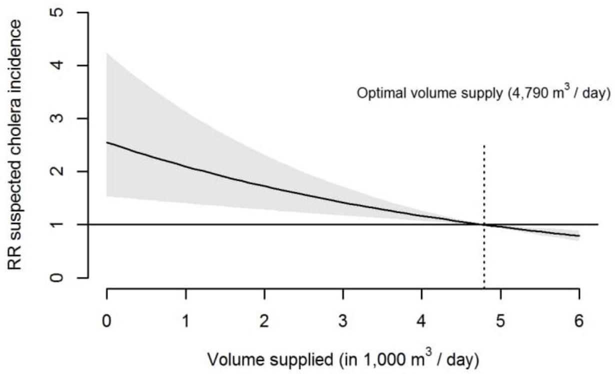


Fig 3. Twelve-day cumulative association of volume of tap water supplied with suspected cholera incidence in Uvira. Shaded area: 95% CI for relative risk (RR)–model predictors: daily volume of water supplied, cubic spline of date, day of the week, precipitation rate of the day and number of CTC admissions on the previous 12 d.

doi:10.1371/journal.pmed.1001893.g003

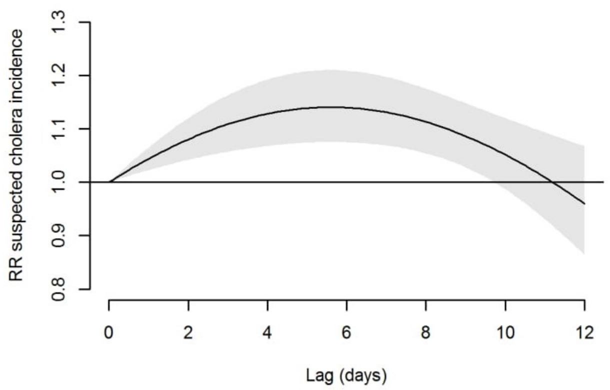


Fig 4. Association of absence of water supply (0 m^3 of water supplied at day 0) with suspected cholera incidence on the 12 following days. Shaded area: 95% CI of RR.

doi:10.1371/journal.pmed.1001893.g004

where a single tap serves, on average, nearly 60 people. These results substantiate the increased annual risk of infection due to water supply failure estimated by Hunter and colleagues for other diarrhoeal pathogens by means of Quantitative Microbiological Risk Assessment (QMRA) [21]. These calculations showed a 13% increase in annual risk of infection for rotavirus or Enterotoxigenic *Escherichia coli* after a single day of water supply failure.

Interruptions in piped water supply are likely to increase cholera incidence in an endemic setting through multiple pathways. In the absence of other improved sources of water in Uvira, households can be expected to revert to unsafe water sources when tap water is not available, or they may reduce their use of water and thereby restrict hygiene behaviours that can reduce person-to-person cholera transmission. The unreliability of tap water supply may also encourage households to store large amounts of water for longer periods, leading to an increased risk of water contamination, as was observed in East Africa [22]. In Bangladesh, *Vibrio cholerae*

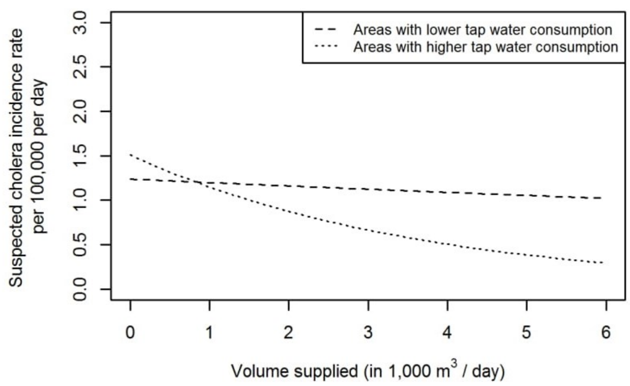


Fig 5. Predicted incidence rate of suspected cholera for 10,000 as a function of water volume supplied, stratified by neighbourhoods with higher (≥2.8 l per person per day) and lower (<2.8 l per person per day) tap water consumption. Model covariates for this prediction were set to the dataset midpoint (September 1, 2011), median precipitation rate value (0.4 mm/h), median number of suspected cholera cases admitted in the other neighbourhood (1 case), and a Thursday.

doi:10.1371/journal.pmed.1001893.g005

Table 2. Sensitivity analysis of the effect of the number of df/year in cubic spline of date and lag-response function on the 12-day cumulative association of volume supplied with admissions to CTC.

Model	df/year in cubic spline of date	Lag-response function	12 day cumulative association of absence of water supplied—RR (95% CI)
Main model	12	Polynomial 2nd degree, no intercept	2.55 (1.54–4.24)
Model A	6	Polynomial 2nd degree, no intercept	6.66 (4.31–10.29)
Model B	18	Polynomial 2nd degree, no intercept	2.44 (1.45–4.10)
Model C	12	Polynomial 2nd degree, with intercept	3.08 (1.69–5.60)
Model D	12	Polynomial 3rd degree, no intercept	2.76 (1.60–4.76)
Model E	12	Polynomial 3rd degree, with intercept	2.99 (1.65–5.45)
Model F	12	3 strata with breaks at days 3 and 6	3.15 (1.75–5.69)
Model G	12	5 Strata with breaks at days 2, 4, 6, and 8	3.11 (1.72–5.62)
Model H	12	Polynomial 2nd degree, no intercept*	2.86 (1.47–5.58)

* Model H adjusts for residual chlorine levels in the water supplied, modelled with the same lag structure as volume.
df, degrees of freedom

doi:10.1371/journal.pmed.1001893.t002

was found in 1.2% of water storage vessels in households that used a safe water source free of contamination [23]. A recent study also showed that stored water collected from improved sources in Cambodia represented a similar risk for health to that from stored water collected from other sources, when comparing microbial quality [24].

The results of this study highlight the need to take into account the reliability of tap water supply when considering it as an "improved" water source. The term "improved" was adopted by the WHO/UNICEF Joint Monitoring Program (JMP) in charge of monitoring progress toward the Millennium Development Goals (MDGs) for water and sanitation, and uses the technology of the household water supply as a proxy for the water's quality. According to the JMP, improved water sources include piped supplies, protected wells, and boreholes, but exclude surface water sources or unprotected springs. JMP monitoring figures showed that approximately 2.3 billion people have gained access to an improved water source in the last 10 y, while 70% of that number gained access to piped water on premises, although these definitions fail to give any indication on the continuity of the water source used [25]. Recognizing this limitation, the Post-2015 Water Working Group assigned by the JMP to develop new metrics for monitoring global water access, now includes the concept of supply reliability in its definition of a "safely managed" water source. Households having access to an improved source of water "on-premises" that fails, on average, less than 2 d in 2 wk; provides water in sufficient quantity for domestic use; meets WHO guidelines for *E. coli*, arsenic, and fluoride; and is subject to a risk-management plan would be considered as using a safely managed water source, while households using an improved water source within 30 min, including queuing, of the house would be considered as having access to basic drinking water [26]. Our study findings support the inclusion of supply reliability in the definition of a higher level of access to drinking water and highlight the health implications of piped water supply unreliability. Our findings suggest that a single day of supply interruption may translate in a significant increase in suspected cholera incidence, especially in neighbourhoods where people rely more heavily on piped water, and the health implications of the JMP threshold of 2 d of failure within a fortnight can be questioned. In addition, the issue of perceived reliability of the supply by users may also be examined, when considering that the health benefits associated with access to piped water supply on premises derive partly from the decreasing need for water storage in the household. Indeed, if households still experience regular unplanned piped supply interruptions, they are likely to keep storing water as a coping strategy for these unexpected cuts, potentially reducing the health benefits expected with piped water on premises. This suggests that avoiding interruptions of piped supplies or mitigating the impact of such interruptions—by informing users ahead of their occurrence when interruptions are planned, or of the expected duration of the interruption when unexpected for example—should be a major element of the risk management plan for these safely managed sources.

Our results also underline the importance of appropriately characterizing water supply used when quantifying the health risk of using "unimproved" water sources. Indeed, the counterfactual often relates to population groups using piped water on premises [27]. These "reference" groups may actually bear an increased health risk linked to the unreliability of the water supply, which remains uncontrolled for in most health impact studies of water interventions. This may explain why an updated review on water-related interventions, published in *The Lancet's* "Global Burden of Disease" series, found no evidence of a significant diarrhoea risk reduction from access to "on-plot" water supply compared to access to other "improved" sources of water [27].

The presence of residual chlorine in the water fed into the distribution network in Uvira indicates that the chlorination at the plant was sufficient to remove existing microbial contamination from the source of the treatment plant. Due to the complexity of chlorine decay in the distribution system between the treatment plant and the tap, these figures are unlikely representative of the amount of residual chlorine at the taps, although chlorine residuals at the taps can only be lower than those at the treatment plant [28]. These levels of residual chlorine in the water supplied, most of the time comprised between 0.5 and 0.7 mg/l, may not be enough to prevent water contamination in the distribution network, especially after interruptions and at low pressure [29]. Water contamination by consumers during collection and storage may also be a significant means of cholera transmission that could be reduced with higher levels of residual chlorine in the supplied water. Our analysis did not reveal any association between levels of residual chlorine at the treatment plant and suspected cholera incidence, but it did not have the required statistical power to investigate a potential interaction between volume supplied and chlorination in detail.

One limitation of this study is that the case definition used may have included patients admitted to the CTC for diarrhoea of aetiology other than cholera infection. Indeed, CTC admissions were rarely confirmed by laboratory tests. A study on cholera rapid diagnostic tests in DRC found that 73% of suspected cholera cases tested during an outbreak in Katanga province were laboratory-confirmed by culture and PCR [30]. Misclassification of admissions as cholera cases would not pose a threat to the validity of comparisons over time as performed by the present study, except if the proportion of non-cholera cases decreases during higher incidence periods, which cannot be excluded.

Another limitation of this study is the assumption that all cholera cases in Uvira seek health care at the CTC. No evidence is available to confirm or reject this assumption, but little alternative to treat severe diarrhoea and associated dehydration is available in Uvira besides the CTC. Furthermore, the CTC provides health care free of charge, and direct costs of treatment should not act as a factor in socioeconomic selection bias. It is unlikely that the potential bias introduced by using health facility data would vary significantly over time, which is important for the present analysis.

Similarly, this analysis assumes that all cholera cases presenting to the CTC are admitted, even during outbreak periods when the CTC is reaching its capacity limit. However a mechanism has been in place since 2009 to strengthen CTC capacity when the number of admissions exceeds 25 per week. It is therefore reasonable to consider that the CTC did not refuse patients due to limited bed space.

Time-series regression modelling does not directly incorporate variation in immunity in the population though the inclusion as regressor of a cubic spline of time will indirectly allow for such fluctuation, if smooth, so the limitation should be minor. This applies, as well, to other potential unmeasured confounders linked with seasonality or showing smooth variations, such as population nutritional status or disposable income. In particular, we considered that the spline function of time-captured variations of the economic cost and affordability of water, as no sudden official water tariffs changes were recorded during the study period. In addition, price paid for water in Uvira is highly heterogeneous and unregulated due to the sharing of private taps and reselling of tap water. Although power supply interruptions in Uvira can be considered as abrupt and are possibly more frequent during the dry season due to less energy production in hydroelectric dams, power supply interruptions should be considered as a cause of variations of the exposure of interest rather than a potential confounder. In a town where less than 8% of the population reports owning a refrigerator, power supply interruptions are unlikely to be associated with suspected cholera incidence by means of inappropriate food preservation. A moderate overdispersion in model residuals (scale parameter of 1.88) was observed and was accounted for by modelling with a quasi-Poisson family. Assuming that the distribution of daily cases would otherwise follow a Poisson distribution, this overdispersion more generally attests to the model not capturing all of the systematic causes of variation in cholera incidence. However, there seems little reason to expect such causes to be associated with water volume, which would be necessary to cause bias in the relative risks reported. Residual confounding cannot be excluded, but a causal relationship seems the most likely explanation for the observed association. In reference to Bradford Hill's criteria for causality, our results indeed demonstrate a reasonably strong effect occurring after exposure, with a dose-response pattern and a stronger association in areas with higher tap water consumption, and along a plausible and coherent biological pathway [31].

To conclude, our results showed a clear association between the poor reliability of the water supply system in Uvira and the incidence of suspected cholera in the entire city. They suggest that supply interruptions and unreliability increase population exposure to unsafe sources of water, encourage water storage in households and restrict hygiene practices, all of which translate into increased suspected cholera transmission. They argue in favour of water supply investments that focus on the delivery of a reliable and sustainable water supply and not only on point-of-use water quality improvements as are often seen during cholera outbreaks.

Supporting Information

S1 Table. Dataset.
(CSV)

S1 Text. Data analysis strategy and timeline.
(DOCX)

S2 Text. STROBE checklist.
(DOC)

S3 Text. R script.
(R)

Acknowledgments

We thank all the field staff who contributed to this study, in particular, the Regideso staff in Uvira. We are also grateful to OXFAM Great Britain staff for their invaluable support in Uvira. In addition, we truly appreciate the valuable input provided by M. Lemenager and F. Bedecarrats from the Agence Française de Développement during the study.

Author Contributions

Conceived and designed the experiments: AJ JHJE SC AG BA. Analyzed the data: AJ AG BA. Contributed reagents/materials/analysis tools: AG. Wrote the first draft of the manuscript: AJ. Contributed to the writing of the manuscript: JHJE SC TV AG BA. Agree with the manuscript's results and conclusions: AJ JMS AK MB FB TV AG BA SC JHJE. Collected, checked and managed the data: AJ JMS AK MB FB TV. All authors have read, and confirm that they meet, ICMJE criteria for authorship.

References

1. World Health Organization. (2013) Weekly Epidemiological Record. 88: 321–336. PMID: 23980290
2. Bompangue D, Giraudoux P, Handschumacher P, Piarroux M, Sudre B, et al. (2008) Lakes as source of cholera outbreaks, Democratic Republic of Congo. Emerg Infect Dis 14: 798–800. doi: 10.3201/eid1405.071260 PMID: 18439365
3. Rebaudet S, Sudre B, Faucher B, Piarroux R (2013) Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. J Infect Dis 208 Suppl 1: S46–54. doi: 10.1093/infdis/jit195 PMID: 24101645
4. Muyembe JJ, Bompangue D, Mutombo G, Akilimali L, Mutombo A, et al. (2013) Elimination of cholera in the democratic Republic of the Congo: the new national policy. J Infect Dis 208 Suppl 1: S86–91. doi: 10.1093/infdis/jit204 PMID: 24101651
5. WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation. (2015) Estimates on the use of water sources and sanitation facilities. http://www.wssinfo.org/data-estimates/
6. Codeco CT, Coelho FC (2006) Trends in cholera epidemiology. PLoS Med 3: e42. PMID: 16435891
7. Feachem RG (1981) Environmental aspects of cholera epidemiology. I. A review of selected reports of endemic and epidemic situations during 1961–1980. Trop Dis Bull 78: 675–698. PMID: 7022972
8. Hartley DM, Morris JG Jr., Smith DL (2006) Hyperinfectivity: a critical element in the ability of V. cholerae to cause epidemics? PLoS Med 3: e7. PMID: 16318414
9. World Health Organization. Prevention and control of cholera outbreaks: WHO policy and recommendations. http://www.who.int/cholera/technical/prevention/control/en/
10. Cairncross S, Cliff JL (1987) Water use and health in Mueda, Mozambique. Trans R Soc Trop Med Hyg 81: 51–54. PMID: 3445322
11. Wright J, Gundry S, Conroy R (2004) Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Trop Med Int Health 9: 106–117. PMID: 14728614
12. Shaheed A, Orgill J, Montgomery MA, Jeuland MA, Brown J (2014) Why "improved" water sources are not always safe. Bull World Health Organ 92: 283–289. doi: 10.2471/BLT.13.119594 PMID: 24700996
13. Azurin JC, Alvero M (1974) Field evaluation of environmental sanitation measures against cholera. Bulletin of the World Health Organization 51: 19–26. PMID: 4549038
14. Wolf J, Pruss-Ustun A, Cumming O, Bartram J, Bonjour S, et al. (2014) Systematic review: Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression. Trop Med Int Health 19: 928–942. doi: 10.1111/tmi.12331 PMID: 24811732
15. Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B (2013) Time series regression studies in environmental epidemiology. Int J Epidemiol 42: 1187–1195. doi: 10.1093/ije/dyt092 PMID: 23760528
16. Gasparrini A (2014) Modeling exposure-lag-response associations with distributed lag non-linear models. Stat Med 33: 881–899. doi: 10.1002/sim.5963 PMID: 24027094
17. Armstrong B (2006) Models for the relationship between ambient temperature and daily mortality. Epidemiology 17: 624–631. PMID: 17028505

18. Azman AS, Rudolph KE, Cummings DA, Lessler J (2013) The incubation period of cholera: a systematic review. J Infect 66: 432–438. doi: 10.1016/j.jinf.2012.11.013 PMID: 23201968
19. Gasparrini A, Leone M (2014) Attributable risk from distributed lag models. BMC Med Res Methodol 14: 55. doi: 10.1186/1471-2288-14-55 PMID: 24758509
20. Gasparrini A (2011) Distributed Lag Linear and Non-Linear Models in R: The Package dlnm. J Stat Softw 43: 1–20. PMID: 22003319
21. Hunter PR, Zmirou-Navier D, Hartemann P (2009) Estimating the impact on health of poor reliability of drinking water interventions in developing countries. Sci Total Environ 407: 2621–2624. doi: 10.1016/j.scitotenv.2009.01.018 PMID: 19193396
22. Thompson J, Porras I, Tumwine J, Mujwahuzi M, Katui-Katua M, et al. (2000) Drawers of Water II: Thirty Years of Change in Domestic Water Use and Environmental Health in East Africa. Russell Press.
23. Spira WM, Khan MU, Saeed YA, Sattar MA (1980) Microbiological surveillance of intra-neighbourhood E1 Tor cholera transmission in rural Bangladesh. Bull World Health Organ 58: 731–740. PMID: 6975182
24. Shaheed A, Orgill J, Ratana C, Montgomery MA, Jeuland MA, et al. (2014) Water quality risks of 'improved' water sources: evidence from Cambodia. Trop Med Int Health 19: 186–194. doi: 10.1111/tmi.12229 PMID: 24252094
25. WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation. (2013) Progress on sanitation and drinking-water: 2013 update. Geneva: World Health Organization. 39 pp.
26. WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation. (2014) WASH targets and indicators post-2015: Recommendations from international consultations. Comprehensive recommendations—Update April 2014. http://www.wssinfo.org/fileadmin/user_upload/resources/post-2015-WASH-targets-factsheet-12pp.pdf
27. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, et al. (2012) A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380: 2224–2260. doi: 10.1016/S0140-6736(12)61766-8 PMID: 23245609
28. Goyal R, Patel HM (2014) Analysis of residual chlorine in simple drinking water distribution system with intermittent water supply. Applied Water Science: 1–9.
29. Kumpel E, Nelson KL (2014) Mechanisms affecting water quality in an intermittent piped water supply. Environ Sci Technol 48: 2766–2775. doi: 10.1021/es405054u PMID: 24459990
30. Page AL, Alberti KP, Mondonge V, Rauzier J, Quilici ML, et al. (2012) Evaluation of a rapid test for the diagnosis of cholera in the absence of a gold standard. PLoS One 7: e37360. doi: 10.1371/journal.pone.0037360 PMID: 22666350
31. Hill AB (1965) The Environment and Disease: Association or Causation? Proc R Soc Med 58: 295–300. PMID: 14283879

Editors' Summary

Background

Every year, outbreaks (epidemics) of cholera—a bacterial gut infection caused by *Vibrio cholerae*—make 2–3 million people ill and kill about 100,000 people. Because people get cholera by drinking water or eating foods contaminated by the feces of infected individuals, cholera outbreaks usually occur in places with poor sanitation and poor access to clean water, such as slums and refugee camps. Natural disasters that disrupt sanitation and water systems can also trigger cholera outbreaks. Most people who become infected with *V. cholerae* have no or mild symptoms, but they may shed bacteria in their feces for up to two weeks. Other infected individuals develop severe diarrhea, producing profuse watery feces. With prompt treatment—replacement of lost fluids and salts by drinking an oral rehydration fluid or by injection of rehydration fluid into a vein in the worst cases—less than 1% of infected individuals die. Without treatment, people with severe cholera can die from dehydration within hours of developing symptoms.

Why Was This Study Done?

Cholera control strategies generally focus on the provision of safe drinking water through treatment of water supplies at source or at point of use. But could addressing the reliability of household water sources in regions with poor access to piped treated water and other improved water supplies help to control cholera? A water source that provides an unpredictable amount of clean water is likely to increase the incidence of cholera (the number of new cases in a population in a given time period) through causing the occasional use of unsafe water sources, unsafe water storage, and limited water availability for handwashing. In this time-series regression (a statistical method for analyzing data collected over time), the researchers investigate the temporal association between water supply interruptions and admissions to the cholera treatment center in Uvira, a town in a province of the Democratic Republic of the Congo where cholera is endemic (always present). Although 80% of Uvira's population regularly or occasionally uses the town's chlorinated water supply, only a quarter of the population has access to a water tap in their yard or dwelling, and the water supply is intermittent and unreliable.

What Did the Researchers Do and Find?

The researchers examined data on the daily number of people over 5 y old admitted to Uvira's cholera treatment center with diarrhea in relation to daily variations in the volume of water supplied by the town's water treatment plant between 2009 and 2014. In the 12 d following a day without tap water, the incidence of suspected cholera increased by 155%. That is, compared to the 12 d following a day with optimal water production, there were 2.55 times as many cases of cholera in the 12 d following a water supply interruption. Moreover, 23.2% of suspected cholera cases over the study period were potentially attributable to a suboptimal tap water supply. Compared to neighborhoods with low tap water consumption, the incidence of suspected cholera was lower in neighborhoods with high tap water consumption. However, in these neighborhoods, water supply interruptions increased the incidence of suspected cholera in the following 12 d nearly 4-fold and 31.4% of suspected cholera cases were attributable to water supply interruptions. Finally, there was no association between the levels of residual chlorine in water fed to the distribution network and the incidence of suspected cholera.

What Do These Findings Mean?

These findings show an association between reduced availability of tap water and increased incidence of suspected cholera in an urban setting in the Democratic Republic of the Congo where cholera is endemic. Thus, improving the reliability of the water supply might substantially reduce the incidence of cholera, particularly in neighborhoods with greater access to tap water. Some aspects of the study design may affect the accuracy of these findings. For example, because laboratory confirmation of cholera was not available, some of the patients admitted to the cholera treatment center with diarrhea may not have had cholera. However, these findings argue in favor of focusing on the delivery of a reliable and sustainable water supply in cholera-endemic areas as a strategy for cholera control and, more generally, highlight the public-health importance of providing access to reliable tap water supplies, a concept now included by the WHO/UNICEF Joint Monitoring Program (an official body tasked with monitoring global water access) in its definition of a safely managed water source.

Additional Information

This list of resources contains links that can be accessed when viewing the PDF on a device or via the online version of the article at http://dx.doi.org/10.1371/journal.pmed.1001893.

- This study is further discussed in a *PLOS Medicine* Perspective by Clarissa Brocklehurst and Tom Slaymaker
- The World Health Organization (WHO) provides information about cholera and about diarrhea; it also provides information on water, sanitation and health and on drinking water quality (in several languages)
- The US Centers for Disease Control and Prevention provides information about cholera (including information on cholera control strategies) for the public, medical professionals and travelers
- The UK National Health Service (NHS) Choices website provides information about cholera
- MedlinePlus provides links to further resources about cholera (in English and Spanish)
- The not-for-profit organization Médecins Sans Frontières (MSF) is tackling several cholera outbreaks around the world; its website includes a description of the ongoing humanitarian crisis in the Democratic Republic of the Congo
- Personal stories about cholera in the Democratic Republic of the Congo are available from UNICEF, a not-for-profit organization that protects the rights of children and young people around the world
- The WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation (JMP) monitors global access to safe drinking water and improved sanitation; the JMP Update Report 2015 is available online
- The non-governmental organization Practical Action provides information and approaches and technologies for improving urban water supplies

6

The influence of tap water service on time and space patterns of admissions to the Uvira cholera treatment centre

Chapter 3 showed that cholera treatment centre (CTC) admission rates vary substantially between neighbourhoods in Uvira and over time, and that *V. cholerae* and enterotoxigenic *E. coli* (ETEC) are the most frequently detected pathogens amongst CTC patients. Chapter 4 documented water-related practices in households, and their association with a spatially explicit tap water service indicator. Probability of drinking water faecal contamination and quantity of water used at home for domestic activities, known risk factors for diarrhoeal diseases, can be predicted, albeit imperfectly, by a household's location relative to alternative sources of water and its tap water service indicator. Chapter 5 demonstrated that the daily quantity of tap water produced for the entire town of Uvira appears to have an important impact on the number of patients admitted to the CTC over the next 12 days. The following chapter explores the combined time and space patterns of CTC admissions in Uvira and their relationship with tap water availability.

6.1 Introduction

Faecal-oral transmission routes of cholera, and other diarrhoea-causing enteropathogens to a large degree, can be divided into two types of transmission cycles: a short-cycle, more direct route within a household and its close proximity, with an infected person contaminating food, drinking water and the very local environment (sanitation facilities, fomites), thus transmitting the infection to others that share these vectors of transmission; a long-cycle, more indirect route, with infection originating from a contaminated wider environment acting as a reservoir, through contaminated water at the source mostly **[1,2]**. Short-cycle transmission can also be referred to as "human-to-human" transmission while long-cycle transmission is also called "environmental-to-human" transmission, although the short- and long-cycle distinction implies a difference in time and space proximity between cases that is not strictly related to the exact transmission route.

Genomics studies have recently become a powerful tool to explore the dynamics of cholera strains transmission at global, regional and community levels, and shed light on the roles of both transmission cycles. The seventh cholera pandemic worldwide and especially in Africa has been linked to several occurrences of introduction of cholera strains, which originated in the South Asian subcontinent, through human movements **[3,4]**. Local and regional transmission of these strains was then established for varying lengths of time – "waves" **[4]**, and evidence of strains clustering within households and communities has been brought to light **[5]**. The role of aquatic reservoirs outside of South Asia, remains debated: although the above studies and others tend to downplay the role of environmental to human transmission **[6]**, findings in Haiti point to recurrent environmental evolution of strains identified from clinical cases **[7]**, and epidemiological data in Africa highlight a higher disease burden around the African Great Lakes region and the Lake Chad basin **[8-10]**.

Both short and long transmission cycles can be exacerbated by poor and intermittent access to safe water, with the short-cycle representing in large part "water-washed" transmission, prevented by ensuring tap water quantity is not a limiting factor for adequate hand, food, personal and domestic hygiene practices; and with the long-cycle, environment-mediated transmission, corresponding to "water-borne" transmission, prevented by eliminating the consumption of unsafe, potentially contaminated water from unsafe sources. Insufficient levels of residual chlorine at the tap may affect both, by allowing water contamination after collection during transport or storage (short-cycle) or by failing to inactivate existing pathogens within a piped distribution network (long-cycle).

Numerous modelling studies have been conducted to characterise the temporal and spatial patterns of diarrhoeal diseases incidence, often focusing on a particular age group or a specific pathogen **[11-17]**. Many of these have focused on cholera and outbreak situations, with a wide range of methodological approaches, from simple cluster detection to complex mechanistic mathematical formulations **[13,15,18-26]**. Accounting for both short and long-cycle transmissions has long been recognised as more realistic in representing the true phenomena underlying the observed data, although determining their respective contributions in different settings remains challenging **[18,27-29]**. Further challenges lie in the generalisability and interpretation of such models to formulate relevant recommendations for control interventions **[2,21,26,30-32]**.

"HHH4" models for infectious diseases, proposed by Held et al. **[33]**, build explicitly on the "endemic-epidemic" distinction and postulate that the number of cases occurring at a certain time in a particular area depends 1) on endemic transmission specific to this time and area (long-cycle transmission); 2) on the number of cases generated by previous cases in the same or neighbouring areas (short-cycle transmission). HHH4 models have the advantage of being formulated for discrete time and space units of data aggregation, such as surveillance data **[34,35]**. HHH4 models have been applied to and evaluated on various infectious diseases and settings, such as dengue and foot and mouth disease in China **[36]**, influenza-related pneumococcal disease in England **[37]**, measles, campylobacteriosis, rotavirus and Lyme borreliosis in Germany **[38,39]**, leptospirosis in Sri Lanka **[40]** and leishmaniasis in Afghanistan **[41,42]**.

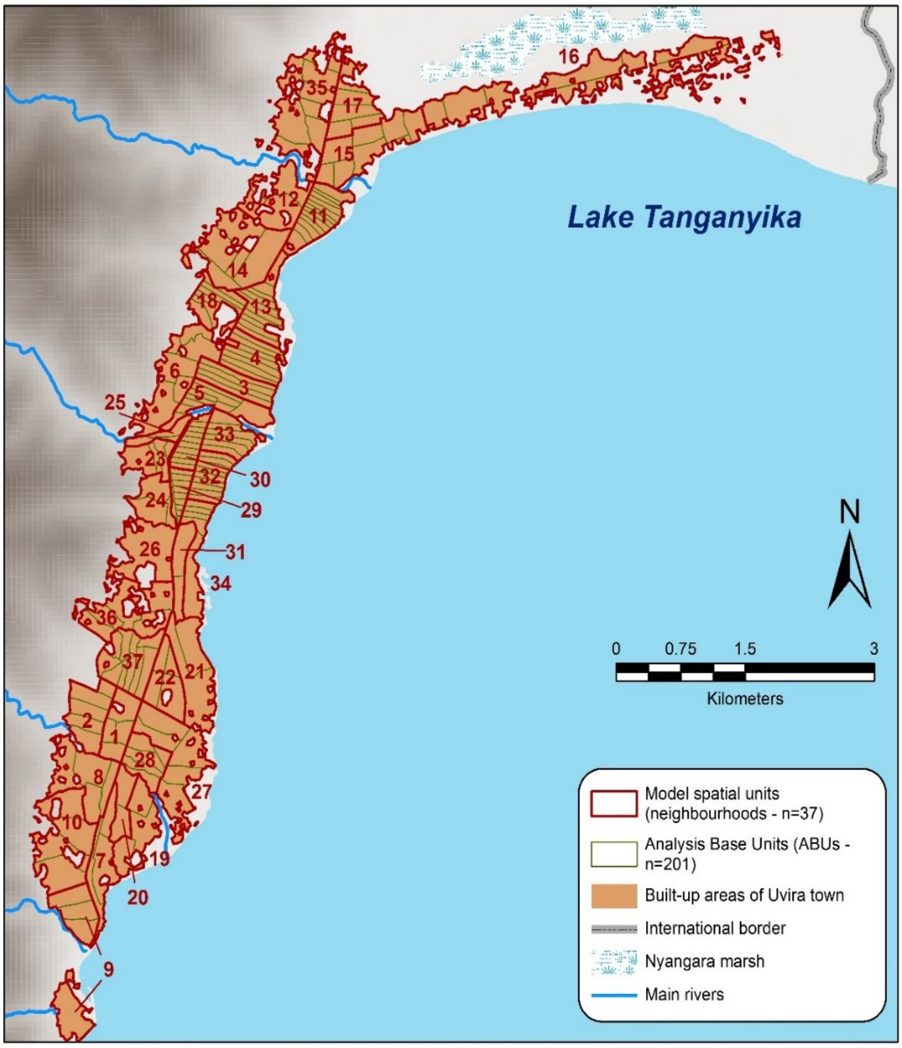
In the context of Uvira, where cholera and less outbreak-prone diarrhoeal diseases occur concurrently throughout the year, understanding the relative contributions of both short and long cycle transmissions could guide a more effective response to reduce acute diarrhoeal diseases incidence. The objectives of the present research were to explore whether time and space variations in tap water service in neighbourhoods of Uvira were associated with the observed time and space patterns of CTC admissions from these neighbourhoods, and whether the relative contributions of short- and long-cycle transmissions in this context could be elicited from the results.

6.2 Methods

The study site and general data collection methods are described in detail in the chapter *Data and methods: an overview*. Specific data sources and preparation for the present analysis are described below.

6.2.1 Data sources and preparation

Data described below were initially generated at the level of 201 Analysis Base Units (ABUs) that were aggregated into 37 larger units for analysis. These 37 units included between 1 and 12 ABUs (**Figure 6-1**). Admissions to the CTC and population estimates were simply tallied. Tap water service indicator, wealth index and distances were combined using population-weighted averages.



1	Kabindula_East	11	Kalundu_West	20	Kilibula_West	29	Rombe1_Centre
2	Kabindula_West	12	Kasenga_Northeast	21	Kimanga_East	30	Rombe1_North
3	Kakombe_East	13	Kasenga_Northwest	22	Kimanga_West	31	Rombe1_South
4	Kakombe_North	14	Kasenga_Southeast	23	Mulongwe_Centrenorth	32	Rombe2_Centre
5	Kakombe_South	15	Kasenga_Southwest	24	Mulongwe_Centresouth	33	Rombe2_North
6	Kakombe_West	16	Kavimvira_Centre	25	Mulongwe_North	34	Rombe2_South
7	Kibondwe	17	Kavimvira_East	26	Mulongwe_South	35	Rugenge
8	Kalundu_East	18	Kavimvira_North	27	Nyamianda_East	36	Songo_North
9	Kalundu_North	19	Kilibula_East	28	Nyamianda_West	37	Songo_South
10	Kalundu_South						

Figure 6-1 Map of the 37 spatial units (neighbourhoods) used for analysis. Grey areas represent topographical relief in Uvira's locality.

Admissions to the CTC

The number of patients admitted daily at the Uvira CTC, along with their residence location in or outside Uvira municipality, was collected from the CTC register and aggregated weekly by residence location from the 1st of January 2009 to the 28th of April 2019.

Population estimates

As described in *chapter 2*, we used population census data available for 10 time points between 2008 and 2017 for 14 administrative units (quartiers) and a detailed census for each street of the 14 quartiers available in November 2017, to estimate the weekly population for 201 ABUs over the study period.

Tap water service indicator

The tap water service indicator was derived following the method described in *Research Paper 2*, based on the amount of tap water invoiced in March 2018 for each of 3,685 georeferenced taps in Uvira (**Figure 6-2**). To summarise, the volume of tap water invoiced at each tap was spread over a circular surface with a chosen radius (here 250 m, 500 m and 750 m) surrounding the tap location. Over this surface, the tap water volume was distributed along a Gaussian function, with the highest value at the tap location and lowest at the edges. In order to account for non-circular edges of built-up areas delineation, a Diggle edge adjustment was used **[43]**.

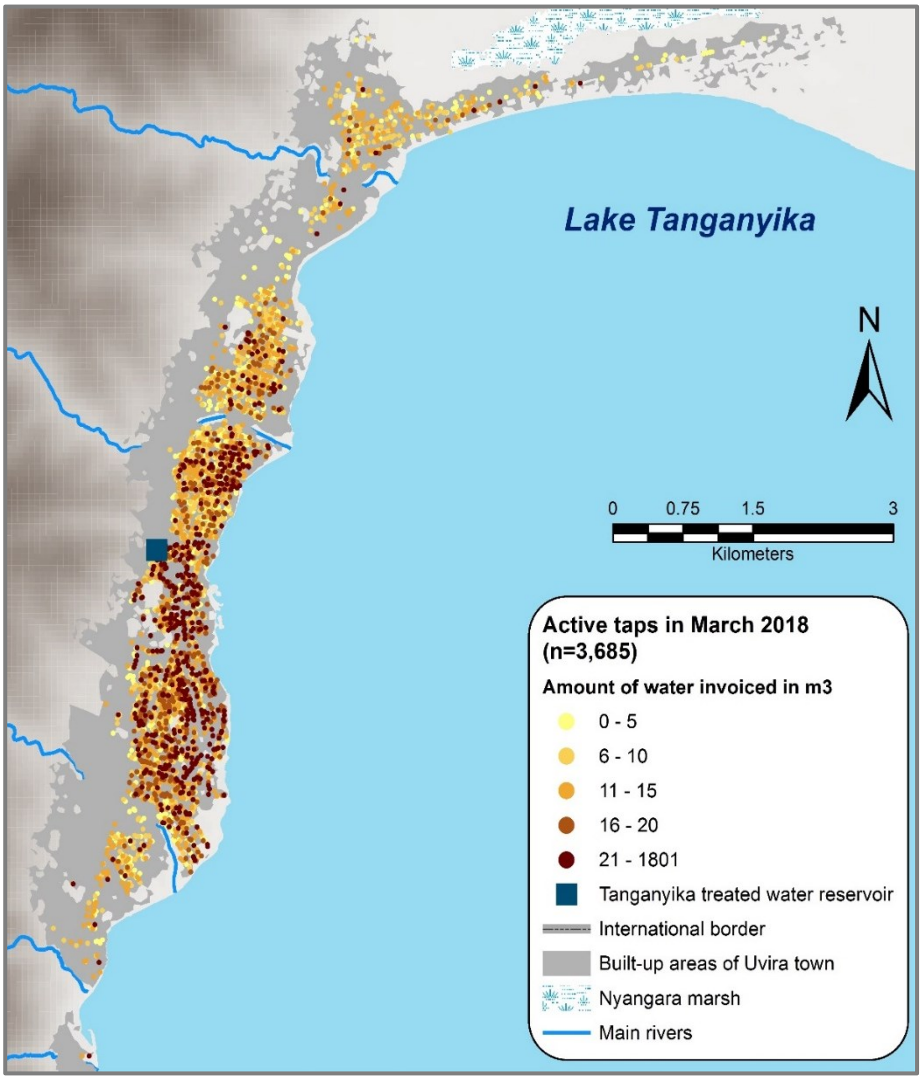


Figure 6-2 Map of the active taps (n=3,685) and amount of water invoiced in Uvira in March 2018

Amounts of tap water attributed to each cell within the circular area - cell size being defined by the resolution used for computation, here approximately 17 m x 21.4 m- by overlapping circular areas were added up, to result in a continuous map of tap water service in March 2018, expressed in litres per m^2 (**Figure 6-3**). This tap water service was then divided by the total amount of tap water distributed over the month of March 2018, to give the proportion of tap water distributed per m^2. A mean value for each ABU was extracted and divided by the mean population density in population per m^2 and multiplied by the mean weekly amount of tap water distributed over the study period, to obtain a tap water service indicator expressed in litres of tap water distributed per capita for each ABU.

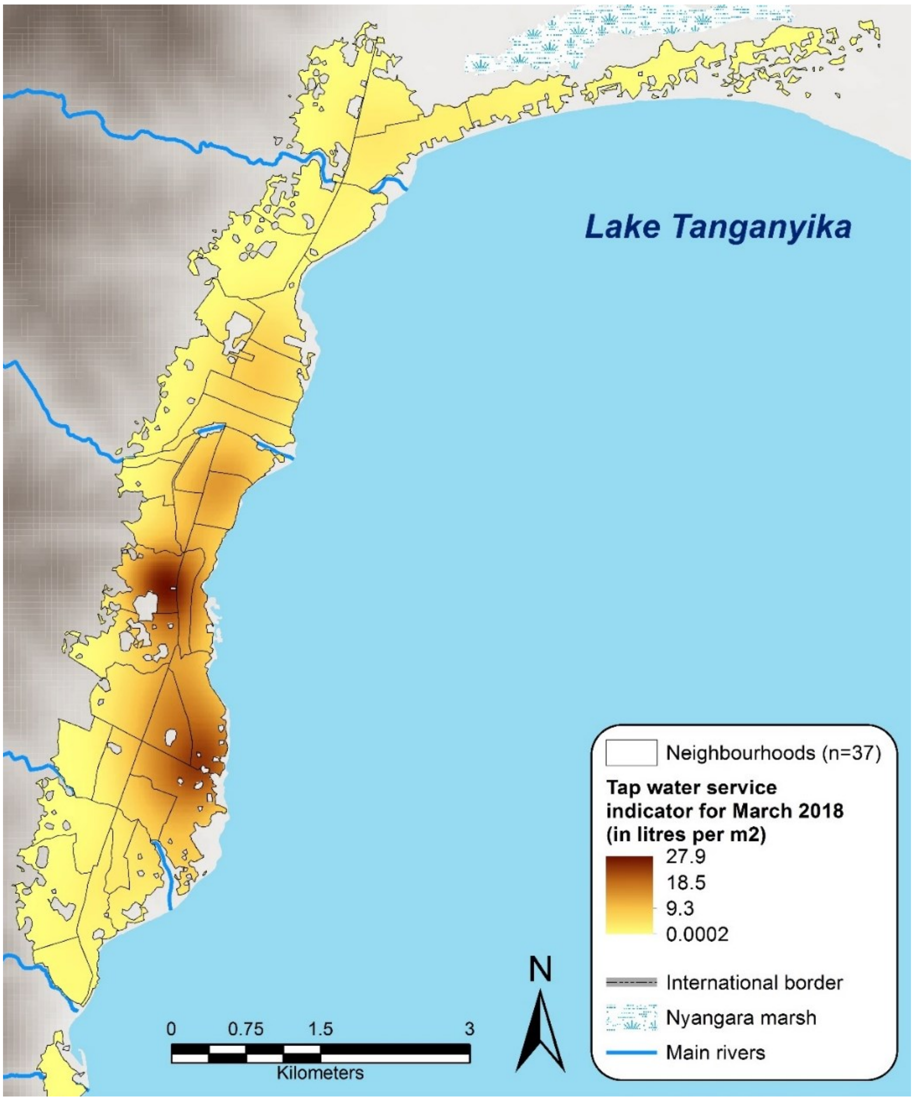


Figure 6-3 Tap water service indicator expressed in litres per m^2 (radius 250 m)

Volume of tap water produced and distributed

Recorded daily volumes of water treated at the Regideso plant were collected from the Regideso register and aggregated into weekly amounts. In order to account for losses between production and distribution, we multiplied the volume produced by a factor 0.8

to reach the distributed volume, as assessed from the total amount invoiced in March 2018 and previous plant productivity assessments (see chapter *Data and methods: an overview*, section *2.5.3*).

Wealth index

A wealth index for each ABU was derived from a continuous wealth index surface generated from 511 households' wealth indices previously obtained. In brief, a polychoric principal component analysis (PCA) was run on 17 household characteristics, mostly pertaining to asset ownership and dwelling construction but excluding characteristics related to water. The PCA-drived index for each of the 511 households was mapped and point data were interpolated by means of Empirical Bayesian Kriging (EBK).

This kriging method estimates spatial relationships between point data through iterative semi-variogram models, weighted and combined using Bayes' theorem. Several EBK parametrisations in terms of the range of points to be included in local models and potential overlaps were applied to a random selection of 60% of the households (training set) and assessed over the other 40% (testing set). EBK parameters that resulted in the smallest standard deviation of residuals for testing households were selected to generate a continuous map of wealth, from which the mean value for each ABU was extracted for the data analysis. Details of the wealth index building process and interpolation are available in **Appendix 4-3** and **Appendix 6-1**

Distances to the CTC and alternative water sources

Distance to the CTC was extracted from the Geographical Information System (GIS) as the sum of the shortest distance from each ABU centroid to the main roads (Routes Nationales RN 4 and RN 5), and the distance remaining on the main roads to the CTC.

Distances from the nearest river and from the lake were estimated as the shortest straight-line distance from each ABU centroid, adjusted for elevation by a Tobler function **[44]**.

6.2.2 Exploratory analysis

Initial analyses were carried out to explore both geographical patterns of counts of suspected cholera cases cumulated over the entire study period for each neighbourhood, and the temporal patterns of weekly counts for the entire town. In both cases, negative binomial regression of counts with log-transformed population offsets was used.

In the purely spatial analysis, the cumulative number of cases admitted to the CTC for each neighbourhood over the entire study period was modelled in terms of the neighbourhood's average tap water service. Confounding of the relationship between CTC admissions and tap water service by wealth and distance to alternative water sources, as suggested by the relationship between tap water service and households' water-related practices (see chapter *Tap water service and households' water-related practices*) was assessed.

In the purely temporal analysis, variations in admissions over time were examined in relation to the varying amounts of tap water distributed weekly, as suggested by a previous time-series regression analysis **[45]**. In the present analysis, seasonal variations in the weekly numbers of cases admitted to the CTC were taken into account by including pairs of harmonic functions (sine/cosine wave of time, with up to 3 pairs per year). Auto-correlation, expected for an outcome related to an infectious disease, was adjusted for by including a term for model residuals at lags showing partial auto-correlation greater than or equal to 0.1.

The Bayesian Information Criterion (BIC) was used to select which lags to include for tap water volume distributed. We limited the lags examined to 2 weeks and assessed model specification with Pregibon's test **[46]**.

Finally, the same formulation was used to run separate time-series regressions of admissions stratified into two sets of neighbourhoods, based on the mean tap water service indicator. A term for the log-transformed admissions rate in the other set was added to represent a potential influence of incidence in one set of neighbourhoods on the incidence in the other set.

6.2.3 Time and space modelling framework

In order to model both spatial and temporal effects of tap water distribution and service simultaneously, weekly counts of suspected cholera cases admitted at the CTC and residing in 37 neighbourhoods of Uvira were modelled using the framework proposed by Held et al for multivariate time-series of infectious disease areal counts and the R package "surveillance" **[33,47,48]**.

In this "endemic-epidemic" modelling framework, the expected number of cases μ_{it} in neighbourhood i at time t is defined by an endemic component ν_t multiplied by the neighbourhood population e_{it}, an autoregressive component $\lambda_t\, y_{i,t-1}$ based on observed counts at week t-1, noted $y_{i,t-1}$, in the same neighbourhood i and a spatial interaction component representing the influence of cases occurring in other neighbourhoods j at week t-1 (equation 1). Spatial interaction weights w_{ji} can be formulated in several ways (see below).

$$\mu_{it} = e_{it}\nu_{it} + \lambda_t\, y_{i,t-1} + \phi_t \sum_{j \neq i} w_{ji}\, y_{j,t-1} \qquad (1)$$

Held et al suggested inclusion in the endemic component of an intercept α, a trend parameter β and S sinusoidal harmonics of Fourier frequency $\omega_k = 2\pi k/52$ to model seasonal variations (equation 2a). Trend and seasonal variations of a similar form can also be included in the autoregressive and spatial interaction components λ_t and φ_t (equations 2b and 2c) to represent seasonal variations in the disease transmission pathways underlying those two components as well.

$$log(\nu_t) = \alpha_{end}{}^{\nu} + \beta_{end}\, t + \sum_{k=1}^{S_{end}} \gamma_{end\,k}\, sin(\omega_k t) + \sum_{k=1}^{S_{end}} \delta_{end\,k}\, cos(\omega_k t) \quad (2a)$$

$$log(\lambda_t) = \alpha_{ar}{}^{\lambda} + \beta_{ar}\, t + \sum_{k=1}^{S_{ar}} \gamma_{ar\,k}\, sin(\omega_k t) + \sum_{k=1}^{S_{ar}} \delta_{ar\,k}\, cos(\omega_k t) \quad (2b)$$

$$log(\phi_t) = \alpha_{ne}{}^{\phi} + \beta_{ne}\, t + \sum_{k=1}^{S_{ne}} \gamma_{ne\,k}\, sin(\omega_k t) + \sum_{k=1}^{S_{ne}} \delta_{ne\,k}\, cos(\omega_k t) \quad (2c)$$

Overdispersion of counts $y_{i,t}$ compared with that expected under the Poisson distribution can be accounted for by fitting the model with a negative binomial distribution with the same overdispersion ψ for all neighbourhoods, or by fitting the model with overdispersion ψ_i varying between neighbourhoods.

Spatial interaction weights between neighbourhoods w_{ij} can also be formulated in several ways. Spatial influence on a neighbourhoods' incidence can be assumed to be limited to directly adjacent neighbourhoods, in which neighbourhoods' adjacency order o_{ij} is equal to 1. Longer range influence can also be accounted for by applying a power law function to o_{ij} ($w_{ij} = o_{ij}{}^{-d}$ where d is a decay parameter to be estimated). The same structure can also be normalized to $w_{ij}/\sum_k w_{jk}$ (normalized power law).

Time- and/or space-varying covariates can also be included in any of the components to reflect dynamic conditions potentially affecting endemic incidence or auto-regressive and spatial transmission. An example of covariate inclusion in the auto-regressive component is given in equation 3, with lcd_i denoting the weekly volume of tap water available per person in neighbourhood i.

$$log(\lambda_{it}) = \alpha_{ar}{}^{\lambda} + \theta_{\frac{end}{lcpd}}\, lcpd_i + \sum_{k=1}^{S_{end}} \gamma_{end\,k}\, sin(\omega_k t) + \sum_{k=1}^{S_{end}} \delta_{end\,k}\, cos(\omega_k t) \quad (3)$$

In the present study, a 1-week interval was deemed appropriate for suspected cholera cases, considering that the average incubation period for cholera is estimated to 1 – 3 days, with 95% of cases under 5 days, and most infected people are infectious for up to 2 weeks **[49,50]**. Incubation and infectious periods are broadly similar for ETEC, the second most frequently detected enteropathogen in Uvira CTC patients (see chapter *Cholera and other enteric infections amongst patients admitted to the cholera treatment centre*, section *3.2.3*).

6.2.4 Model development and selection

A base model was first developed with a formulation informed by the exploratory analysis. In order to avoid overparameterising the base model, the BIC was used to identify the best fitting spatial interaction structure between neighbourhoods and the most parsimonious number of harmonics in components' seasonal variations.

Potential time- and space-covariates to be considered were suggested by the exploratory analysis. A backward selection strategy based on likelihood ratio tests (threshold taken p=0.05) was used to select covariates in the three components. Covariates to be considered were identified by the exploratory analysis and included the estimated weekly volume of tap water available per person, at lags 0, 1 and 2 and averaged over the entire study period, as well as an interaction term between linear time trend.

We expected wealth and distances to alternative water sources to confound any causal relationship between tap water service and CTC admissions, as observed in the relationship between tap water service indicator and water-related practices in households (see chapter *Tap water service and households' water-related practices*). They were therefore adjusted for in the model by adding wealth index, distance to the lake and distance to the nearest river as covariates in each component.

6.2.5 Residual error and attributable fraction estimates

The absolute residual error, calculated as the difference between fitted and observed number of CTC admissions, was used as a simple measure of model fit to the data, across study weeks or neighbourhoods.

Assuming that the association between tap water availability and CTC admissions in the model adjusted for confounders represents causation, the number of cases attributable to a tap water availability lower than a chosen reference value was estimated. This number of attributable cases was then divided by the total number of predicted cases to obtain the attributable fraction.

6.2.6 Sensitivity analysis

A sensitivity analysis was conducted to evaluate the robustness of results to changes in model specification. We examined the effect of changes in the number of harmonics in seasonal variations and overdispersion parameters. We also assessed whether changing the tap water service estimate to radiuses 500 m and 750 m and applying a population estimate correction – suggested by an outlier identification at the exploratory analysis stage - had any effect.

6.3 Results

6.3.1 Admissions to the CTC

A total of 13,425 patients were admitted at Uvira CTC between the 1st January 2009 and the 28th of April 2019. 12,421 of these patients (92.5 %) were residing at the time of admission in one of the 37 neighbourhoods used in the present analysis. 1,004 records were excluded: 746 patients were not resident in Uvira at the time of their admission, and 258 (1.9 % of total) records were missing patients' residence information.

The cumulative incidence of CTC admissions for the entire town over 539 weeks was 567.9 cases per 10,000. Cumulative CTC admissions by neighbourhood of residence are shown in **Figure 6-4**.

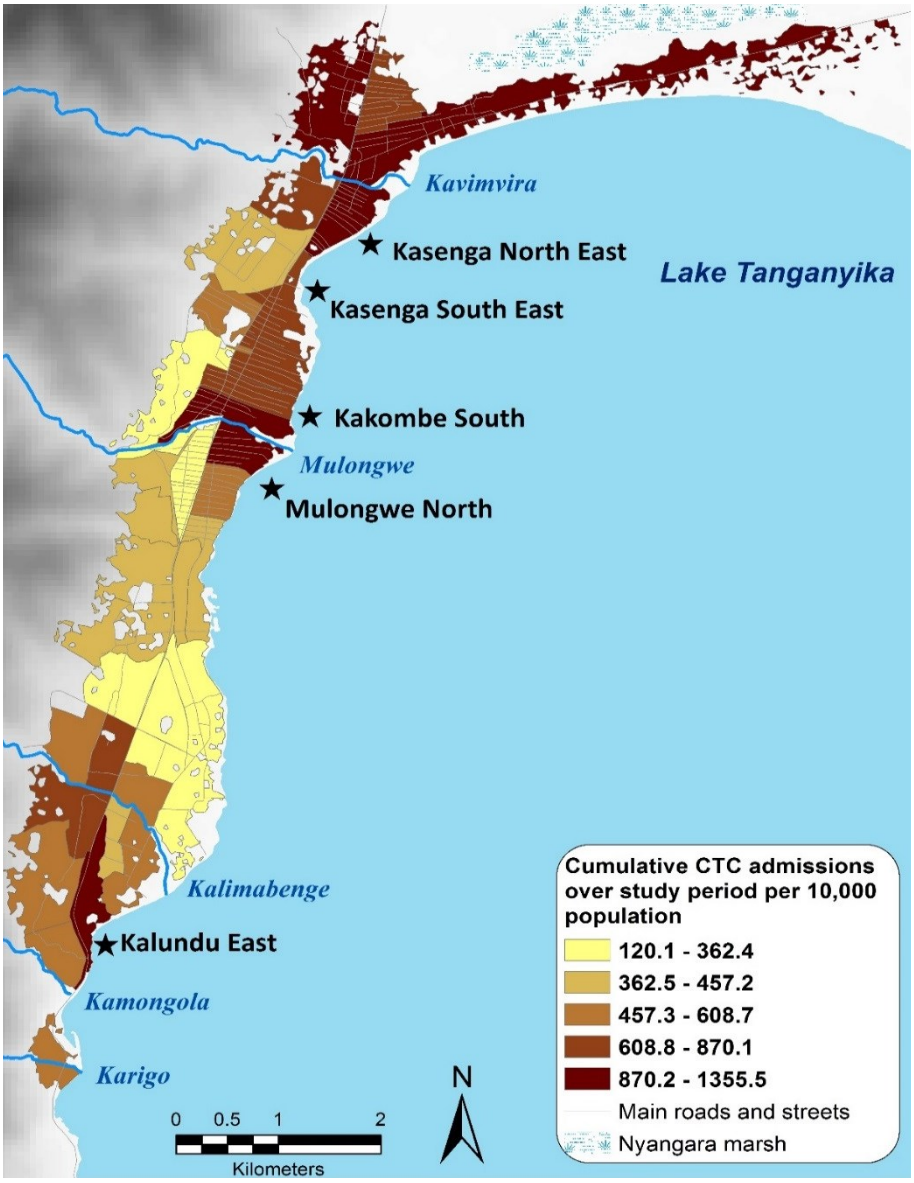


Figure 6-4 Map of cumulative CTC admissions per 10,000 population over 539 weeks, in 37 neighbourhoods of Uvira. Colour scale represents quintiles. Grey areas represent topographical relief in Uvira's locality.

For the entire town, weekly incidence ranged from 0 to 7.6 cases per 10,000, with a median of 0.75 cases per 10,000 (Interquartile range IQR 0.39 – 1.4). There were only three weeks in which there were no CTC admissions of patients residing in Uvira. Figure 6-6 shows weekly admissions per 10,000 for Uvira over the study period.

The highest recorded weekly admission rates occurred in neighbourhoods Kalundu East, Mulongwe North and Kasenga Southeast, with more than 40 admissions per 10,000 on one occasion. Only Kasenga Northeast and Kakombe South experienced admissions in more than half of the weeks included in the study period, with median weekly incidence rates of 1.38 and 1.28 per 10,000 respectively (**Figure 6-5**).

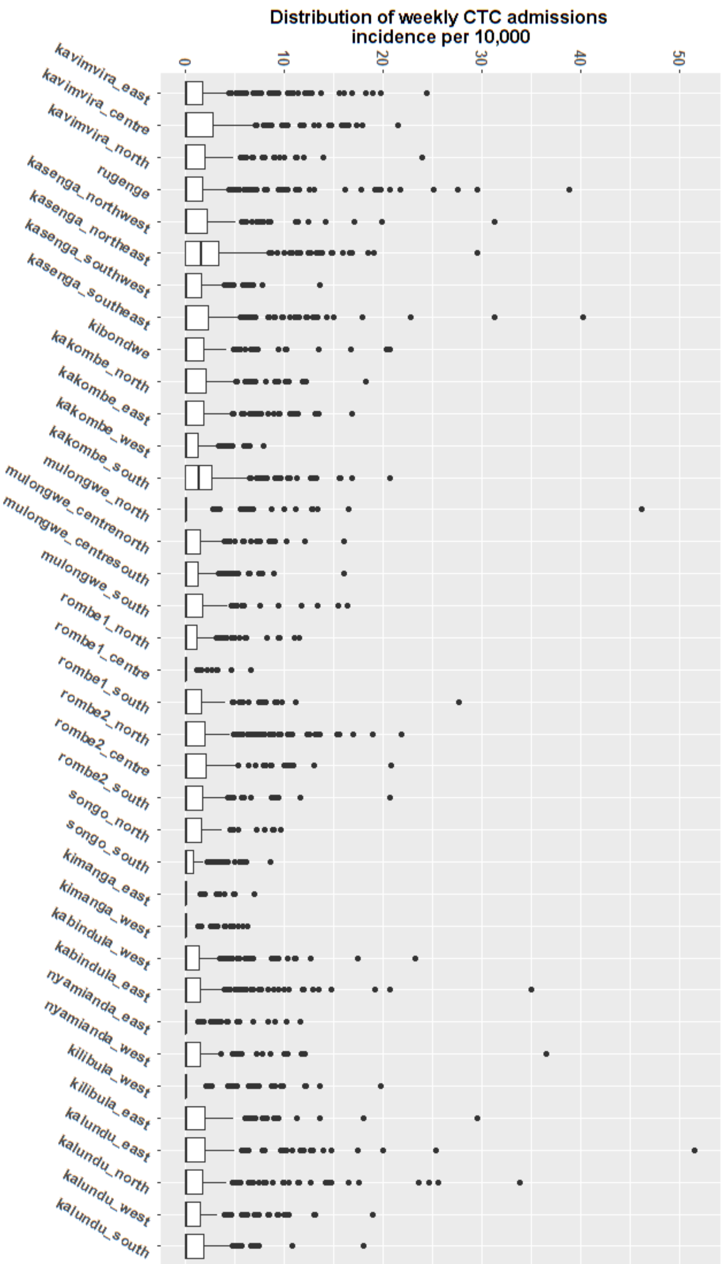


Figure 6-5 Box plots of weekly CTC admissions incidence per 10,000 over the study period, for each neighbourhood (n=37)

Weekly admissions for Uvira showed a broadly similar seasonal pattern each year, with consistently fewer admissions occurring in the month of May, while higher rates tended to occur during the months of February, March, August and September (**Figure 6-6** & **Figure 6-7**).

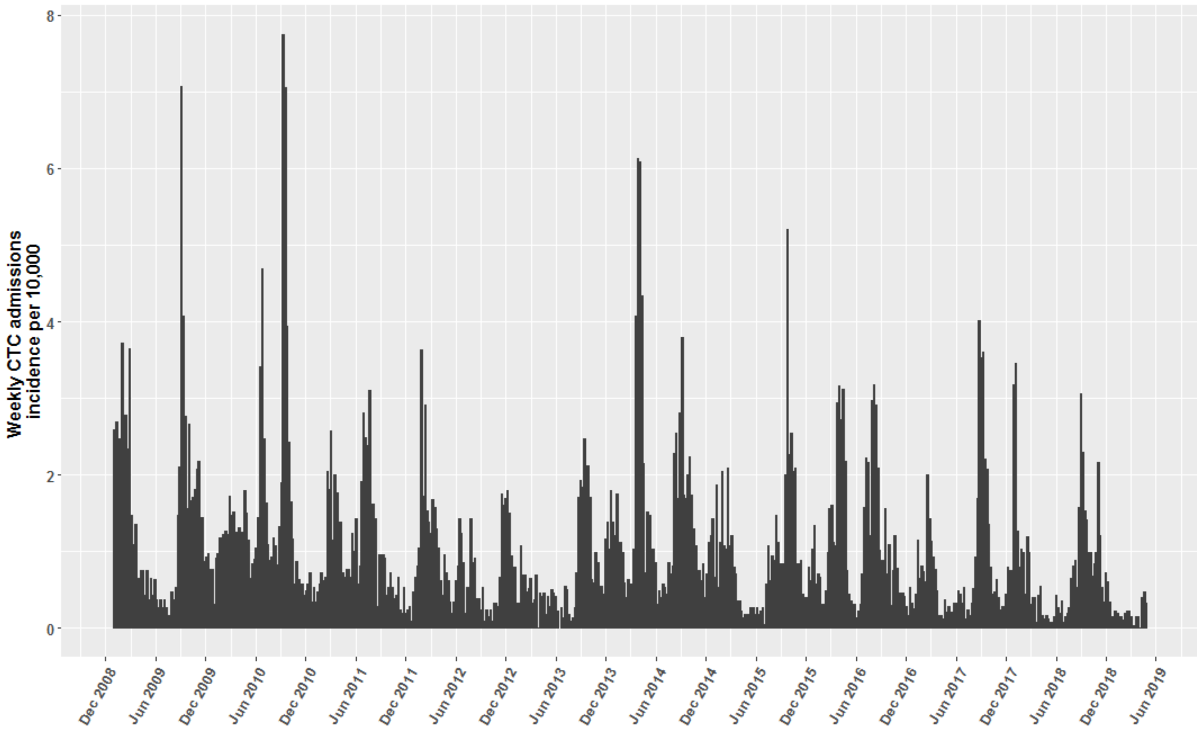


Figure 6-6 Weekly CTC admissions incidence per 10,000 during the study period, for all neighbourhoods

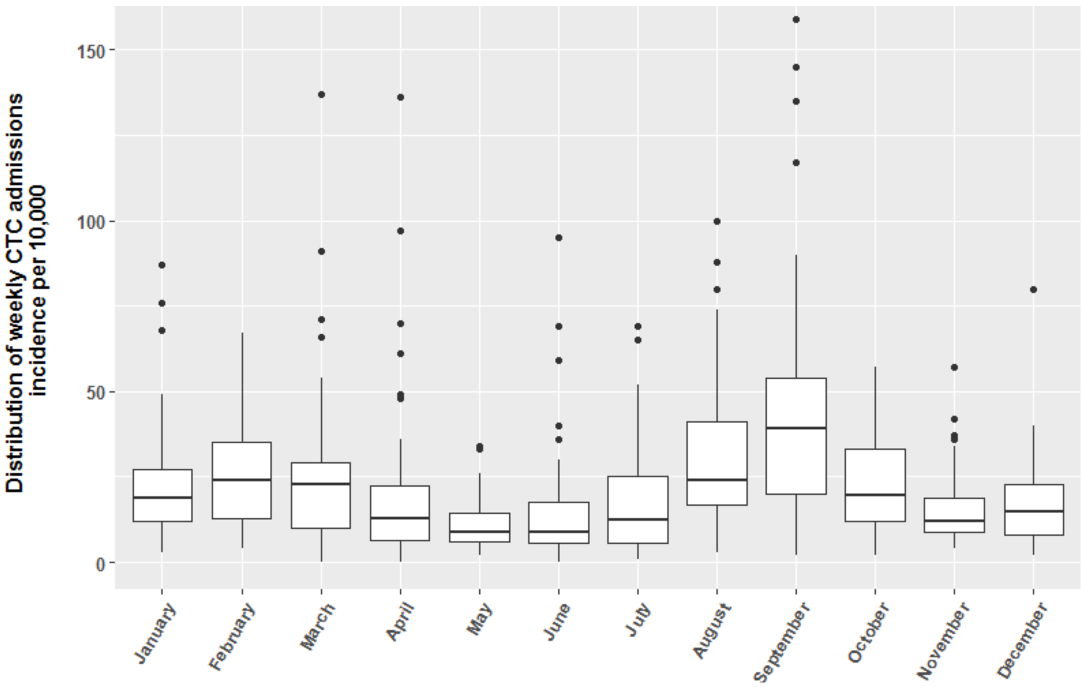


Figure 6-7 Distribution of weekly CTC admissions incidence per 10,000, for all neighbourhoods, by month of the year

6.3.2 Tap water service and water-related factors

Water volume distributed weekly varied hugely over the study period, from 3,756 m^3 to 27,102 m^3 with a median of 17,618 m^3 (IQR 14,294 m^3– 20,892 m^3). A plot of tap water volume distributed weekly over time is presented in **Figure 6-8**.

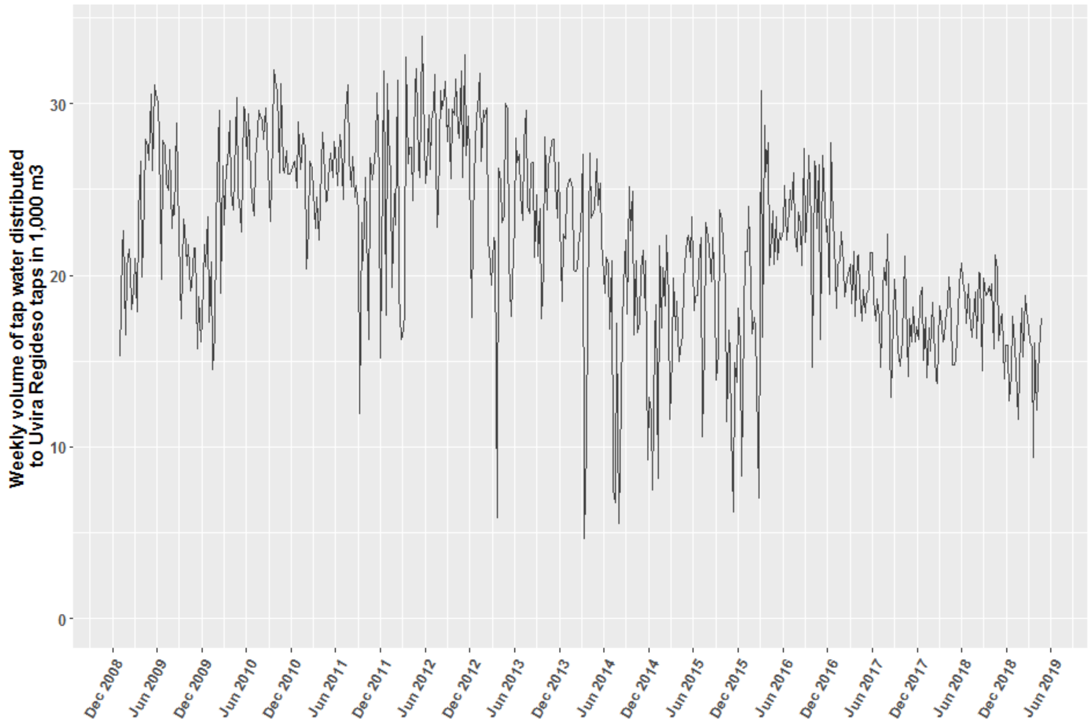


Figure 6-8 Weekly volume of tap water distributed to Uvira Regideso taps in 1,000 m3

On average over the study period, tap water service ranged from 1.2 litres per capita per day (lcd) to 41.3 lcd in Kabindula West and Kimanga East respectively, as shown in **Figure 6-9**. Over the entire study period and taking into account the weekly variations in the total amount of tap water distributed, this range extended to 0.3 to 64 lcd (median = 5.2 lcd; IQR 2.8 – 11.6).

The average wealth index of neighbourhoods varies between 0.34 in Rugenge and 1.26 in Kimanga East (**Figure 6-10**). The mean distance of neighbourhoods to the CTC, potentially associated with lower CTC admissions, ranges from 376 m (Mulongwe South) to 7,768 m (Kavimvira East). Neighbourhood distance to the nearest river ranged from 60 m (Mulongwe North) to 2,780 m (Kavimvira East), while distance to the lake ranged from 120 m (Rombe South) to 1,454 m (Kabindula West).

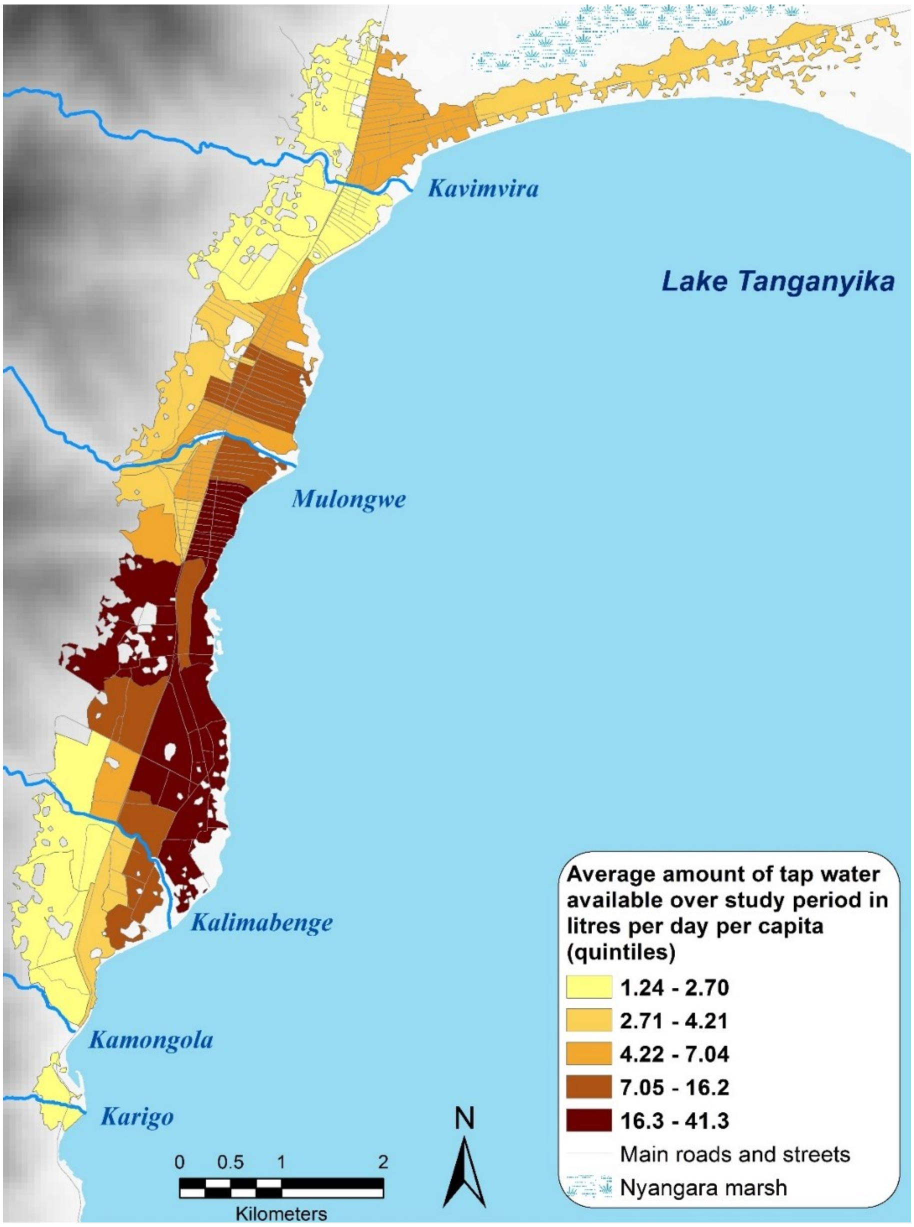


Figure 6-9 Average tap water service over the study period

As expected, wealth index and tap water service are correlated (Pearson's correlation = 0.66; p-value<0.001), highlighting the need to include wealth as a potential confounder of any relationship between tap water service and admissions incidence.

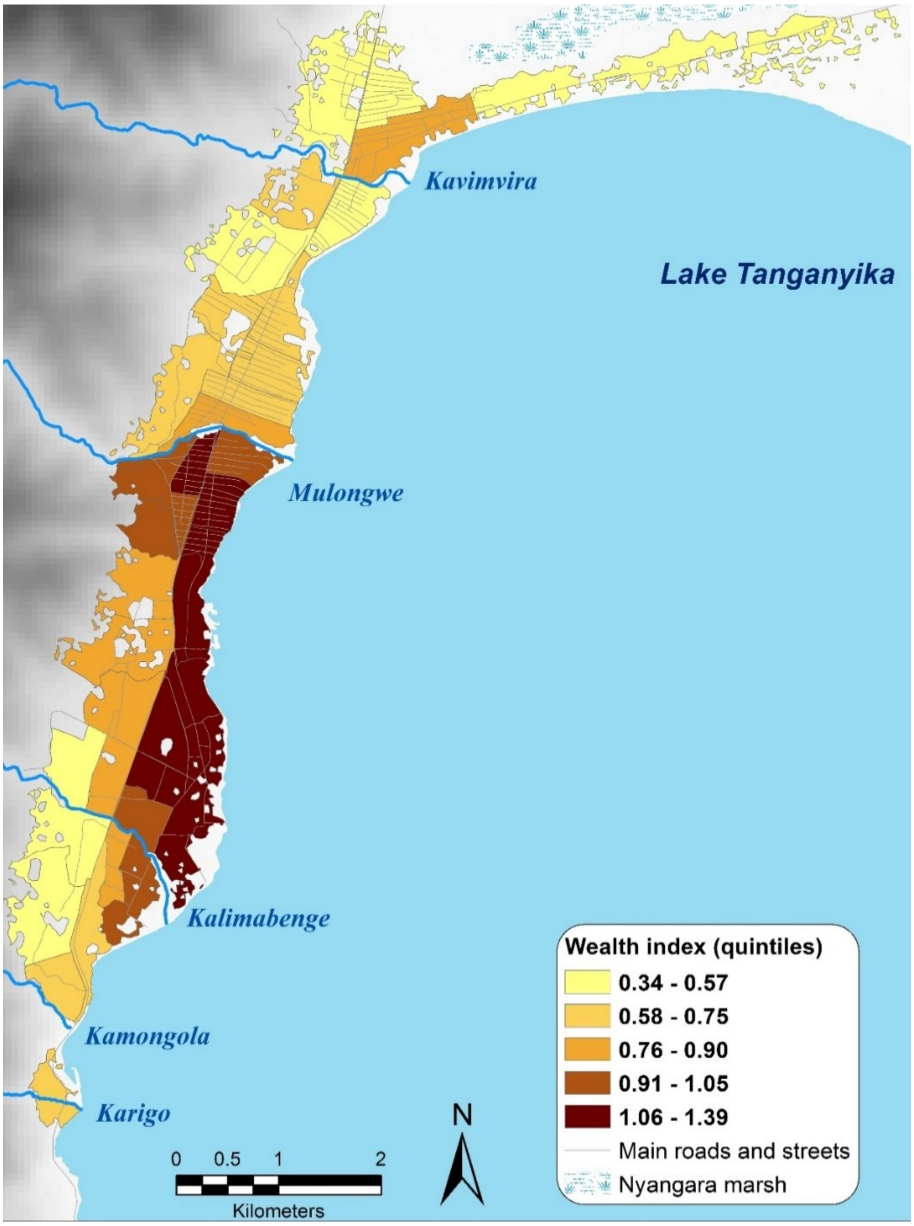


Figure 6-10 Neighbourhoods wealth indices

6.3.3 Exploratory analysis

Space

A plot of the log-transformed cumulative incidence rate of admissions to the CTC by neighbourhood against tap water service suggested a negative crude association (**Figure 6-11**). The results of fitting negative binomial models of the cumulative incidence, against tap water service only and adjusted for wealth index and distances to alternative water sources are presented in **Table 6-1**.

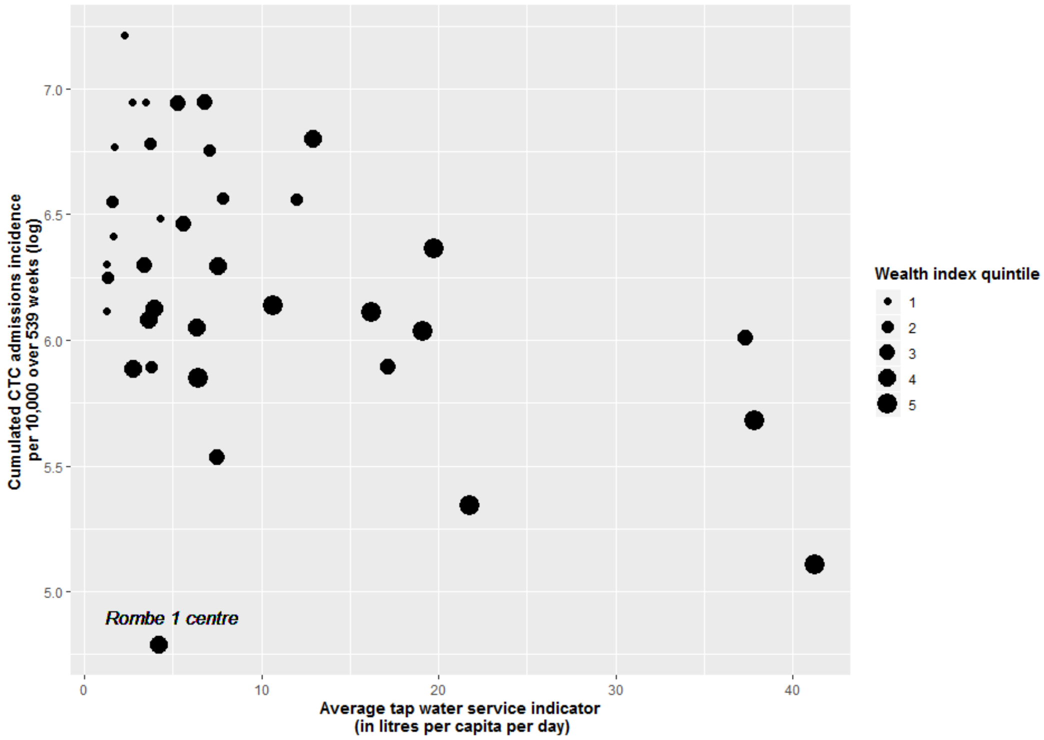


Figure 6-11 Plot of log cumulated CTC admissions per 10,000 over the entire study period (539 weeks) against tap water service. Wealth index quintile is represented by the size of the dot.

Unadjusted for wealth, a 5 lcd increase in average tap water service is associated with a 12% reduction in admissions (Incidence rate ratio IRR 0.88; 95% CI 0.82 – 0.95). The results are strongly confounded by wealth, with an IRR of 0.97 (95%CI 0.89 – 1.06) once adjusted for wealth index. They are further confounded by distances to alternative water sources (IRR = 0.99; 95% CI 0.92 - 1.07 when adjusted for wealth and distances to alternative sources). There was no evidence of misspecification of the models when assuming a linear relationship on the log scale.

A plot of the log-transformed cumulative incidence of admissions per neighbourhood as a function of distance to the CTC showed no evidence that this distance was negatively associated with admissions. On the contrary, the admissions rate appeared to increase with increasing distance to the CTC (**Figure 6-12**). Although this is not sufficient to totally exclude the possibility of distance to the CTC affecting the number of cases admitted, we assumed that we could omit this factor in further analyses.

The plot (**Figure 6-11**) of the log-transformed cumulative admission rate to the CTC by neighbourhood against average tap water service highlighted an outlier, identified as the neighbourhood Rombe 1 Centre, having both a low tap water service and the lowest admission rate. Data inspection revealed that Rombe 1 Centre is also an outlier in terms of population density with a very high estimate of 80,280 capita per km^2 while the mean population density for Uvira is 18,660 capita per km^2. This suggested that discrete neighbourhood populations may be overestimated for small and highly populated areas of Uvira. As the modelling framework does not support a missing unit, sensitivity analyses included modelling data based on a smoother spatial distribution of population estimates. Details of this approach are available in **Appendix 6-2** .

Table 6-1 Results of negative binomial models between cumulated CTC admissions over the entire study period and tap water service in 37 neighbourhoods

	Entire study period (539 weeks)		
	Estimate or IRR (95% CI)	**p-value**	**Covariate increment**
Unadjusted			
Cumulated CTC admissions per 10,000	557 (484 - 641)	p<0.001	
Average tap water service *	0.88 (0.82 - 0.95)	p<0.001	+5 lcd
Overdispersion	1.09		
Residual deviance	38 on 35 degrees of freedom		
Adjusted for wealth			
Cumulated CTC admissions per 10,000	545 (481 - 617)	p<0.001	
Average tap water service *	0.97 (0.89 - 1.06)	0.52	+5 lcd
Wealth index *	0.9 (0.84 - 0.96)	0.001	+0.1
Overdispersion	1.11		
Residual deviance	37.8 on 34 degrees of freedom		
Adjusted for wealth and distance to alternative water sources			
Cumulated CTC admissions per 10,000	532 (480 - 589)	p<0.001	
Average tap water service *	0.99 (0.92 - 1.07)	0.87	+5 lcd
Wealth index *	0.85 (0.8 - 0.9)	p<0.001	+0.1
Distance to the nearest river *	0.97 (0.95 - 0.993)	0.008	+100m
Distance to the lake *	0.93 (0.89 - 0.962)	p<0.001	+100m
Overdispersion	1.17		
Residual deviance	37.5 on 32 degrees of freedom		

** centred at the mean*

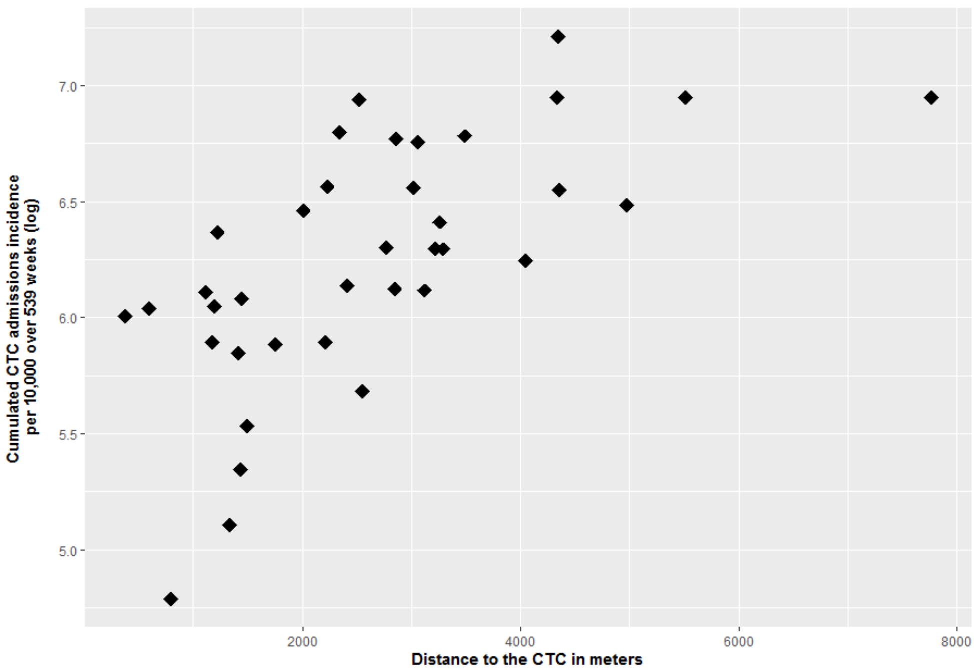


Figure 6-12 Plot of log cumulated CTC admissions per 10,000 over the entire study period (539 weeks) against distance to the CTC

Time

With a time-series regression of total weekly admissions to the CTC against time only, represented by a linear trend and harmonic seasonal functions, the best fit to the data was achieved with 2 highs and 2 lows per year. There was strong evidence of autocorrelation with admissions the previous week (lag 1), that was adjusted for. When weekly tap water volume distributed per capita was added, the best fit based on the BIC was obtained when including the volume at lag 1. There was no evidence of model misspecification when assuming a linear relationship on the log scale.

The model above provided evidence of both a decrease in incidence over the study period (IRR per year = 0.91; 95% CI 0.90 – 0.93) and of a decrease in incidence as the weekly amount of water distributed per capita increased (IRR for a 10 lcd increase 0.96, 95%CI 0.94 – 0.99).

Time and space

When the above model was fitted to separate time-series of admissions in neighbourhoods with lower or higher tap water service levels, temporal patterns and associations between volume of tap water distributed and weekly incidence of admissions varied noticeably between the two sets of neighbourhoods. Results of both non-stratified and stratified analyses are presented in **Table 6-2**.

At mean volume of tap water distributed, weekly admissions per 10,000 at the beginning of the study was estimated to be higher in neighbourhoods with lower tap water service (rate = 0.72; 95% CI 0.60 - 0.86) than in the neighbourhoods with a better tap water service (rate = 0.61; 95% CI 0.51 - 0.72). Evidence of a strong association between volume of tap water distributed per capita and rate of admissions was observed in the better served neighbourhoods only (IRR per 10 lcd increase = 0.95; 95% CI 0.92 – 0.97), with no evidence of a similar relationship in the neighbourhoods with a lower tap water service (IRR per 10 lcd increase = 1.00; 95% CI 0.97 – 1.02). Better served neighbourhoods saw the rate of admissions reduce faster over the study period (IRR per year = 0.91; 95% CI 0.89 - 0.93) than those with a poorer service (IRR per year = 0.98; 95% CI 0.95 – 1.00).

Table 6-2 Results of time-series regression of weekly CTC admissions and relationship with the weekly volume of tap water distributed, for all neighbourhoods (top) and stratified by tap water service (bottom)

		All neighbourhoods (n=37)		
Variables		Estimate or IRR (95% CI)	p-value	IRR increment
Weekly CTC admissions per 10,000		1.51 (1.37 - 1.67)	p<0.001	
Week in study period		0.91 (0.90 - 0.93)	p<0.001	+52 weeks
Tap water volume distributed per capita*		0.96 (0.94 - 0.99)	0.001	+10 lcd
Autocorrelation term (se; p-value)	Lag 1	0.56 (0.02; p<0.001)		
Overdispersion parameter		1.06		
Residual deviance		560.1 on 529 degrees of freedom		
BIC		-2765		

		Neighbourhoods with lower tap water service (n=19)			Neighbourhoods with higher tap water service (n=18)		
Variables		Estimate or IRR (95% CI)	p-value	IRR increment	Estimate or IRR (95% CI)	p-value	IRR increment
Weekly CTC admissions per 10,000		0.72 (0.60 - 0.86)	p<0.001		0.61 (0.51 - 0.72)	p<0.001	
Week in study period		0.98 (0.95 - 1.00)	0.02	+52 weeks	0.91 (0.89 - 0.93)	p<0.001	+52 weeks
Tap water volume distributed per capita*		1.00 (0.97 - 1.02)	0.95	+10 lcd	0.95 (0.92 - 0.97)	p<0.001	+10 lcd
Incidence in other stratum (Lag 1)		2.59 (2.18 - 3.07)	p<0.001		2.58 (2.24 - 2.97)	p<0.001	
Autocorrelation term (se; p-value)	Lag 1	0.27 (0.03; p<0.001)			0.19 (0.03; p<0.001)		
	Lag 2	0.12 (0.03; p<0.001)			0.10 (0.03; p<0.001)		
Overdispersion parameter		1.06			1.1		
Residual deviance		559.2 on 526 degrees of freedom			588.5 on 526 degrees of freedom		
BIC		-2746.2			-2729		

** centred at the mean*

6.3.4 Spatio-temporal model

Basic model

Following the exploratory analysis, we first fitted a basic model with a linear time trend and two seasonal harmonic pairs per year in each component, and a negative binomial distribution of admission counts with a single overdispersion parameter for all neighbourhoods. Based on the BIC, the best fitting formulation for spatial interaction between neighbourhoods with the above parameters was weighting contributions of neighbouring units along a normalised power law function up to the maximum adjacency order of 7. The combination of a single seasonal harmonic pair in the endemic component and two pairs in the spatial component led to a lower BIC. Similarly, the linear time trend terms in the autoregressive and spatial components did not improve the model fit to the data and were removed to reach model 0 formulation (**Equations 4a to 4d**).

$$\mu_{it} = e_{it}\nu_t + \lambda_t\, y_{i,t-1} + \phi_t \sum_{j\neq i} w_{ji}\, y_{j,t-1}$$

$$\textit{with}\ \mu_{it} \sim NegBin(\mu,\psi)\ \textit{and}\ w_{ij} = o_{ij}^{\ -d} / \textstyle\sum_k w_{jk}\,, \textit{where}\ 1 \leq o \leq 7 \qquad (4a)$$

$$log(\nu_t) = \alpha_{end}^{\ \nu} + \beta_{end}\, t + \gamma_{end}\ sin\left(\frac{2\pi}{52}\, t\right) + \delta_{end}\ cos\left(\frac{2\pi}{52}\, t\right) \qquad (4b)$$

$$log(\lambda_t) = \alpha_{ar}^{\ \lambda} \qquad (4c)$$

$$log(\phi_t) = \alpha_{ne}^{\ \phi} + \gamma_{ne1}\ sin\left(\frac{2\pi}{52}\, t\right) + \delta_{ne1}\, cos\left(\frac{2\pi}{52}\, t\right) + \gamma_{ne2}\ sin\left(\frac{4\pi}{52}\, t\right) + \delta_{ne2}\, cos\left(\frac{4\pi}{52}\, t\right) \qquad (4d)$$

Model 0 estimates an endemic weekly incidence of 0.4 admissions per 10'000 (95% CI 0.35 – 0.46) at the beginning of the study period for the entire town of Uvira, which decreases annually by 6% (IRR=0.94; 95% CI 0.92 – 0.95). The auto-regressive component estimates that each admission from a specific neighbourhood i at week t leads to 0.28 admissions (95% CI 0.26 – 0.3) from the same neighbourhood i at week t+1. The spatial interaction component estimates that each admission in neighbourhoods j at week t leads to 0.26 admission (95% CI 0.21 – 0.3) in neighbourhood i at week t+1, when neighbourhoods i and j are directly adjacent (o_{ij}=1). The influence of other neighbourhoods decreases with adjacency order along a normalised power law function with parameter d = 1.99 (95% CI 1.75 - 2.23).

Model 0 includes a single seasonal sinusoidal curve per year in the endemic component, with the endemic incidence at its highest at week 49 of each year, when it is estimated to be 1.27 times the mean annual endemic incidence. Endemic incidence is estimated to be lowest at week 23, at 0.79 times the mean annual endemic incidence. Two sinusoidal curves represent annual seasonal variations in the spatial influence between neighbourhoods, with higher values at weeks 8 and 32 of each year (factors 1.2 and 1.76 respectively). Spatial influence is estimated to be lowest at weeks 19 and 47 (factors 0.77 and 0.61 respectively). Seasonal variations for each component are visualised in **Figure 6-13**.

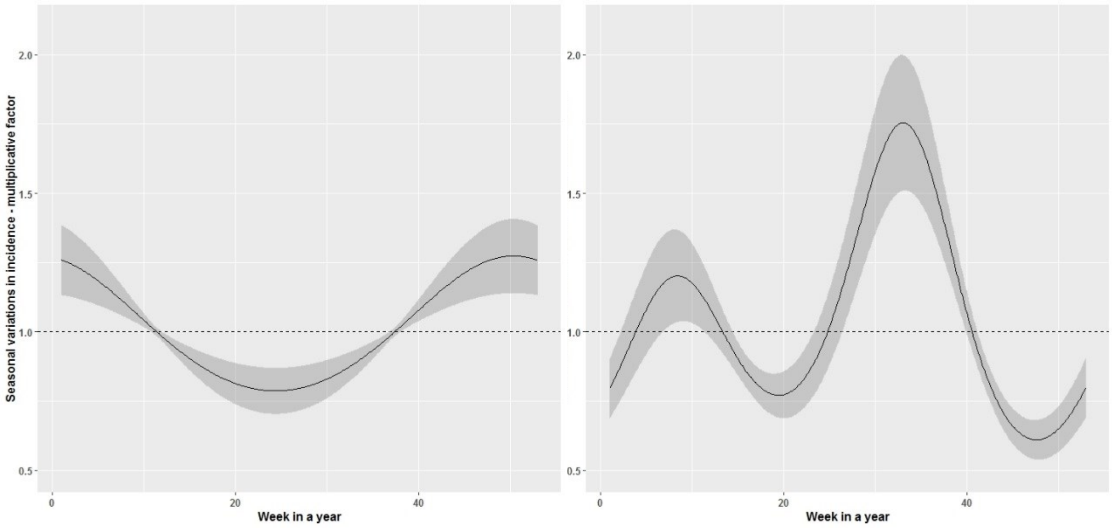


Figure 6-13 Seasonal variation factors estimated by model 0 (left: endemic component; right: spatial interaction component). Grey areas represent 95% confidence intervals.

The overdispersion parameter ψ=1.02 (95% CI 0.95 – 1.09) indicates that the variance in admissions is very slightly higher than the mean. Coefficients and parameters for model 0 are reported in **Table 6-3**.

Table 6-3 Parameters estimates for models 0,1 and 2

	MODEL 0		MODEL 1		MODEL 2
	Estimate or IRR (95% CI)	**IRR increment**	**Estimate or IRR (95% CI)**	**IRR increment**	**Estimate or IRR (95% CI)**
Endemic component					
Fixed incidence per 10,000	0.4 (0.35 - 0.46)		0.41 (0.36 - 0.47)		0.41 (0.36 - 0.47)
Annual trend [1]	0.94 (0.92 - 0.95)		0.919 (0.9 - 0.939)		0.916 (0.898 - 0.935)
Time trend x mean tap water volume distributed [2]	na	52 weeks x 4.8 lcd	0.981 (0.968 - 0.994)	52 weeks x 4.8 lcd	0.983 (0.972 - 0.994)
Mean tap water volume distributed [3]	na	50% (4.8 lcd)	0.95 (0.9 - 1.01)	50% (4.8 lcd)	1.04 (0.97 - 1.11)
Wealth Index [4]	na		na	0.1	0.86 (0.83 - 0.9)
Distance to the lake [4]	na		na	100m	0.91 (0.88 - 0.93)
Distance to the nearest river [4]	na		na	100m	1 (0.99 - 1.01)
Autoregressive component					
Fixed incidence	0.28 (0.26 - 0.3)		0.26 (0.24 - 0.29)		0.24 (0.22 - 0.26)
Tap water volume distributed the previous week [5]	na	10 lcd	0.84 (0.76 - 0.94)	10 lcd	0.78 (0.67 - 0.91)
Wealth Index [4]	na		na	0.1	1.02 (0.97 - 1.08)
Distance to the lake [4]	na		na	100m	0.97 (0.94 - 1.007)
Distance to the nearest river [4]	na		na	100m	1.02 (1 - 1.03)
Spatial interactions component					
Incidence from direct neighbouring areas	0.26 (0.21 - 0.3)		0.23 (0.19 - 0.28)		0.19 (0.15 - 0.23)
Tap water volume distributed the previous week [5]	na	10 lcd	0.71 (0.61 - 0.82)	10 lcd	0.74 (0.64 - 0.85)
Mean tap water volume distributed [3]	na	50% (4.8 lcd)	1.13 (1.06 - 1.21)	50% (4.8 lcd)	1.27 (1.18 - 1.37)
Wealth Index [4]	na		na	0.1	0.87 (0.84 - 0.9)
Distance to the lake [4]	na		na	100m	0.99 (0.97 - 1.02)
Distance to the nearest river [4]	na		na	100m	0.94 (0.92 - 0.96)
Neighbourhood adjacency order decay	1.99 (1.75 - 2.23)		1.73 (1.51 - 1.95)		1.43 (1.22 - 1.63)
Other model parameters					
Overdispersion parameter	1.022 (se 0.034)		0.989 (se 0.035)		0.927 (se 0.036)
BIC	37285		37164		36921

Table notes: 1) centered at week 1; 2) centered at week 1 & mean tap water volume distributed over time and space (9.6 lcd); 3) expressed as relative difference with and centered at mean tap water volume distributed over time and space (9.6 lcd); 4) centered at the mean over space; 5) centered at mean tap water volume distributed over time and space (9.6 lcd)

Models including tap water service

Model 1 was reached after a backward selection strategy, with an initial model including in each of the components the estimated weekly volume of tap water available per person in each neighbourhood, at lags 0, 1 and 2 and averaged over the entire study period per neighbourhood, as well as an interaction term with the linear time trend in the endemic component. After iterative removal of terms with p-value > 0.05 for likelihood ratio test, Model 1 includes the estimated volume of tap water available per person the previous week (lag 1) in both the auto-regressive and the spatial components. It also includes the estimated weekly volume of tap water averaged over the entire study period in the endemic and spatial components. An increase in tap water availability the previous week (lag 1) is associated with a decrease in the auto-regressive and spatial components. An increase in the average volume of tap water over the study period is associated with a slight decrease in endemic incidence and an increase in spatial component. Model 1 also suggests that the linear decrease in endemic incidence over time is modified by average tap water availability, with better served neighbourhoods experiencing a sharper decrease over the study period. Adding the above tap water availability covariates improved the BIC, from 37,285 to 37,164. Coefficients and parameters for model 1 are reported in **Table 6-3**.

Model 1 indicates that a neighbourhood (denoted I_0) with an average overall tap water availability (9.6 lcd), had a weekly endemic incidence of 0.41 CTC admission per 10,000 at the beginning of 2009, which decreased by 8.1 % (95% CI 6.1 – 10 %) annually over the study period (2009 – 2019). In this neighbourhood I_0, one admission occurring in a directly adjacent neighbourhood led to 0.23 (95% CI 0.19 – 0.28) admissions the following week when the amount of tap water available the previous week did not vary from the average amount. Under the same conditions, one admission from this neighbourhood I_0 gave rise to 0.26 (95% CI 0.24 – 0.29) admissions from the same neighbourhood I_0 the following week.

Model 1 implies that a neighbourhood I_1 with 50 % or 4.8 lcd more tap water availability on average than the neighbourhood I_0 described above, would have a 5 % lower endemic incidence at week 1 (ie 0.39 CTC admission per week; 95% CI 0.32 – 0.48). It also suggests that this weekly incidence decreased 0.19 % faster annually (ie 9.8 % annual decrease; 95% CI 6.7 – 12.8 %). A case in a directly adjacent neighbourhood contributed slightly more to the following week's incidence in this better served neighbourhood I_1 than in the average one I_0 (0.26 admission; 95% CI 0.2 – 0.34) when the amount of tap water available the previous week did not vary from the average amount. If the amount of tap water available the previous week was 10 lcd higher than the average week, in either neighbourhood I_0 or I_1, an admission from a directly adjacent neighbourhood had less influence on the number of cases occurring the following week (IRR = 0.71; 95% CI 0.61 – 0.82; ie 0.16 cases in I_0 and 0.19 in I_1). An increase of 10 lcd above the average the previous week decreased the number of cases observed the next week in the same neighbourhood (IRR = 0.84; 95% CI 0.76 – 0.94; no difference between neighbourhoods with different average tap water availability).

In model 1, the influence of incidence in other neighbourhoods decreases less sharply with adjacency order (power law function parameter d = 1.73 vs 1.99 in model 0).

In model 2, wealth index and distances to the lake and the nearest river were added to each component. Model 2 suggests strong confounding of the effect of average tap water availability by wealth and distance to alternative water sources. This effect decreases in the endemic component but increases in the spatial component. Evidence for confounding of the effect of weekly availability of tap water by wealth and distance to alternative water sources is present in the autoregressive component, with a larger decrease in autoregressive component as the amount of tap water available the previous week increases. There is little evidence for confounding of the effect of weekly tap water availability in the spatial component. Coefficients and parameters for model 2 are reported in **Table 6-3**.

Assuming the two hypothetical neighbourhoods I_0 and I_1 used previously share the same mean wealth index, mean distance to the lake and mean distance to the nearest river, model 2 suggests that they had an endemic weekly incidence at the beginning of the study of 0.41 and 0.43 per 10,000 respectively (95% CI 0.36 – 0.47 and 0.35 – 0.52). Endemic incidence decreased annually by 8.4 % and 10 % respectively (95% CI 6.5 – 9.2 % and 7 – 12.7 %). One admission in a directly adjacent neighbourhood led to 0.19 admissions in neighbourhood I_0 and 0.24 in neighbourhood I_1 when tap water availability the previous week did not vary from the mean. As in model 1, a weekly tap water availability the previous week 10 lcd higher than the mean reduced the influence of an admission in directly adjacent neighbourhoods by 26 % (IRR 0.74; 95% CI 0.64 – 0.85) in either neighbourhood. A case admitted from either neighbourhood I_0 or I_1 will lead to 0.24 admission the following week (95% CI 0.22 – 0.26) at the mean weekly tap water availability. An increase of 10 lcd the previous week will reduce this autoregressive component by 22 % in either neighbourhood (IRR = 0.78; 95% CI 0.67 – 0.91).

Independently of tap water availability and distances to alternative water sources, endemic and spatial interaction components decrease noticeably with increasing neighbourhood wealth, with a wealth index increase of 0.1 (out of a spanning range of 0.92) leading to an IRR of 0.86 and 0.87 respectively (95% CIs 0.83 – 0.9 and 0.84 – 0.9 respectively). For two neighbourhoods at mean values for all characteristics but wealth index, one of average wealth and the other 0.1 better off, it represents a difference in weekly endemic incidence at week 1 of 0.41 per 10,000 vs 0.36 per 10,000 in the better-off. An admission in a directly adjacent neighbourhood leads to 0.19 admissions the following week, vs 0.17 in the better-off. There is no evidence however of association between the autoregressive incidence and wealth independently of tap water availability and distances to alternative water sources (IRR for 0.1 increase = 1.02; 95% CI 0.97 – 1.08).

Once tap water availability and wealth index have been adjusted for, model 2 suggests that increasing distance to the lake is associated with a reduction in endemic incidence, while increasing distance to the nearest river is associated with an increasing auto-regressive component and a decreased spatial component. Adjusting for wealth and distances to alternative water sources improves noticeably the goodness of fit of model 2 (BIC = 36,921 vs BIC = 37,164 for model 1). Seasonal sinusoidals in model 1 and 2 are similar to that of model 0 (**Figure 6-14**).

The overdispersion parameter estimates is marginally lower than model 0 for model 1 (ψ = 0.99; 95% CI 0.92 - 1.06) and even lower for model 2 (ψ = 0.93; 95% CI 0.87 – 1.00).

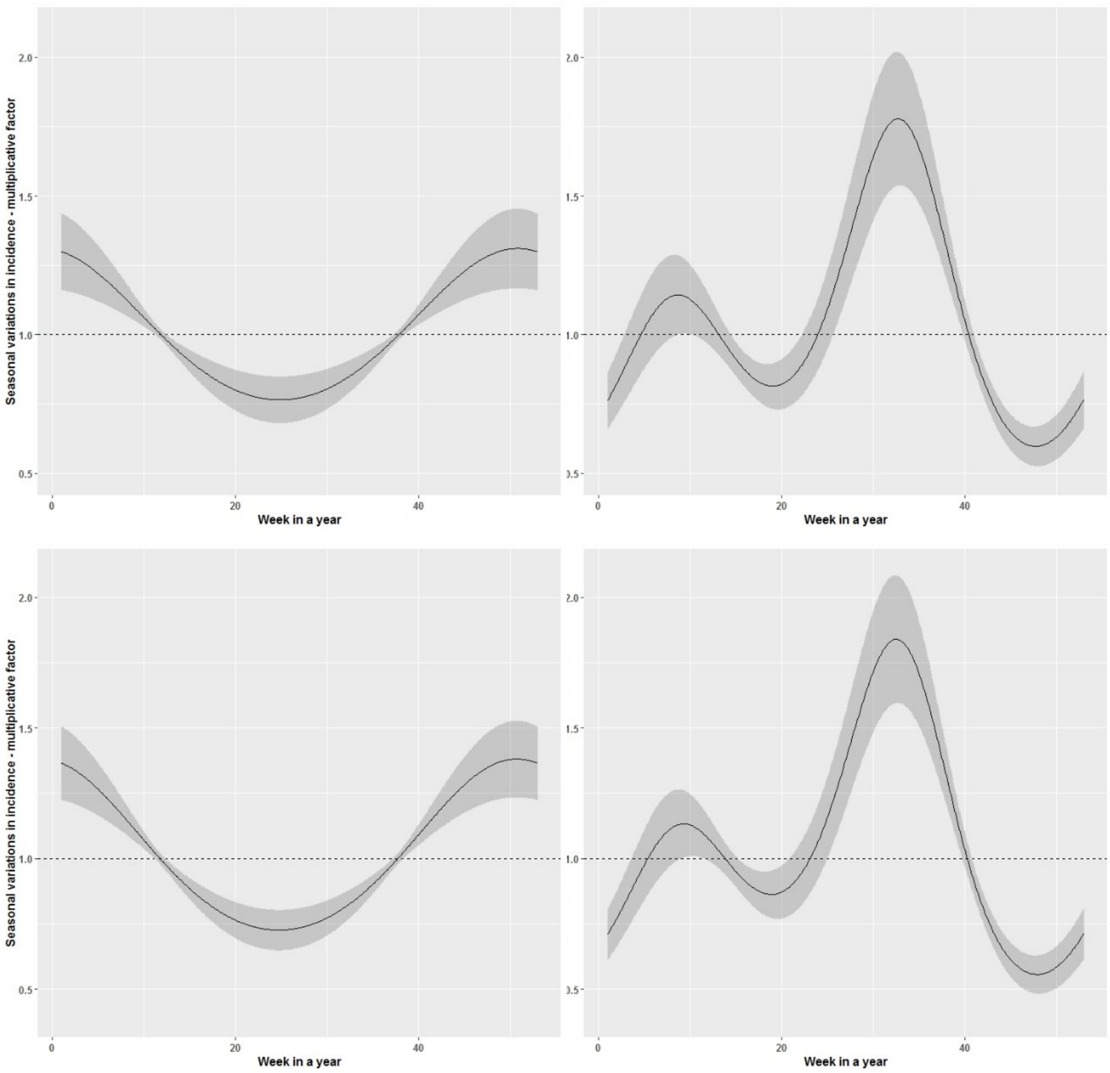


Figure 6-14 Seasonal variation factors estimated by models 1 (top) and 2 (bottom) for the endemic component (left) and the spatial component (right). Grey areas represent 95% confidence intervals.

A plot of observed and fitted values for weekly CTC admissions stratified by component is available in **Figure 6-15**. **Figure 6-16** represents the total observed and fitted weekly CTC admissions within the model uncertainty envelope.

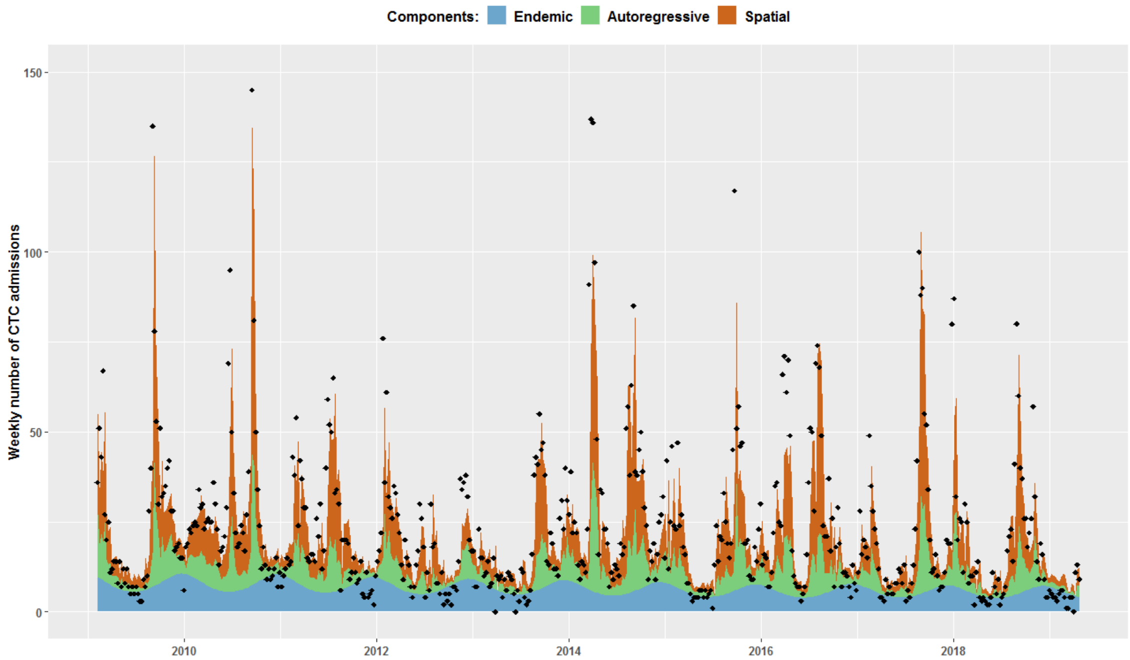


Figure 6-15 Observed (♦) and fitted weekly number of CTC admissions for all neighbourhoods and stratified by component, as estimated by model 2

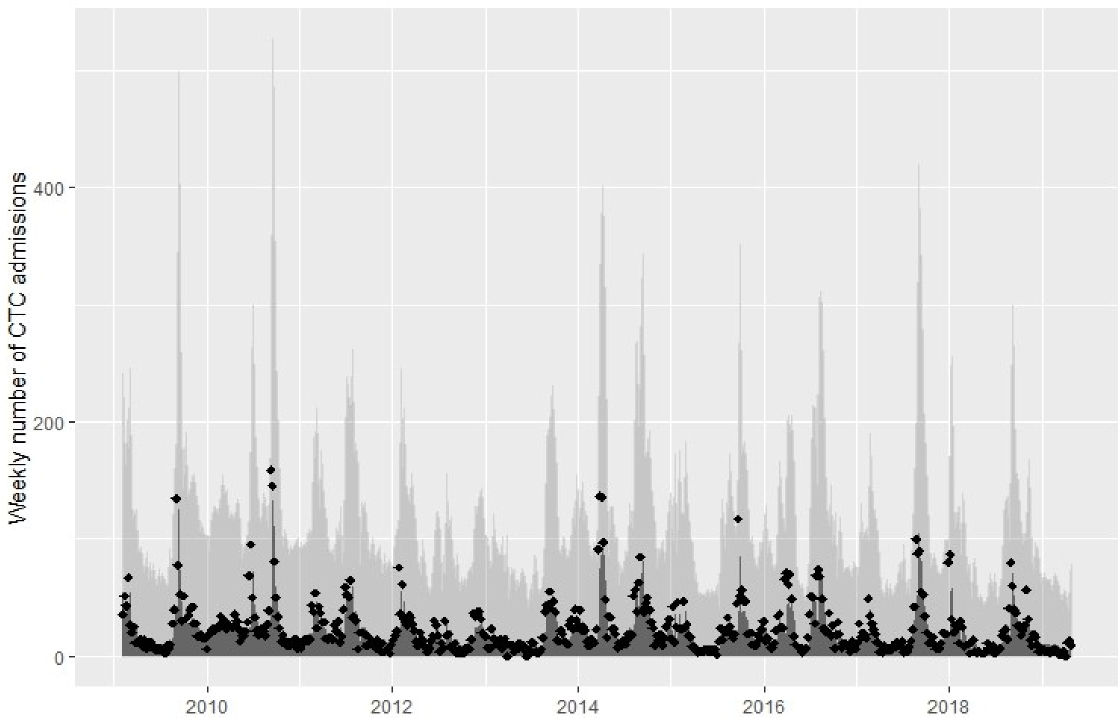


Figure 6-16 Observed (♦) and fitted weekly number of CTC admissions for all neighbourhoods, as estimated by model 2. Light area represents 95% confidence interval for fitted values (lower boundary at 0 at all points).

Overall, model 2 predicts that endemic, autoregressive and spatial components contribute to 28.7 %, 25.3 % and 46 % respectively of the CTC admissions. The relative contribution of the spatial component increases noticeably as the weekly number of observed CTC admissions increases (**Figure 6-17**).

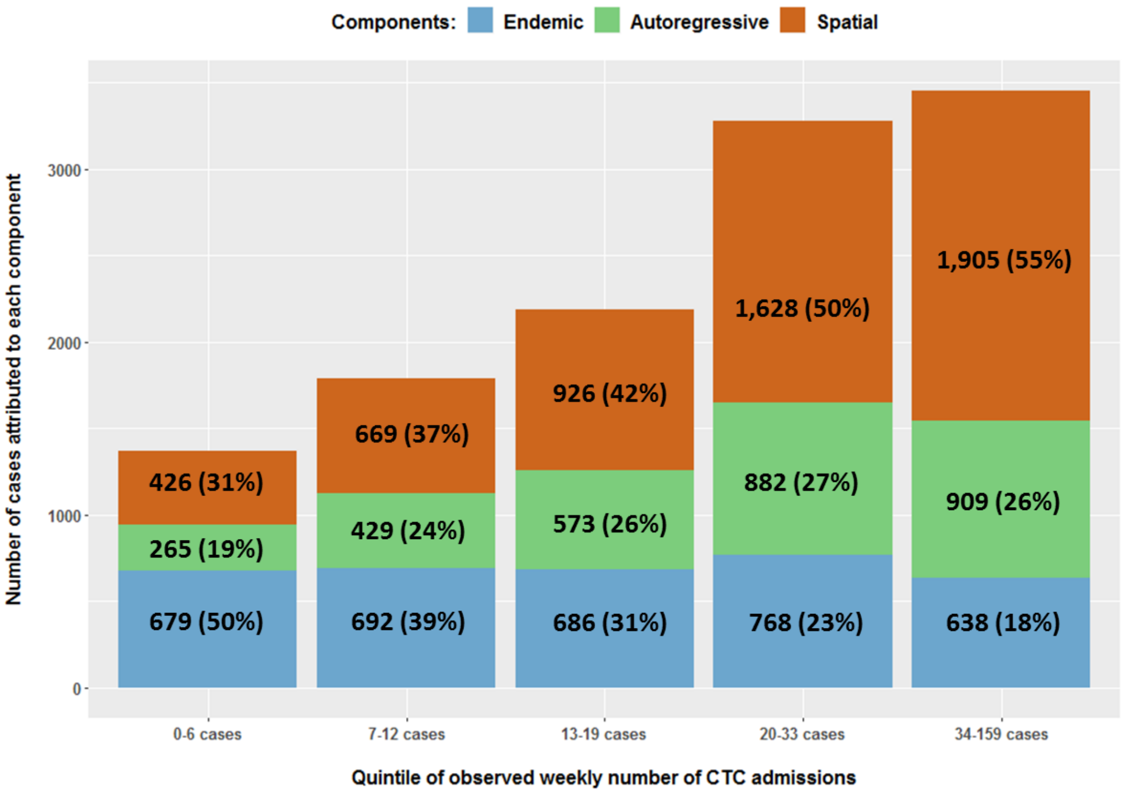


Figure 6-17 Contributions of Model 2 components by quintile of weekly observed number of CTC admissions

Model residuals

The total number of CTC admissions predicted by model 2 is 12,078, for 12,180 observed. Weekly absolute error between fitted and observed CTC admission counts ranges from -119 (week 37-2010) to +49 (week 37- 2009) and varies widely over quintiles of weekly observed CTC admissions (**Figure 6-18**). Model 2 tends to overestimate the number of CTC admissions for weeks with less than 20 admissions observed, while it tends to underestimate the number of admissions at weeks with admissions >33.

Absolute error at the neighbourhood level varies between -303 (Kakombe South) and +137 (Rombe1 centre), with a median of 0 (IQR = -39 - +44). Model 2 tends to overestimate the number of CTC patients residing in neighbourhoods with lower observed admissions, and to underestimate the number of CTC patients residing in neighbourhoods reporting more CTC cases.

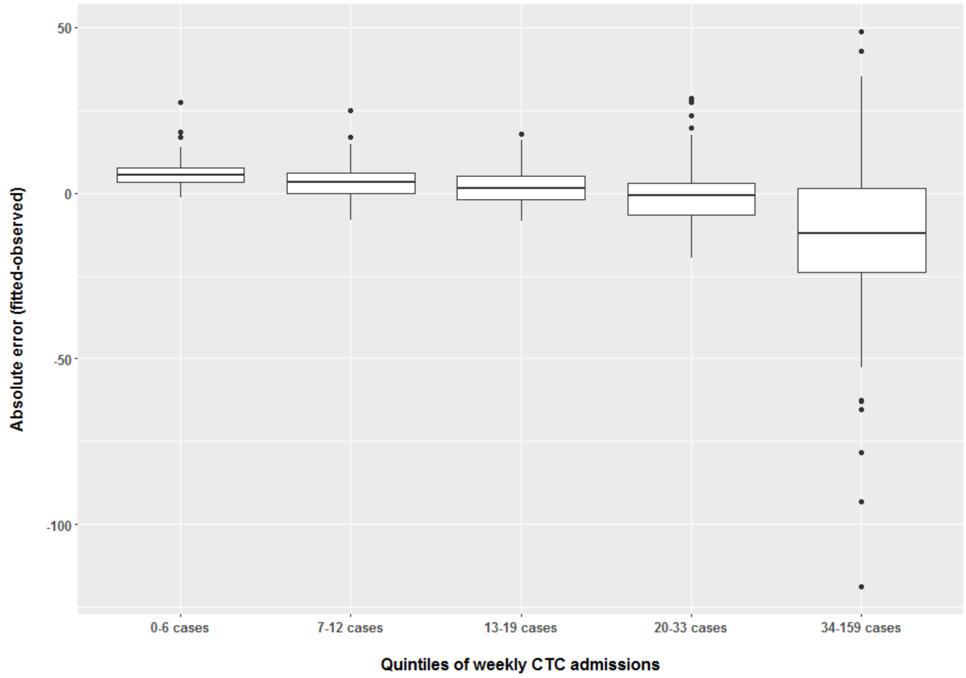


Figure 6-18 Model 2 absolute error distribution over quintiles of weekly incidence

Incidence attributable to insufficient tap water availability

The actual situation was compared with two hypothetical scenarii: 1) the volume of tap water distributed is stable at the 95th percentile of the distribution of the actual weekly amount distributed (30,000 m^3), no changes to the tap service indicator ; 2) all neighbourhoods are distributed 20 lcd or their actual tap water volume, whichever is the higher, during the entire study period.

Model 2 estimates that 1,217 admissions to the CTC (10.1 %) can be attributed to a weekly amount of distributed water lower than 30'000 m^3 (95% CI 639 – 1,711 admissions; 5.3 – 14.2 %). Under scenario 2, a total of 2,555 admissions to the CTC (21.1 %) would be averted (95% CI 413 – 4,266 admissions; 3.4 – 35.3 %).

Transmission hotspots

Taking into account each neighbourhood's adjacency orders with the other neighbourhoods and the normalised power law decay estimate d from model 2 ($\sum_{j \neq i} w_{ji}$ with $w_{ij} = o_{ij}^{-d} / \sum_k w_{jk}$, where $1 \leq o \leq 7$ in the spatial component formulation), the contribution of each neighbourhood to the incidence of others during the study period was mapped (**Figure 6-19**). It highlights the higher spatial influence of central neighbourhoods, as they have more neighbouring areas at higher adjacency order.

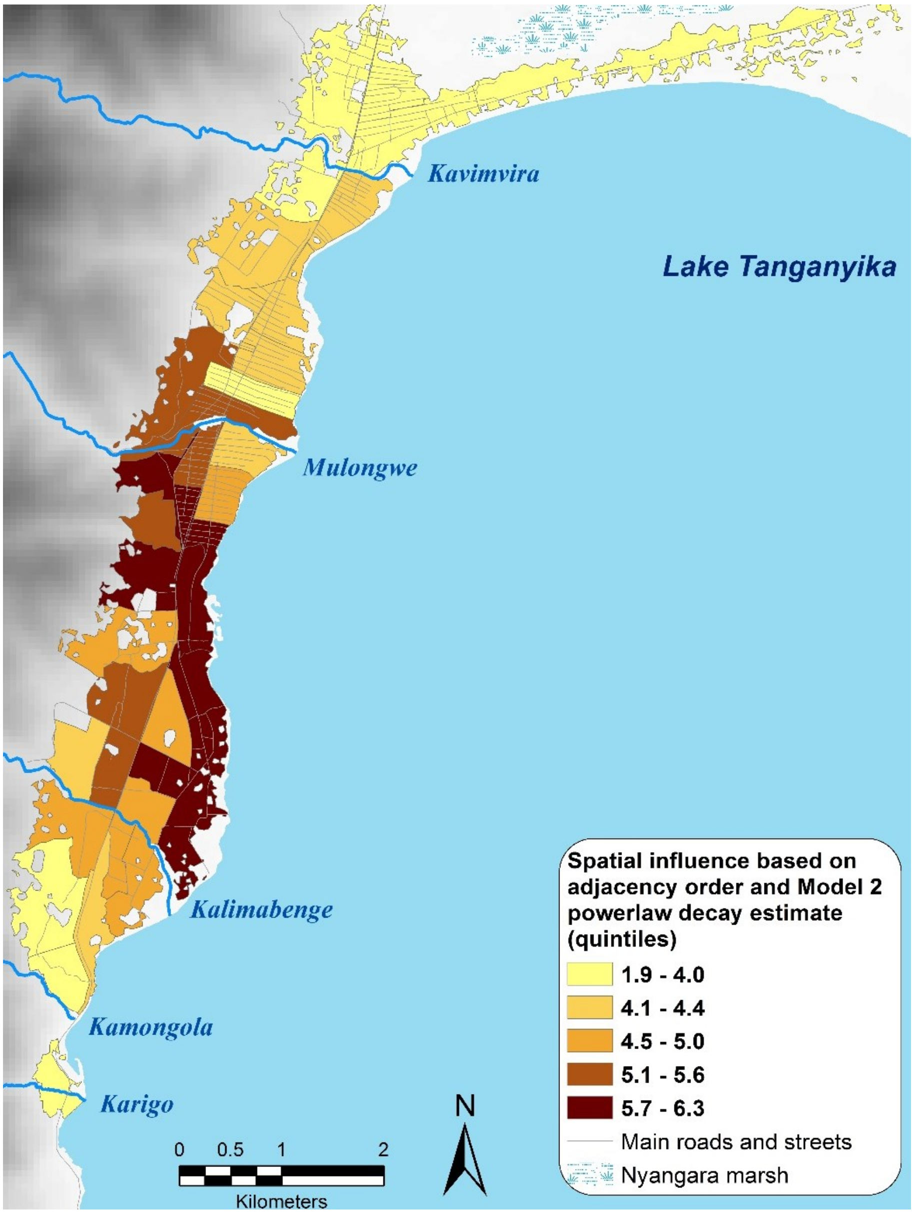


Figure 6-19 Spatial influence of neighbourhoods based on adjacency order and Model 2 normalised powerlaw decay estimate d

When the spatial transmission coefficient and CTC admissions rate are taken into account ($\phi_t \sum_{j \neq i} w_{ji}\, y_{i,t-1}$ in the spatial component formulation), some neighbourhoods can be identified as "hotspots" of transmission to other areas of town (**Figure 6-20**). Model 2 suggests that cases from Kakombe South and Kasenga South East strongly influenced the incidence in other areas of town over the study period, with 1,059 and 1,038 admissions respectively contributing to other neighbourhoods' admission rates.

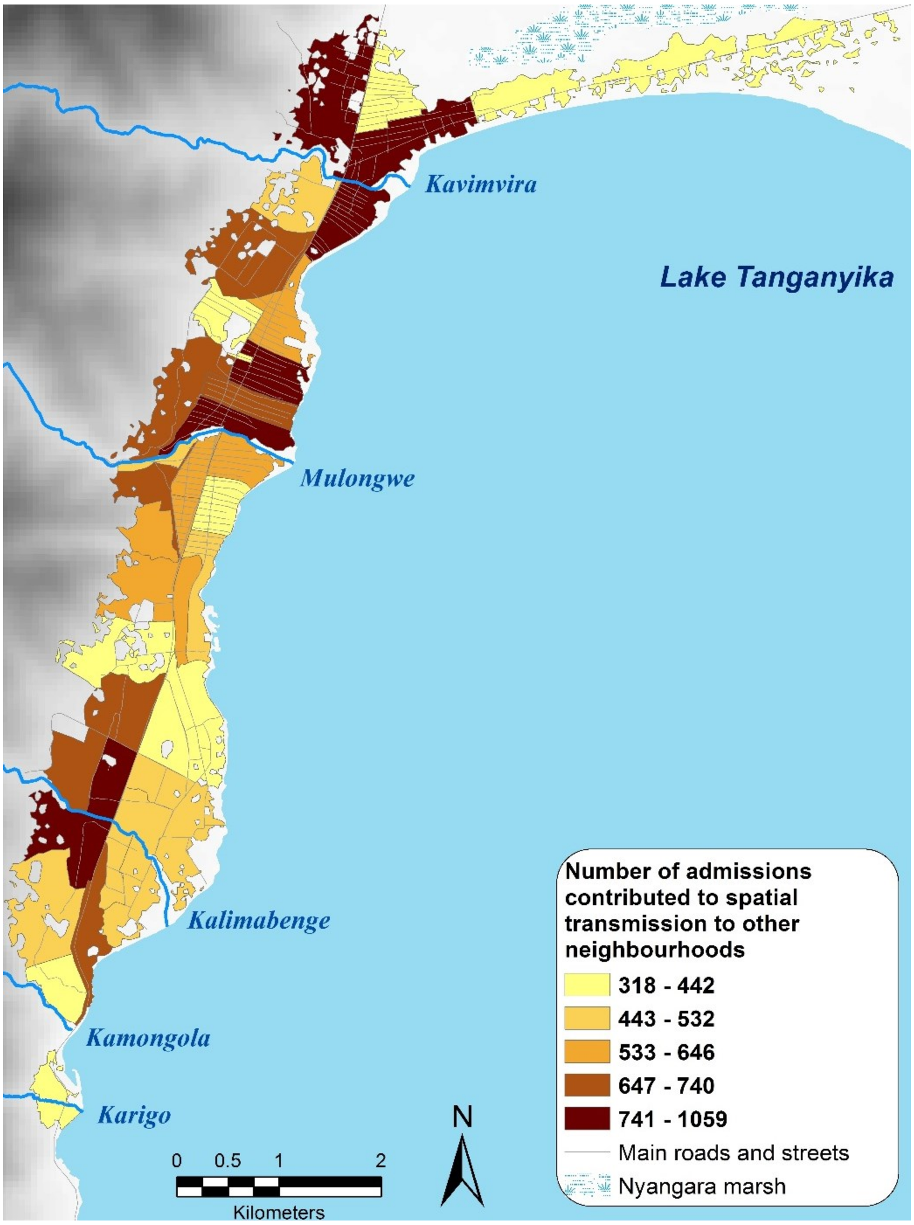


Figure 6-20 Map of "transmission hotspots": number of admissions contributed to spatial transmission to other neighbourhoods over the study period

Sensitivity analysis

Allowing for a unit-specific overdispersion parameter (model 3) or including more seasonal terms in all three components, in the best fitting combination suggested by Akaike Information Criterion (AIC) instead of BIC (model 4), led to only minor changes in parameters estimates or model performance. Similarly, using a tap water service indicator estimated over 500 m or 750 m radiuses instead of 250m (models 5 and 6) only leads to minor changes in magnitude of weekly tap water availability effect over the autoregressive component. Finally, smoothed population estimates (model 7) strengthen very slightly the results obtained by model 2, with an overall narrowing of the confidence intervals around parameter estimates. Results of the sensitivity analysis are reported in **Table 6-4**.

Table 6-4 Sensitivity analysis results

	Model 2 (main model)	Model 3	Model 4	Model 5	Model 6	Model 7
Varying model parameters						
Seasonal sinusoidals	1 in endemic component, 2 in spatial component		3 in endemic component, 2 in both autoregressive and spatial components			
Overdispersion	Single parameter for all neighbourhoods	Neighbourhood-specific parameter				
Tap water service indicator	250m radius			500m radius	750m radius	
Population estimates	Discrete at ABU level					Smoothed continuous
Endemic component IRRs [1]						
Time trend x mean tap water volume distributed	0.983 (0.972 - 0.994)	0.981 (0.97 - 0.993)	0.981 (0.97 - 0.992)	0.983 (0.97 - 0.997) [3]	0.984 (0.969 - 0.999) [3]	0.981 (0.969 - 0.992) [5]
Mean tap water volume distributed	1.04 (0.97 - 1.11)	1.03 (0.96 - 1.1)	1.04 (0.98 - 1.11)	1.04 (0.96 - 1.13) [4]	1.04 (0.95 - 1.14) [4]	1.01 (0.95 - 1.08) [6]
Autoregressive component IRR [1]						
Tap water volume distributed the previous week	0.78 (0.67 - 0.91)	0.78 (0.66 - 0.92)	0.78 (0.66 - 0.91)	0.73 (0.61 - 0.87)	0.69 (0.57 - 0.84)	0.77 (0.66 - 0.91)
Spatial component IRRs [1]						
Tap water volume distributed the previous week	0.74 (0.64 - 0.85)	0.73 (0.63 - 0.84)	0.72 (0.63 - 0.84)	0.71 (0.6 - 0.84)	0.71 (0.6 - 0.86)	0.71 (0.61 - 0.83)
Mean tap water volume distributed	1.27 (1.18 - 1.37)	1.26 (1.17 - 1.35)	1.3 (1.21 - 1.41)	1.37 (1.25 - 1.5) [4]	1.42 (1.28 - 1.57) [4]	1.31 (1.21 - 1.42) [6]
BIC	36,921	37,175	36,951	36,914	36,910	36,892
Attributable number [2]	2,555 (413 - 4,266)	2,752 (555 - 4,494)	2,594 (441 - 4,314)	2,675 (307 - 4,530)	2,776 (301 - 4,683)	2,717 (500 - 4,469)
Attributable fraction [2]	20.6% (3.3 - 34.3%)	22.2% (4.5 - 36.2%)	20.9% (3.5 - 34.7%)	21.5% (2.5 - 36.5%)	22.3% (2.4 - 37.7%)	21.9% (4 - 36%)

Table notes: 1) same increments used as table 6.2 unless specified; 2) Comparison reference: minimum of 20 lcd each week in all units; 3) 52 weeks x 4.7 lcd; 4) 50% or 4.7 lcd; 5) 52 weeks x 4.6 lcd; 6) 50% or 4.6 lcd

6.4 Discussion

In the present study, time and space patterns of more than 12,000 admissions to the CTC in Uvira over 10 years were explored using an endemic-epidemic modelling framework. Results provide insights into the relative contributions of short- and long-cycle transmissions of acute diarrhoeal diseases leading to CTC admission in a cholera-endemic urban setting and the influence of tap water provision on these cycles.

6.4.1 Results summary

In a model without covariates, the endemic incidence of CTC admissions at the beginning of the study was estimated to approximately 0.4 admissions per week per 10,000 overall and halved over the study period to 0.20 admissions per week per 10,000. Each admission from a specific neighbourhood led to 0.28 admissions in the same neighbourhood and 0.26 admission from directly adjacent neighbourhoods the following week. It also led to 0.13 admissions from more distant neighbourhoods. Endemic incidence varied annually, with an increase of about 25% in December and a decrease of about 20% in June. Admissions generated from admissions in other neighbourhoods also showed seasonal patterns, with a notable 75 % increase in September and decreases in May and November by 20 % and 40 % respectively.

When augmented with covariates and potential confounders, long-term tap water service, here represented by the mean of tap water service indicator in each neighbourhood over the whole study period, was shown to be independently associated with the spatial component and to interact with the endemic component decrease over the study period. Short-term, weekly, variations in tap water availability, that reflect both spatial heterogeneity in tap water service and variability in tap water volume distributed over time, were associated with both epidemic components, spatial and autoregressive.

Neighbourhoods with a better long-term tap water service experienced a sharper decrease in endemic incidence over the study period. Their endemic incidence level at the beginning of the study was however not significantly different from other neighbourhoods once wealth and distance to alternative water sources were taken into account. A better long-term tap water service however increased noticeably the incidence attributed to admissions in neighbouring areas, independently of wealth or distance to alternative sources.

In all neighbourhoods, an increase in the amount of tap water available in a specific week decreased the number of admissions generated the following week from admissions in the same or neighbouring areas. This effect was only slightly confounded by neighbourhoods' wealth and location relative to alternative sources of water.

Our results highlight the relative importance and distinctive nature of epidemic components – short-cycle transmission - in comparison with the endemic one – long-cycle transmission. Overall, the number of CTC admissions originating from the previous week's admissions represent more than two thirds of all admissions, and the number of cases unrelated to previous week's admissions, represent less than a third of all admissions.

Assuming that the above described relationships between average tap water service and weekly amount of tap water available in neighbourhoods with long and short-cycle transmission represent a causal association, we established that a stable weekly volume of tap water distributed at the 95th percentile of its observed maximum – ie with no increase in production capacity but an improvement of its reliability – may reduce the number of CTC admissions by nearly 10 %. Improvements in production capacity and reliability combined with an improved service in the poorly served areas in order to reach a minimum service indicator of 20 lcd every week, may reduce the number of CTC admissions by more than 20 %.

6.4.2 Interpretation

Our results attest to a distinct impact of tap water service level on short- and long-cycle transmission of acute diarroeal diseases leading to CTC admissions in Uvira.

Formulations of model components define the time and space boundaries of short- and long-cycle transmissions underlying endemic and epidemic components. The use of a single lag for autoregressive and spatial components formulation indeed limits the occurrence of short-cycle transmission to within one week of a case admission to the CTC. The spatial component also restricts short-cycle transmission to the boundaries of Uvira, as the spatial lag includes CTC admissions from all 37 neighbourhoods but not from outside Uvira. The endemic component is therefore referring to any transmission occurring outside of these time and space boundaries: transmission may be related to cases occurring outside of Uvira or to an indirect pathway with more than a week between index and secondary cases, possibly involving an environmental "reservoir" or an asymptomatic case (or one not seeking care at the CTC).

The association of a better tap water service on average with a lower endemic incidence is confounded by both wealth and location relative to surface water sources, suggesting that endemic incidence and long-cycle transmission of acute diarrhoeal diseases are not exclusively related to contaminated water consumption **[51]**, but are most likely also linked to other exposures to a contaminated environment, known to be associated with wealth as well. Higher wealth index is indeed related to better sanitation and some hygiene practices, in Uvira (see *Chapter 4 Tap water service and households' water-related practices*) and elsewhere **[52,53]**.

The sharp decline in endemic incidence over the study period, sharper in neighbourhoods with a better tap service, suggests that environment-mediated transmission of diarrhoeal diseases decreased over the study period. There is no evidence of such a trend in the autoregressive and spatial components, implying that this decrease in transmission was indeed limited to long-cycle transmission. A possible explanation for this decline in endemic incidence could be a decreasing consumption of contaminated water from surface sources as a result of several communication campaigns on drinking water quality, perhaps combined with a reduction of exposure to environmental contamination over the study period, especially in households residing in areas where a better tap water service enabled it further.

The influence of the weekly amount of tap water availability on the autoregressive component, independently of neighbourhood wealth and distance to other water sources, indicates that the actual amount of tap water available in a specific week, rather than in average over time, impacts the short-cycle transmission of cholera and acute diarrhoeal diseases leading to CTC admissions within a neighbourhood. It shows that the lack of water availability at the taps and the coping mechanisms it leads to, influences the water-washed transmission routes and hand / food / domestic hygiene at the household and communal levels. This influence may be related to a higher risk of drinking water in the household being contaminated after collection by an infected individual as the household reverts to a source with insufficient residual chlorine. It may also reflect households reducing temporarily the amount of water used for hygiene activities. This second interpretation is consistent with the increase in autoregressive case generation as distance to the nearest river increases, that would indicate that in neighbourhoods further away from rivers, the lack of tap water is harder to compensate in quantity terms, by using river water. A third explanation for an increase in short-cycle transmission resides in pathogen ingress into the local tap water distribution network under intermittent pressure.

The spatial component represents an extension of the autoregressive component over longer distances, and short-cycle transmission at a less localised scale, outside of the household and local community. Although counter-intuitive at first, the increasing number of admissions attributed to previous cases in other neighbourhoods in areas better served on average with tap water may be explained by the attraction of this higher level of tap water service towards residents of other neighbourhoods. Inflow of residents from other neighbourhoods may play a role in the short-cycle transmission from infected residents of poorly served neighbourhoods. Evidence of negative confounding by neighbourhood wealth suggest that wealth and associated factors in a neighbourhood may reinforce the attraction of better served neighbourhoods to inhabitants of other areas. In Uvira, higher levels of tap water service and wealth can be found in the historical centre of town where administration offices and commercial activities are concentrated, which attract many residents of other neighbourhoods for reasons other than tap water itself. Spatial transmission could also be related to contaminated wastewater and sewage flowing between neighbourhoods. In Uvira in particular, where sanitation infrastructure is limited to onsite technology and open defecation is still frequently reported, all waste- and rainwater flows downwards to the rivers and lake, through neighbourhoods at a lower elevation. Elevation is also a hindrance for tap water distribution and higher altitude neighbourhoods are generally less well served by tap water. Although elevation is not an explicit confounder in our model, it is highly correlated with distance to the lake and somewhat with distance to the rivers, which are accounted for. As distances to the lake and rivers, and therefore elevation, increase, our model suggests that spatial transmission decreases, consistent with this interpretation.

Increased tap water availability in a given week is however associated with reduced spatial transmission from other neighbourhoods, perhaps by allowing better hygiene practices in households and local public spaces, as for the shorter distance autoregressive transmission. Contamination of the water supply distribution network with pathogens shed by infected individuals, may also be responsible for longer distance transmission, depending on the network design.

Figure 6-21 An open drainage channel in the Rombe 2 neighbourhood

6.4.3 Limitations

There are several limitations to this analysis, and it is challenging to predict when and how potential biases may have affected our results.

We first base our modelling on the number of reported CTC admissions, and healthcare seeking behaviour may be an important source of bias. Although increasing distance of neighbourhoods from the CTC is not obviously associated with a lower admission rate in our data, we cannot rule out that having to travel further to receive healthcare remains a deterrent for households living in further afar neighbourhoods. Household wealth may induce a similar bias, as the poorest may not seek healthcare because of its direct and opportunity costs, whilst the wealthiest may seek care from private higher-end healthcare facilities, even though this is strongly discouraged by the health authorities and healthcare providers themselves because of the required infection prevention and control (IPC) measures. Biases may also change over time, with healthcare-seeking behaviour possibly changing when treatment at the CTC is perceived as better or worse – for example when an international non-governmental organisation (INGO) is deployed to support the CTC during a surge in cases – or when communities are made aware of an ongoing outbreak **[54]**.

The potential for differential healthcare seeking behaviour in the population raises questions about the interpretation of the number of CTC admissions as a proxy for disease transmission intensity. Even though one may suppose that healthcare seeking behaviour is less likely to vary as the severity of symptoms increases, and that symptom severity increases with exposure to a higher dose of pathogen, the assumed relationship between exposure and symptom severity, hence between exposure and CTC admissions, is likely to suffer from fluctuating upwards or downwards biases.

In addition, our results are likely affected by the aetiology of the diarrhoea leading to CTC admissions. In particular, the variations in average serial intervals – the time interval between onset of symptoms in primary and secondary cases – between pathogens may introduce misclassification between endemic and epidemic components in a model based on a discrete time unit of 7 days[6]. Results from patients' stool analysis presented in chapter *Cholera and other enteric infections amongst patients admitted to the cholera treatment centre* showed that a majority of CTC patients tested positive for either *V. cholerae* or ETEC, for which average serial intervals are estimated to be shorter than 7 days (about 4 days for cholera **[55]**, and possibly slightly shorter for ETEC), suggesting that our estimate of the ratio of epidemic:endemic components in our model may be an underestimate. Using a weekly count of CTC admissions has the advantage however of removing potential variations in healthcare seeking behaviour and CTC admissions process across days of the week.

Both pathogens share many biological and epidemiological characteristics, such as infectious dose, incubation period, infectious duration and shedding, high proportion of asymptomatic infections, clinical presentation features, individual risk factors and

[6] For pathogens with an average serial interval shorter than 7 days, one case can lead to another case within the same time interval used in the model, and the second case may be included in the endemic component estimate rather than epidemic components. The inverse applies to pathogens with an average serial interval longer than 7 days.

seasonal incidence peaks. They are likely transmitted along similar short- and long-cycle transmission pathways but the reason why ETEC has not been implicated in explosive large-scale outbreaks has yet to be elucidated. Our results may therefore largely represent a mix of the patterns of transmission of both *V. cholerae* and ETEC, though it cannot be excluded that some of these patterns apply specifically to one or the other, and that epidemic component estimates are overestimated if both pathogens are responsible separately for a peak in CTC admissions at the same time. Multiple aetiologies also make it difficult to grasp the role of immunity acquired from infection with a specific pathogen, and our model does not account for expected variations in susceptible population over time. Duration and levels of protection conferred by a natural infection are unknown for many enteropathogens **[56]**. A natural *V. cholerae* infection is however expected to provide substantial protection against subsequent symptomatic cholera infection of the same serogroups for 3 to 7 years **[57]**, although mild or asymptomatic infections may not confer the same level of protection **[58]**.

The use of weekly cumulated amounts of tap water distributed may also mask the impact of short interruptions or reductions in distribution, and explain the much lower fraction of CTC admissions attributed to weekly variations in tap water service than from daily variations (10 % vs 25 % - **[45]**). Also, the tap water service indicator calculation is based on the monthly proportion of total water distributed at each tap, masking even further how daily variations in water distribution may affect some parts of the town more than others. Further, the indicator does not reflect the potential impact of distribution intermittency over tap water quality, likely to vary geographically with the distribution network, its pressure profile and pipe damages.

The present results are also limited by the use of several variables which are estimated based on a single time-point measure, such as wealth, built-up surfaces and tap water service level, de facto assuming that neighbourhoods' relative characteristics remained constant over the study period. This assumption may be particularly invalid for neighbourhoods at the periphery of town, that developed more recently, or areas suffering from a temporary absence of tap water distribution following major pipe damage. Like other possible sources of bias for this study, it is difficult to predict when this bias may have affected the analysis, and how it is likely to have influenced the estimated model parameters.

Finally, this model uses aggregated measures of exposure, confounders and disease outcome for each discrete spatial unit, and is therefore vulnerable to the ecological fallacy when inferring that model results pertain to individual levels of exposure and risk **[59-61]**. Although our findings do not contradict individual risk factors for acute diarrhoea and cholera identified through other studies, further analyses could be performed to exclude the ecological fallacy as an explanation, by addressing in particular the variability of exposure (here tap water service) and confounders (here wealth and distances to surface water sources) within spatial units **[62]**. Running the same model with different geographical aggregations of ABUs, with the same or a varying number of geographical units, would be a first step to investigate whether ecological bias is indeed present, and whether more complex modelling approaches are necessary and sufficient to reduce it **[63,64]**.

6.4.4 Implications of the results

Despite the above limitations, our results support an important role for tap water service on endemic and epidemic transmission of diarrhoeal diseases in the cholera-endemic urban setting of Uvira. Although further research is needed to better elucidate the impact of tap water service on specific transmission pathways, our findings hint at several opportunities to improve cholera and diarrhoeal diseases control strategies for Uvira.

Time-sensitive response to prevent short-cycle transmission

Our findings highlight the importance of short-cycle transmission within and across neighbourhoods, supporting intervention strategies that prevent short-cycle transmission of cholera and other diarrhoeal diseases in affected households and their neighbourhood, but also in the wider community.

Case area targeted interventions (CATI) strategies have increasingly been recommended and rolled out during cholera outbreak responses in recent years **[1,21]** with the clear objective of reducing transmission within affected households and their surroundings, for which a much higher risk of infection and disease was observed **[5,30,31,65-67]**. CATI most often integrate some or all of the following prevention components: promotion of point-of-use water treatment, improved hand and food hygiene practices through intense messaging, home visits and provision of supplies – soap, hand washing station, water treatment product - and actual interventions to enhance domestic hygiene – disinfectant spraying for example. Antibiotic prophylactic treatment and oral cholera vaccination may also be part of the package. Part of a CATI strategy could also be delivered directly to CTC patients or their carers **[68]**.

Although our findings and those of the household surveys (see chapter *Tap water service and households' water-related practices*) do not question the relevance of these interventions for a CATI response in Uvira, they support the development of an additional intervention, to address the water quantity issue, as was implemented in e recent outbreak response in Kinshasa **[69]**, that would ensure that occasional reduced access to water, does not act as a barrier to the enhanced water quality and hygiene practices promoted. However, results of the households' surveys conducted in 2016 and 2017 are not without limitations (see end of section *4.4*) and such intervention development would require a better understanding of 1) current practices in affected households and how limited water quantity use occasionally affects their compliance with targeted behaviours; 2) how best to bridge the gap in short-term clean water access, for instance by temporarily bringing clean water closer to the targeted households through additional treated water transport and distribution points, or increasing households' means for safe storage of water.

Compensating for the degradation of water quality at the taps due to an intermittent supply, especially in the vicinity or downstream of affected households, may also be part of a targeted response. Additional chlorination at the tap as water distribution resumes may reduce the faecal contamination resulting from pathogen ingress during periods of low pressure and biofilm detachment as pressure and velocity peaks in the distribution network **[70-73]**.

The relative importance of inter-neighbourhood short-cycle transmission in our model indicates that consideration should be given to expanding the response to an increase in CTC admissions from a specific neighbourhood to a wider geographical area. Mapping the recurrent and seasonal patterns of movement of the population between neighbourhoods, in relation to tap water interruptions and tap water service geographic inequalities, but also relating to other public gathering places and events, would be key for an appropriate targeting of such a wider response, as would be an investigation of the precise transmission routes that can be prevented in the short-term. In addition to population mobility, documenting the course of wastewater, in existing open drainage channels or free flowing, including seasonal variations in such paths, would complement the wider community targeting for short-cycle transmission prevention.

The seasonal patterns identified in both long- and short-cycle transmission should be explored, as they may contribute to further elucidation of transmission pathways and the most effective means to prevent them. Seasonal patterns may also provide help to predict surges in CTC admissions and better plan and prepare for timely and targeted responses. Climate-related variability in enteric diseases and cholera incidence continues to be researched globally, especially as remote-sensing data are increasingly available along with methods for analysis **[74]**. The relationship between planktonic bloom and cholera outbreaks in Bangladesh is a well-known example **[75]**. Seasonal patterns in CTC admissions may however be related to other variations in human behaviours that should also be considered; for instance, social mixing patterns or diet may change with religious celebrations, school year, harvesting and fishing seasons **[28,76-78]**.

Figure 6-22 Planktonic bloom in Lake Tanganyika in September 2018 (photo courtesy VL Kapepula)

Sustainable improvements of tap water service

The need to sustainably improve tap water service has long been promoted to prevent diarrhoeal diseases and is an integral part of the Global Task Force for Cholera Control (GTFCC) roadmap to cholera elimination by 2030 **[79,80]**. The present modelling results as well as the documented households' water-related practices, substantiate the need to target the "safely managed" level of water access as defined by the Sustainable Development Goals (SDG) and target 6.1 on drinking water), with a reliable tap water supply free of contamination on premises **[81]**.

Although improving average tap water service alone in some neighbourhoods may not be sufficient to reduce long-cycle transmission of cholera and acute diarrhoeal diseases, it is necessary in order to enable key practices and behaviours that limit exposure to enteric pathogens in the environment, as it seems to have in better served areas over the study period. Priority should be given to reducing service inequalities by targeting the poorest areas of Uvira, where tap water service is also generally the lowest and is present alongside other risk factors (see *Chapter 4 Tap water service and households' water-related practices*). Special attention should be paid as well to areas nearest surface water sources where determinants of households' water source choices may be intricate **[51]**. Reducing geographical disparity in average tap water service may also reduce population mobility across neighbourhoods for tap water collection and therefore decrease some of the short-cycle transmission between neighbourhoods.

Evidently, improvements of tap water service in Uvira should not be limited to increasing coverage in tap connections but should also ensure that tap water is indeed available at the taps when needed. Infrastructure investments into water treatment, storage capacity and the distribution network to meet higher coverage and demand should be carefully planned to maximise supply reliability. If intermittent supply is temporarily unavoidable, due to engineering, power supply and funding constraints, further research should be done on households' coping strategies with tap water interruptions and on potential benefits of predictable interruptions in service over unexpected ones– for instance when a scheduled, rotational supply by neighbourhoods is in place or phone messages ahead of planned interruptions in service are sent to customers. Ideally, the infrastructure and its management should allow as well for short-term and targeted supply increase to complement other short-cycle transmission prevention interventions in response to a surge in CTC admissions or before such a surge is predicted to occur (see above in *Time-sensitive response to prevent short-cycle transmission*).

More data on the impact of supply intermittency on the quality of the water distributed should be collected and lead to the implementation of mitigation strategies to maintain water quality when and where necessary. Additional treatment points on the distribution network may be needed to guarantee a sufficient concentration of residual chlorine at all taps, especially as many customers may still be using connections outside their compound.

6.5 Chapter 6 - References

1. D'Mello-Guyett L, Gallandat K, Van den Bergh R, et al. *Prevention and control of cholera with household and community water, sanitation and hygiene (WASH) interventions: A scoping review of current international guidelines*. **PLoS One**. 2020;15(1):e0226549.
2. Phelps MD, Azman AS, Lewnard JA, et al. *The importance of thinking beyond the water-supply in cholera epidemics: A historical urban case-study.* **PLoS Negl Trop Dis**. 2017;11(11):e0006103.
3. Ramamurthy T, Mutreja A, Weill FX, Das B, Ghosh A, Nair GB. *Revisiting the Global Epidemiology of Cholera in Conjuction With the Genomics of Vibrio cholerae*. **Frontiers in public health**. 2019;7:203.
4. Weill FX, Domman D, Njamkepo E, et al. *Genomic history of the seventh pandemic of cholera in Africa.* **Science**. 2017;358(6364):785-789.
5. Domman D, Chowdhury F, Khan AI, et al. *Defining endemic cholera at three levels of spatiotemporal resolution within Bangladesh.* **Nature genetics**. 2018;50(7):951-955.

6. Weill FX, Domman D, Njamkepo E, et al. *Genomic insights into the 2016-2017 cholera epidemic in Yemen*. **Nature**. 2019;565(7738):230-233.

7. Mavian C, Paisie TK, Alam MT, et al. *Toxigenic Vibrio cholerae evolution and establishment of reservoirs in aquatic ecosystems*. **Proceedings of the National Academy of Sciences of the United States of America**. 2020;117(14):7897-7904.

8. Rebaudet S, Sudre B, Faucher B, Piarroux R. *Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes*. **J Infect Dis**. 2013;208 Suppl 1:S46-54.

9. Bompangue D, Giraudoux P, Handschumacher P, et al. *Lakes as source of cholera outbreaks, Democratic Republic of Congo*. **Emerg Infect Dis**. 2008;14(5):798-800.

10. Bompangue Nkoko D, Giraudoux P, Plisnier PD, et al. *Dynamics of cholera outbreaks in Great Lakes region of Africa, 1978-2008*. **Emerg Infect Dis**. 2011;17(11):2026-2034.

11. Berhe HW, Makinde OD, Theuri DM. *Modelling the dynamics of direct and pathogens-induced dysentery diarrhoea epidemic with controls*. **Journal of biological dynamics**. 2019;13(1):192-217.

12. Beyene H, Deressa W, Kumie A, Grace D. *Spatial, temporal, and spatiotemporal analysis of under-five diarrhea in Southern Ethiopia*. **Tropical medicine and health**. 2018;46:18-18.

13. King AA, Ionides EL, Pascual M, Bouma MJ. *Inapparent infections and cholera dynamics*. **Nature**. 2008;454(7206):877-880.

14. Lal A, Fearnley E, Wilford E. *Local weather, flooding history and childhood diarrhoea caused by the parasite Cryptosporidium spp.: A systematic review and meta-analysis*. **Sci Total Environ**. 2019;674:300-306.

15. Mukandavire Z, Liao S, Wang J, Gaff H, Smith DL, Morris JG. *Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe*. **Proceedings of the National Academy of Sciences**. 2011;108(21):8767-8772.

16. Nilima, Kamath A, Shetty K, Unnikrishnan B, Kaushik S, Rai SN. *Prevalence, patterns, and predictors of diarrhea: a spatial-temporal comprehensive evaluation in India*. **BMC public health**. 2018;18(1):1288.

17. Tien JH, Earn DJ. *Multiple transmission pathways and disease dynamics in a waterborne pathogen model*. **Bulletin of mathematical biology**. 2010;72(6):1506-1533.

18. Chao DL, Longini IM, Jr., Morris JG, Jr. *Modeling cholera outbreaks*. **Current topics in microbiology and immunology**. 2014;379:195-209.

19. Che EN, Kang Y, Yakubu AA. *Risk structured model of cholera infections in Cameroon*. **Mathematical biosciences**. 2020;320:108303.

20. Dangbe E, Irepran D, Perasso A, Bekolle D. *Mathematical modelling and numerical simulations of the influence of hygiene and seasons on the spread of cholera*. **Mathematical biosciences**. 2018;296:60-70.

21. Finger F, Bertuzzo E, Luquero FJ, et al. *The potential impact of case-area targeted interventions in response to cholera outbreaks: A modeling study*. **PLoS Med**. 2018;15(2):e1002509.

22. Lin J, Xu R, Tian X. *Global dynamics of an age-structured cholera model with multiple transmissions, saturation incidence and imperfect vaccination*. **Journal of biological dynamics**. 2019;13(1):69-102.

23. Lin JZ, Xu R, Tian XH. *Transmission dynamics of cholera with hyperinfectious and hypoinfectious vibrios: mathematical modelling and control strategies*. **Mathematical biosciences and engineering : MBE**. 2019;16(5):4339-4358.

24. Nyabadza F, Aduamah JM, Mushanyu J. *Modelling cholera transmission dynamics in the presence of limited resources*. **BMC research notes**. 2019;12(1):475.

25. Pasetto D, Finger F, Camacho A, et al. *Near real-time forecasting for cholera decision making in Haiti after Hurricane Matthew*. **PLoS computational biology**. 2018;14(5):e1006127-e1006127.

26. Yang C, Wang X, Gao D, Wang J. *Impact of Awareness Programs on Cholera Dynamics: Two Modeling Approaches*. **Bulletin of mathematical biology**. 2017;79(9):2109-2131.

27. Codeço CT. *Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir*. **BMC Infectious Diseases**. 2001;1(1):1.

28. Mari L, Bertuzzo E, Righetto L, et al. *Modelling cholera epidemics: the role of waterways, human mobility and sanitation*. **Journal of The Royal Society Interface**. 2012;9(67):376-388.

29. Wang J, Liao S. *A generalized cholera model and epidemic–endemic analysis*. **Journal of biological dynamics**. 2012;6(2):568-589.

30. Azman AS, Luquero FJ, Rodrigues A, et al. *Urban cholera transmission hotspots and their implications for reactive vaccination: evidence from Bissau city, Guinea bissau*. **PLoS Negl Trop Dis**. 2012;6(11):e1901.

31. Debes AK, Ali M, Azman AS, Yunus M, Sack DA. *Cholera cases cluster in time and space in Matlab, Bangladesh: implications for targeted preventive interventions*. **International journal of epidemiology**. 2016;45(6):2134-2139.

32. Lee EC, Azman AS, Kaminsky J, Moore SM, McKay HS, Lessler J. *The projected impact of geographic targeting of oral cholera vaccination in sub-Saharan Africa: A modeling study*. **PLoS medicine**. 2019;16(12):e1003003-e1003003.

33. Held L, Höhle M, Hofmann M. *A statistical framework for the analysis of multivariate infectious disease surveillance counts*. **Statistical Modelling**. 2005;5(3):187-199.

34. Bracher J, Held L. *Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction*. **arXiv preprint arXiv:190103090**. 2019.

35. Wakefield J, Dong TQ, Minin VN. *Spatio-temporal analysis of surveillance data*. **Handbook of Infectious Disease Data Analysis**. 2019:455-476.

36. Bauer C, Wakefield J. *Stratified space–time infectious disease modelling, with an application to hand, foot and mouth disease in China*. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**. 2018;67(5):1379-1398.

37. Chiavenna C, Presanis AM, Charlett A, et al. *Estimating age-stratified influenza-associated invasive pneumococcal disease in England: A time-series model based on population surveillance data*. **PLOS Medicine**. 2019;16(6):e1002829.

38. Herzog SA, Paul M, Held L. *Heterogeneity in vaccination coverage explains the size and occurrence of measles epidemics in German surveillance data*. **Epidemiology and infection**. 2011;139(4):505-515.

39. Stojanović O, Leugering J, Pipa G, Ghozzi S, Ullrich A. *A Bayesian Monte Carlo approach for predicting the spread of infectious diseases*. **PloS one**. 2019;14(12):e0225838-e0225838.

40. Plouffe CC. *Space-time modelling of emerging infectious diseases: Assessing leptospirosis risk in Sri Lanka*. 2016.

41. Adegboye OA, Adegboye M. *Spatially Correlated Time Series and Ecological Niche Analysis of Cutaneous Leishmaniasis in Afghanistan*. **Int J Environ Res Public Health**. 2017;14(3).

42. Adegboye OA, Al-Saghir M, Leung DH. *Joint spatial time-series epidemiological analysis of malaria and cutaneous leishmaniasis infection.* **Epidemiology and infection**. 2017;145(4):685-700.

43. Diggle PJ. *Statistical analysis of spatial point patterns.* London: Academic Press; 1983.

44. Tobler W. Three Presentations on Geographical Analysis and Modeling: Non-Isotropic Geographic Modeling; Speculations on the Geometry of Geography; and Global Spatial Analysis (93-1). 1993; https://escholarship.org/uc/item/05r820mz. Accessed 19/05/2019.

45. Jeandron A, Saidi JM, Kapama A, et al. *Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo.* **PLoS Med**. 2015;12(10):e1001893.

46. Pregibon D. *Goodness of Link Tests for Generalized Linear Models*. **Journal of the Royal Statistical Society Series C (Applied Statistics)**. 1980;29(1):15-14.

47. Meyer S, Held L, Höhle M. *Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package surveillance.* **Journal of Statistical Software; Vol 1, Issue 11 (2017)**. 2017.

48. Paul M, Held L, Toschke AM. *Multivariate modelling of infectious disease surveillance data.* **Statistics in medicine**. 2008;27(29):6250-6267.

49. Azman AS, Rudolph KE, Cummings DA, Lessler J. *The incubation period of cholera: a systematic review.* **The Journal of infection**. 2013;66(5):432-438.

50. Clemens JD, Nair GB, Ahmed T, Qadri F, Holmgren J. *Cholera.* **Lancet**. 2017;390(10101):1539-1549.

51. Jeandron A, Cumming O, Kapepula L, Cousens S. *Predicting quality and quantity of water used by urban households based on tap water service.* **npj Clean Water**. 2019;2(1):23.

52. Rheingans R, Anderson JD, Luyendijk R, Cumming O. *Measuring disparities in sanitation access: does the measure matter?* **Tropical Medicine & International Health**. 2014;19(1):2-13.

53. Waddington H, Snilstveit B. *Effectiveness and sustainability of water, sanitation, and hygiene interventions in combating diarrhoea.* **Journal of Development Effectiveness**. 2009;1(3):295-335.

54. Yakum MN, Ateudjieu J, Guenou E, et al. *Health seeking behaviour among suspected cases of cholera in Cameroonian health districts in Lake Chad basin.* **BMC research notes**. 2017;10(1):433.

55. Phelps M, Perner ML, Pitzer VE, Andreasen V, Jensen PKM, Simonsen L. *Cholera Epidemics of the Past Offer New Insights Into an Old Enemy*. **The Journal of infectious diseases**. 2018;217(4):641-649.

56. Rogawski McQuade ET, Liu J, Kang G, et al. *Protection From Natural Immunity Against Enteric Infections and Etiology-Specific Diarrhea in a Longitudinal Birth Cohort.* **The Journal of infectious diseases**. 2020.

57. Leung DT, Uddin T, Xu P, et al. *Immune responses to the O-specific polysaccharide antigen in children who received a killed oral cholera vaccine compared to responses following natural cholera infection in Bangladesh.* **Clin Vaccine Immunol**. 2013;20(6):780-788.

58. Ali M, Emch M, Park JK, Yunus M, Clemens J. *Natural cholera infection-derived immunity in an endemic setting.* **J Infect Dis**. 2011;204(6):912-918.

59. Idrovo AJ. *Three criteria for ecological fallacy*. **Environ Health Perspect**. 2011;119(8):A332-A332.

60. Piantadosi S, Byar DP, Green SB. *THE ECOLOGICAL FALLACY.* **American Journal of Epidemiology**. 1988;127(5):893-904.

61. Sedgwick P. *Understanding the ecological fallacy*. **BMJ : British Medical Journal**. 2015;351:h4773.

62. Salway R, Wakefield J. *A hybrid model for reducing ecological bias*. **Biostatistics**. 2008;9(1):1-17.

63. Lancaster GA, Green M, Lane S. *Reducing bias in ecological studies: an evaluation of different methodologies*. **Journal of the Royal Statistical Society: Series A (Statistics in Society)**. 2006;169(4):681-700.

64. Wang F, Wang J, Gelfand A, Li F. *Accommodating the ecological fallacy in disease mapping in the absence of individual exposures*. **Statistics in medicine**. 2017;36(30):4930-4942.

65. Azman AS, Luquero FJ, Salje H, et al. *Micro-Hotspots of Risk in Urban Cholera Epidemics*. **J Infect Dis**. 2018;218(7):1164-1168.

66. Bi Q, Azman AS, Satter SM, et al. *Micro-scale Spatial Clustering of Cholera Risk Factors in Urban Bangladesh*. **PLoS Negl Trop Dis**. 2016;10(2):e0004400.

67. Burrowes V, Perin J, Monira S, et al. *Risk Factors for Household Transmission of Vibrio cholerae in Dhaka, Bangladesh (CHoBI7 Trial)*. **Am J Trop Med Hyg**. 2017;96(6):1382-1387.

68. George CM, Sack DA. *Integration of water, sanitation and hygiene intervention delivery at health facilities with a reactive ring vaccination programme to reduce cholera*. **International journal of epidemiology**. 2017;46(6):2093-2094.

69. Bompangue D, Moore S, Taty N, et al. *Description of the targeted water supply and hygiene response strategy implemented during the cholera outbreak of 2017-2018 in Kinshasa, DRC*. **BMC Infect Dis**. 2020;20(1):226.

70. Bivins AW, Sumner T, Kumpel E, et al. *Estimating Infection Risks and the Global Burden of Diarrheal Disease Attributable to Intermittent Water Supply Using QMRA*. **Environmental Science & Technology**. 2017;51(13):7542-7551.

71. Coelho ST, James S, Sunna N, Abu Jaish A, Chatila J. *Controlling water quality in intermittent supply systems*. **Water Supply**. 2003;3(1-2):119-125.

72. Kumpel E, Nelson KL. *Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality*. **Environmental Science & Technology**. 2016;50(2):542-553.

73. Quaye MNA. *Effects of operational strategies on drinking water quality in intermittent water supply systems*. 2018.

74. Liang S. *Comprehensive Remote Sensing*. Elsevier Science; 2017.

75. Islam MS, Islam MS, Mahmud ZH, Cairncross S, Clemens JD, Collins AE. *Role of phytoplankton in maintaining endemicity and seasonality of cholera in Bangladesh*. **Transactions of the Royal Society of Tropical Medicine and Hygiene**. 2015;109(9):572-578.

76. Bengtsson L, Gaudart J, Lu X, et al. *Using Mobile Phone Data to Predict the Spatial Spread of Cholera*. **Scientific Reports**. 2015;5(1):8923.

77. Bwire G, Munier A, Ouedraogo I, et al. *Epidemiology of cholera outbreaks and socio-economic characteristics of the communities in the fishing villages of Uganda: 2011-2015*. **PLoS Negl Trop Dis**. 2017;11(3):e0005407.

78. Finger F, Genolet T, Mari L, et al. *Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks*. **Proceedings of the National Academy of Sciences**. 2016:201522305.

79. Legros D. *Global Cholera Epidemiology: Opportunities to Reduce the Burden of Cholera by 2030*. **J Infect Dis**. 2018;218(suppl_3):S137-s140.

80. Organization WH. *Ending cholera a global roadmap to 2030*. 2017.

81. United Nations Joint Monitoring Program. *JMP methodology 2017 update and SDG baseline*. Geneva 2018.

7

Summary, reflections and way forward

7.1 Main findings

In Uvira, a cholera-endemic town of the Democratic Republic of the Congo (DRC), admissions to the cholera treatment centre (CTC) reach annual rates comparable to those observed during some outbreaks in other countries of Sub-Saharan Africa, with an average of nearly 44 cases per year per 10,000 population. *V. cholerae* O_1 is the single most frequently detected diarrhoea-causing pathogen amongst admitted patients, but cholera was confirmed by rapid diagnostic tests (RDT) in only about 40 % of them over 37 months. Cholera confirmation rates increased with admission rates, but sporadic confirmed cases were also identified between the bi-annual admission peaks usually occurring in the first and third trimester of each year.

The tap water supply system of Uvira delivers an average of 17 litres per capita per day (lcd) through an average of about 14 active tap connections per 1,000 inhabitants. Nearly 80 % of households of Uvira report using the tap water supply system at least occasionally, especially for drinking water, although only 50 % of the population report using exclusively taps for their drinking water. About 63 % of the households use mostly tap water for domestic use, either from taps on the premises (25 %) or outside their compound (38 %). In absence of boreholes and other kinds of wells, other households habitually use the only alternative water sources, namely the five permanent rivers and Lake Tanganyika. Nearly 40 % of stored drinking water samples from households were free of *Escherichia coli*, but another 40 % were contaminated with more than 10 colony forming units (CFU) *E. coli* per 100 ml, which represents a substantial risk for diarrhoeal diseases transmission.

The tap water service is highly heterogeneous across neighbourhoods. A tap water service indicator, constructed to reflect both the density of tap connections within 250 m and the quantity of water drawn at the taps, ranged from 0 to 99 litres per capita per day across town in November 2017. The large variations in tap water service are reflected in households' water-related practices, independently of wealth: those living in the two best served quintiles are more likely to always use tap water for drinking, spend less time collecting water and use more water at home for domestic activities, particularly adult bathing, house cleaning and dishwashing. They are also less likely to store drinking water contaminated with *E. coli*. When combined with distance to rivers and lake, the tap water service indicator can be used to predict the probability of a household to store contaminated drinking water and the amount of water used at home.

Diminutions and interruptions of the daily supply of tap water are associated with an increase in CTC admissions – a day without tap water distribution is associated with a 2.5-fold increase in CTC admissions rate over the following 12 days. Over a little more than 5 years, 23.2 % of CTC admissions were attributed to a daily tap water supply lower than a reference volume set to 4,800 m^3 per day (about 19 lcd). The impact of reduced daily water supply is higher for areas of town with a better tap water service.

A spatio-temporal analysis performed on weekly CTC admissions for 37 neighbourhoods of Uvira over 10 years confirms that reductions in the weekly amount of tap water distributed are associated with an increased number of CTC admissions, specifically those attributed to the epidemic transmission of acute diarrhoeal diseases, within and between neighbourhoods, independently of their wealth and location relative to surface water sources. The same analysis indicates that CTC admissions attributed to endemic transmission are not associated with average tap water service independently of a neighbourhood's wealth and location relative to surface water sources. But endemic CTC admissions halved over 10 years overall, with a larger reduction in neighbourhoods with a better tap water service on average. The latter are however more strongly affected by the CTC admissions rate occurring in other neighbourhoods.

7.2 Strengths and limitations

7.2.1 CTC admissions and cholera

Limitations related to using CTC admissions as a proxy for cholera incidence were discussed in detail in several chapters of this thesis.

This measure is indeed likely to be biased by differences in healthcare seeking behaviour, between individuals and over time. It would have been very valuable to understand better the determinants of healthcare seeking behaviour in different strata of the population – for instance perception of the severity of symptoms, direct and indirect costs or expected quality of care. Combined quantitative and qualitative methods would be most appropriate and would need to be repeated to assess whether these factors changed over time – especially during cholera outbreaks.

In addition, CTC admissions cannot all be attributed to cholera, as demonstrated in *chapter 3*, and many other enteropathogens were detected amongst CTC-admitted patients, despite only testing for a panel of 15 enteropathogens and on samples preserved for up to several months on filter cards. Without a case-control study and possibly a quantitative measure for the number of organisms present in stool samples for each pathogen of interest, our conclusions about the aetiology of acute diarrhoea in CTC patients are necessarily cautious. At present, a cholera case is considered as confirmed under WHO guidelines if *V. cholerae* O_1 or O_{139} is detected by culture or polymerase chain reaction (PCR) in the stools of an individual, regardless of the concomitant presence of other diarrhoea causing pathogens **[1]**. Our results suggest that this may lead to an overestimation of the number of cases of acute diarrhoea truly caused by cholera in endemic settings where cholera transmission occurs simultaneously with that of many other diarrhoea-causing pathogens and many cholera infected individuals remain asymptomatic. They also raise the question of the impact of infections with multiple enteric pathogens on the severity of cholera illness. Coinfections with several enteric pathogens have been shown to be associated with an increase in severity and duration of diarrhoeal episodes in young children **[2-5]**. Coinfections with other enteropathogens have also been shown to explain some of the decrease in rotavirus vaccine efficacy observed in low and middle income countries in comparison with high income countries **[6]** and their negative impact on immunogenicity of oral cholera vaccines (OCV) in young children may be partly responsible for a lower efficacy in that age group **[7,8]**.

The uncertainties related to the fraction of CTC admissions truly attributable to cholera are compounded by those regarding naturally acquired immunity against the disease, and its duration. It seems reasonable to assume that an individual recovering from a cholera episode has a strong immunity against reinfection with a similar strain for a while and that a susceptible-infected-recovered (SIR) model can be applied to study the dynamics of an outbreak over a couple of years. When extended to 10 years or more as in our setting, a transition back to "susceptible" would be needed, in a SIRS type model. There is a lack of evidence however on how long it takes for a recovered individual to become susceptible again, and whether immunity/susceptibility are modulated by other factors, such as age or ingested dose. Uvira CTC admission rates in 2010 and 2014 were substantially higher than other years – could waning immunity in the population and an increased proportion of susceptible individuals account for these higher rates, in addition to potential climatic factors?

Measuring the incidence of acute diarrhoea in the community and testing each symptomatic individual for cholera (and/or other pathogens of interest) would have been necessary to overcome the challenges posed by using CTC admissions data, and a large case-control study with reliable detection of cholera and other pathogens would have been needed to establish a robust estimate of the cholera attributable fraction. However, both these options would have been costly and not without other biases **[9,10]**. Cross-sectional serology surveys have recently been proposed as a means to estimate cholera incidence in a population over the past year **[11]** and this could potentially be a more feasible approach in a context like Uvira.

7.2.2 Qualitative research

The findings of this thesis are based on several sources of quantitative data, and some aspects of households' perceptions and behaviour would have benefitted from further exploration by qualitative data and analysis.

In addition to better discerning healthcare seeking behaviour determinants (see above), in-depth interviews, observations and focus group methods could have been directed at understanding in more detail the determinants of water source choices for drinking and for domestic activities, as well as coping strategies during tap water supply interruptions. Special attention to the role of costs of tap water (direct and indirect) in this investigation, possibly with a willingness-to-pay approach, would allow making specific recommendations on costs and affordability, especially regarding an improved water service.

Finally, similar methods with households affected by acute diarrhoea, targeted at households' hygiene perceptions and practices when one of its members suffers from diarrhoea, would strengthen our comprehension of mechanisms and interventions that would promote and reinforce hygiene measures preventing diseases transmission within households. This could be the opportunity as well to know more about potential stigma surrounding cholera, often referred to in Uvira as the "disease of dirty hands" ("maladie des mains sales").

7.2.3 Exposure measures

Objectives of the present research were deliberately focused on the role of tap water supply on acute diarrhoeal diseases leading to CTC admissions. Isolating the water issue from that of sanitation and hygiene limited the possibilities to account for either confounding or interactions between those three interlinked prevention measures against diarrhoeal diseases and cholera. Reported access to and use of different types of sanitation, ranging from open defecation to flush toilets, was included in the asset-based wealth index constructed from households surveys and spatial estimates for improved sanitation coverage could be produced by using empirical Bayesian Kriging (EBK), as was done for wealth index, in order to explore whether space-time patterns of CTC admissions are associated to these estimates at neighbourhood level. Sanitation access and use would then need to be excluded from the wealth index to retain wealth as a confounder in the analysis. Correlation with distances to the lake and to rivers is to be expected as sanitation technology is likely dependent on the type of soil in each area. A strong correlation between wealth, tap water service and sanitation may pose challenges for model computation and/or interpretation, and our sanitation indicator would remain a single time point estimate. In addition to improved sanitation coverage, another indicator for sanitation-related risk factors for pathogen transmission could be derived from a distance to wastewater channels and temporary streams.

Figure 7-1 The temporary stream Kibondwe, between the Kibondwe and Kasenga neighbourhoods

Another limitation of our findings is the sub-optimal methodology for measuring drinking and environmental water contamination, with a sensitivity higher than 1 CFU *E.coli* per 100 ml due to sample dilution (drinking water samples) and no replicates for quantification of the contamination (both drinking and environmental water samples). The complexity of detecting and characterising *V. cholerae* strains from water samples (e.g. in viable but non-culturable [VBNC] state, free-form or clumped, toxigenic, non-O_1 non-O_{139}) in a field laboratory also prevented us from exploring further the relevance of such contamination.

To overcome the above limitations would require robust laboratory analyses for water microbiological contamination of tap water quality at multiple points of the network and over time, especially after low pressure events. For the present research we considered that the Regideso treatment plant output had undergone sufficient treatment and was free of bacterial contamination. We had however no evidence to demonstrate that water distributed at the tap remained uncontaminated, with many obvious occasions for pathogen ingress in the distribution network.

7.2.4 Spatial data and ecological approach

Our study made use of high-resolution spatial data when investigating an urban environment. This resolution was key to constructing a tap water service indicator and to estimate several other geographical variables of interest (for example wealth, distances to the lake or the rivers, distance to the CTC, altitude) for chosen spatial units.

One of the challenges with spatial data in our setting at this resolution was the level of correlation between exposures of interest: tap water service, wealth, distance the lake, altitude and distance to the CTC are strongly associated in Uvira, as a result of the urbanisation process over several decades in a constrained physical environment. It was therefore difficult to disentangle the role and importance of each of these factors in relation to CTC admissions, without running into collinearity issues.

Another limitation of our spatial data was the absence of the exact geolocation of the residence of CTC admissions. This meant we had to perform analyses related to CTC admissions at a spatially aggregated level and use statistical methods suitable for discrete spatial units rather than continuous geolocations that are commonly used in other studies **[12-16]**. Relationships between exposures (e.g. tap water service) and disease outcome (e.g. CTC admissions) were therefore established using an ecological approach, with no individual exposure directly linked to CTC admission, with the risk of the ecological fallacy in our interpretations **[17-19]**. Further statistical analyses could be performed by using different spatial models to estimate exposure at aggregated levels from our household data, to overcome such ecological bias in our results (see *chapter / section 6.4.3 Limitations*) **[20,21]**.

7.2.5 Generalisability

The research presented throughout this thesis was conducted over more than 6 years and included retrospective data for a study period going back more than 10 years when considering CTC admissions and tap water distribution, with daily data.

The long duration of these two datasets is extremely valuable for an outbreak-prone disease like cholera, whose incidence varies widely over time in an endemic area such as Uvira. Many studies of cholera in Sub-Saharan Africa either focussed on a single outbreak at a relatively high spatial resolution **[12-15,22-28]**, or on cholera incidence and recurrence over several years but at a national or regional level **[29-36]**. With a high resolution in both time and space, the prolonged study period also allowed us to investigate the importance/role of tap water infrastructure and supply in acute diarrhoeal diseases, in contrast with studies exploring short-term risk factors or one-off emergency response interventions **[37-47]**.

Studying, even in-depth, a single setting nonetheless raises questions about the generalisability of the findings to other urban cholera-endemic places. A major particularity of Uvira is the combination of an abundance of surface water across all parts of town and the absence of other water source options (such as boreholes and wells) except surface or tap water. For instance, in Goma, the provincial capital of North-Kivu, the only free water source is Lake Kivu, and the town stretches up to 5 km away from the lake shore over flat terrain. Other water supply options for households are water sellers in trucks and the Regideso network, with similar unreliability and heterogeneous coverage issues as those experienced in Uvira. Do tap water supply interruptions impact admissions to Goma health facilities and CTCs for acute diarrhoea in a comparable way?

7.3 Implications for research and practice

7.3.1 Further analysis of existing data in Uvira

The existing data in Uvira could be further analysed to try to address some of the limitations raised above and in the previous chapters.

First, the time-series regression analysis (*chapter 5*) could be repeated over the entire study period and used to assess whether 1) the results initially published are indeed robust to a different time period; 2) changing the case definition affects the results – the initial study only considered CTC admissions of over 5-year olds. In addition, the initial study did not detect an effect of chlorination level in the treatment plant output, possibly as a consequence of some missing chlorination data. Using a longer time period may elicit a role for chlorination levels, despite a similar proportion of missing data.

The space-time model could also be run with different geographical aggregations of Analysis Base Units (ABU), with the same or a varying number of geographical units, to attempt and reduce the risk of an ecological fallacy when using aggregated measures of outcome and exposure in populations. The model could also be augmented with an indicator for improved sanitation use – derived from household survey data by EBK as for wealth.

Further work could also be performed on the cholera confirmation dataset to develop a model focusing specifically on predicting the probability that patients not tested during the 37 months of confirmation study would have been RDT-confirmed cholera cases had they been tested, rather than exploring associations between RDT-confirmation and characteristics of admitted patients as was done in *Research paper 1*. This would require a strategy to account for missing data – such as sex, age or duration of stay for example - for some of the patients. If the model performs well, these predicted probabilities could then be integrated into the time-series regression and the space-time model to either 1) restrict analysis to those having been tested positive and those being predicted as such with a probability above a set threshold; or 2) use the predicted probability of each individual admission being confirmed to simulate a large number of daily/weekly counts of datasets per aggregated geographical unit (a single one in the case of the time-series regression), and run the same time-series regression or space-time model on all the simulated datasets, to establish whether the models parameters converge towards particular distributions and summary values.

Independently or in combination with the above, it would be worth exploring further the role of seasonality on CTC admissions and on their relationship with tap water service; open source datasets on temperature, rainfall and chlorophyll concentration in the Lake Tanganyika are readily available on remote-sensing platforms. Official dates for school and bank holidays, or agricultural and fishing seasons, should be straightforward to find. Such investigations are not without challenges, but interest in the impact of climate change of WASH-related diseases has already led to many methodological advances **[48-50]**.

Finally, incorporating the current findings and data into the evaluation of the impact of the tap water supply improvements would be advisable. For instance, the analysis of the step-wedge clustered randomised control trial (sw-cRCT) data may benefit from integrating a baseline tap water service indicator measure for each cluster, and a measure of tap water distribution for each step, in order to quantify the intensity of the intervention each cluster received over time.

7.3.2 Expanding the research in Uvira and similar settings

This thesis work also raised questions that would require further primary data collection in Uvira or in other cholera-endemic urban areas.

The impact of tap water supply intermittency on water quality at the taps undoubtedly deserves further investigation, especially in cholera-endemic places, despite the well-known risk of water to be contaminated during transport and storage. Can water supply network modelling accurately identify areas particularly at risk of temporary lack of residual chlorine and increased levels of faecal and/or specific enteric pathogens contamination when water distribution resumes? How frequently can *V. cholerae*, possibly in VBNC form or within rugose colonies, be isolated from pipes biofilms and/or in tap water after pressure and flow velocity peaks in the network? Would a temporary increase in residual chlorine concentration be sufficient to mitigate this risk? Is there a time limit under which interruptions in distribution would be less detrimental to the water quality? Answers to these questions would be relevant to Uvira, but also to many urban centres, particularly where cholera outbreaks occur (for instance in India or Bangladesh), at a time when water supply intermittency may be unavoidable due to increased pressure on water resources for an exponentially growing urban population **[51]**.

Our findings also further substantiate a wealth of evidence on the importance of the short-cycle transmission of cholera and other diarrhoeal diseases, in households and their immediate and broader environment. More research on how specifically this transmission occurs, through water, food, hands, with possibly insect or animal vectors, fomites as intermediaries, could provide evidence for additional specific intervention opportunities in both private and public domains. How much do rain and wastewater channels and streams contribute to environmental contamination in areas they flow through? Are domestic animals (chicken, goats, pigs) important spreaders of enteric pathogens within and between compounds? Are flies significant vectors of cholera in markets and in/between households? Are specific foods, such as street food, dried fish or mangoes for example, particularly at-risk of cholera contamination and bacterial growth? Even though findings of such investigations may be difficult to translate into preventative actions, absence of evidence is surely a barrier to develop appropriate interventions.

In addition, the African Great Lakes have been hypothesised by some as cholera reservoirs recurrently driving outbreaks occurrence in surrounding endemic areas, although this is contested by others on the basis of clinical strains genomics. A clearer answer to this question would facilitate the design of effective, comprehensive cholera control strategies. Evidence to resolve this should be generated as whole genome sequencing (WGS) and other genomics analyses have become more affordable, ideally in a multi-site study across the African Great Lakes region, with a wide-ranging clinical and environmental sampling strategy **[52]**.

Finally, exploring how and under which circumstances the findings from this research are generalisable to other comparable places would provide broader and stronger evidence to adapt and strengthen cholera control and response strategies.

7.3.3 Strengthening acute diarrhoeal diseases surveillance in cholera endemic settings

Our results highlight the need to improve surveillance of acute diarrhoea in cholera endemic settings, to ensure timely detection of actual cholera outbreaks requiring an immediate response with a consideration of other possible aetiologies contributing to incident cases of diarrhoea. More accurate and timely information is key to roll-out cost-effective interventions in the right places and identify priority areas for response and control strategies.

Improving the sensitivity and specificity of the clinical definition of a suspected cholera case in endemic settings, possibly including locally identified criteria such as age, gender or place of residence, could be envisaged, with a capacity building program targeted at the clinical assessment upon patient presentation. Smartphone decision-support tools could be trialled for this purpose **[53]**. Increasing cholera confirmation capacity, with the best available rapid diagnostic tests and a clear set of criteria to determine when and how to use them (e.g. testing x patients above a certain number of admissions from a specific area) would be recommended as well, in order to be able to distinguish between 1) acute diarrhoea cases admitted to a CTC (or presenting to an outpatient clinic); 2) suspected cholera cases (see above); 3) and cholera-confirmed cases.

More accurate data on absolute numbers of cases should be complemented by a more systematic collection of residence location as health facility catchment areas may vary substantially over time (for instance when NGO support is rolled out for an outbreak response and modifies healthcare seeking behaviours). In turn, population estimates for geographical units used for residence recording should be accurate enough and updated or refined if needed. Geographical scale in urban centres should be commensurate with that of control interventions considered, and a higher spatial resolution of data may be necessary to develop case-area targeted interventions (CATI) programs.

7.3.4 Improvements to the response and control strategies

Differences in coverage and intermittency of tap water supply in Uvira have been shown to be determinants of drinking water quality and quantity of water use by households, as well as CTC admissions across neighbourhoods and over time. Our results support a stronger integration of the existing tap water supply system – and its weaknesses - into both the short-term response to a surge in diarrhoeal disease admissions and the long-term control strategy; some recommendations have been made in *chapters 5 and 6*.

Long-term improvements to the tap water infrastructure should address the unreliability and intermittency of the supply: 1) with a larger amount of treated water stored before being distributed, to act as a "buffer" for short interruptions at the water treatment plant; 2) with a reinforced distribution system that improves pressure and flow at the most distant taps; 3) with a clear mitigation plan for power cuts and how/when to best to use generators and fuel resources; 4) ultimately with an independent source of energy to reduce dependency on the unreliable power grid and on a costly generator alternative. Other infrastructure improvements could also compartmentalise the distribution network further than the single partition north/south currently existing, to enable a scheduled rolling supply by areas when production cannot meet the overall demand and ensure that taps in all areas are served in turn, instead of only those nearest to the reservoir.

Increasing geographical coverage through public tap installation is likely to encourage households to use tap water more often as their drinking water source, as most already do, albeit not all of the time if the service remains unpredictable. Areas near the lake and the rivers could be particularly targeted to reduce the attraction of these alternative sources in households' choice for drinking water, although this raises the question of equity of tap water access improvements. To achieve health benefits with increased coverage and use of tap water for drinking, levels of residual chlorine in the water at the tap need to be sufficient to 1) guarantee an absence of recontamination through the distribution network; 2) ensure some protection against recontamination during transport and storage. Additional infrastructure, such as chlorine supplementation pumps at key points of the network, may need to be put in place and operated **[54]**.

The "water consumption" plateau described in *chapter 4* suggests that bringing tap water closer to households through public taps outside their compounds should not be expected to have a notable impact on the quantity of water used for domestic activities, except for those usually spending more than 90 minutes per trip to collect water. Financial incentives for households to subscribe to a tap connection within their premises could be envisaged, possibly by street or restricted area to reduce the unit cost of installation by grouping the necessary works on pipes.

It is worth stressing that all infrastructure improvements would need to be accompanied by a strong capacity building component for the operator (the Regideso) and careful financial adjustments to ensure sustainability.

Preparedness for incidence peaks should include support to the above long-term improvements and assessing the remaining gaps in tap water provision that would temporarily need to be bridged to reduce short-cycle transmission. As was demonstrated in *chapter 6*, interventions cutting short-cycle transmission are not limited to those deployed directly in households, but could include 1) direct support to the water treatment plant to operate the generator during low power grid supply; 2) additional chlorination at public taps after interruptions if needed[7]. Case area targeted interventions (CATI) may be rolled out in and around affected households and designed to support water use for hygiene purposes. Our findings point to the challenges of including vaccination and/or antibiotic prophylaxis in CATI packages for cholera endemic areas such as Uvira, as even during high incidence periods, households identified through CTC admissions may not be affected by cholera but by diarrhoea caused by another pathogen for which vaccination and/or antibiotic prophylaxis may be ineffective.

7.4 Personal reflections on the learning process

This thesis is the result of 6 years of work as a research degree student and a staff member in the Environmental Health Group (EHG), initiated under the supervision of Dr Jeroen Ensink in EHG, and continued in 2016 under that of Prof. Simon Cousens in the Infectious Diseases Epidemiology (IDE) department.

The research was initially funded to conduct an impact evaluation of improvements to the tap water supply and was therefore bound to adapt to the implementation of these large infrastructure works. These resulted in multiple alterations to the research protocol as delays and changes in the actual improvements accumulated. I consequently learned how to adapt to make the best use of available data and additional study duration, in response to the evolving circumstances. The absence of a permanent structure in the field to host the research and my limited time on-site, as a consequence of high costs for logistics and a volatile political and security situation, also posed challenges.

Such an evolving research process, with unavoidable deviations from a clear protocol aimed at testing specific pre-specified hypotheses with identified methods, is particularly vulnerable to data-driven and researcher biases. It seems however to be a necessary compromise to generate evidence in and relevant to unpredictable and complex settings, in which diseases such as cholera thrive. If additional proof was needed, as the improvements works and their impact evaluation were still on-going, unprecedented flooding affected Uvira on the 17th of April 2020, seriously damaging the Regideso water treatment plant and the tap water distribution network, including in areas that had already been targeted by improvements. This happened at the same time as the Covid-19 pandemic, that generated an intense global initiative to promote handwashing with soap, including in humanitarian settings. Both could have notable but opposing effects on the transmission of diarrhoeal diseases and cholera in Uvira, and the impact evaluation approach will have to adjust to these unpredictable events in a pragmatic way.

[7] At the time of writing, several cities of DRC have communicated on an increase in tap water supply by the Regideso to fight against the Covid-19 pandemic.

Figure 7-2 Floodings in the Kilibula and Kimanga neighbourhoods on the 17th of April 2020 (Photo courtesy P. Delaunoy)

Part of the necessary pragmatism has been to rely heavily on data routinely collected by local collaborating parties, such as the health authorities and the Regideso. These routine data were complemented by newly generated data, such as the geolocation of tap connections and rapid diagnostic test confirmation of CTC admissions, along with a more formal register collecting more comprehensive information on admissions and an extensive mapping of the town, but this required a relatively limited effort in comparison with setting up a wholly independent data collection process. It demonstrates that surveillance data and routine administrative or commercial records can be extremely valuable when their quality can be assessed and strengthened, while keeping their limitations in mind when analysing them. Sources of information on populations and behaviours are increasing at a fast pace – mobile phone network data and use of social media are already being used to estimate population movements during outbreaks, while high resolution remote sensing is becoming the norm in many scientific fields. Integrating the use of these data sources into research protocols should encourage research in complex emergencies or any location where specific data collection operations are challenging or comparatively costly for the added value.

Surveillance or routine data are still insufficient to document and understand better what households actually do when it came to domestic and drinking water consumption. The complexity of properly implementing a survey in hundreds of households is barely perceptible from articles reporting their findings, where sampling strategy, questionnaires development, interviewers training, daily logistics and data management rarely merit more than a couple of sentences. Leading, planning and managing such primary data collection in Uvira, with minimal support from the local OXFAM office that had shrunk to a bare minimum by end of 2016 and fully closed by mid-2017, was fraught with obstacles. From interviewer and supervisor selection and recruitment, their training to higher standards than they were used to for rapid assessments, negotiating transport and communication compensations, enforcing starting times and minimum daily number of interviews, to mitigating as much as possible the personal risks each of them was taking in visiting remote areas of Uvira, the human resources aspects were particularly difficult to navigate, even if the trusting and respectful relationships established and renewed after a year were amongst the most rewarding features of my fieldwork.

On a different note, the present work built largely on spatial analysis methods to generate aggregated estimates or construct a survey sampling frame. These methods were invaluable in the context and the data sources described above. Their heavy use was not initially planned, and this came at a cost in terms of the time required by iterative self-learning periods on using the R environment and specific packages, as well as the ArcGIS suite of software – in addition to some free “competitors” such as QGIS, GME, Satscan, GeoDa, Epanet…. Knowing in advance the direction the research would take would have encouraged me to invest time early on in attending advanced training on spatial methods and statistics, as well as on using R, as I have no doubt that spatial methods and statistics with R are increasingly valuable in epidemiology of infectious diseases.

Further, using several statistical and epidemiological approaches – cross-sectional, predictive, time-series, time-space – has possibly led to slower progress and steep learning curves, but enlightened me about the benefits of each of them, their challenges, their limitations and their combined “triangulation” possibilities.

7.5 Concluding remarks

A renewed global commitment to controlling cholera was made in 2017 by the Global Task Force on Cholera Control with the ambitious goals of reducing cholera deaths by 90 % and stopping cholera transmission in 20 countries by 2030. The proposed strategy recognizes the need to combine short and long-term control efforts and to target cholera hotspots, such as the Uvira area. Prevention of cholera will not be achieved in the long run without building or strengthening sustainable tap water supply systems in urban areas, even if cholera vaccination can buy some time by reducing the lives lost to yet another vaccine preventable disease. Cholera is, first and foremost, a disease of poverty (one of many) and one should not lose sight of the great burden of non-cholera diarrheal diseases that is borne by the very same people vulnerable to cholera. We must take every opportunity to address the vicious circle between poverty and disease, for which there is no silver bullet.

7.6 Chapter 7 - References

1. Global Task Force For Cholera Control - Surveillance WG. *Interim Guidance Document on Cholera Surveillance.* 2017.
2. Bhattarai V, Sharma S, Rijal KR, Banjara MR. *Co-infection with Campylobacter and rotavirus in less than 5 year old children with acute gastroenteritis in Nepal during 2017–2018*. **BMC Pediatrics**. 2020;20(1):68.
3. Bhavnani D, Goldstick JE, Cevallos W, Trueba G, Eisenberg JN. *Synergistic effects between rotavirus and coinfecting pathogens on diarrheal disease: evidence from a community-based study in northwestern Ecuador*. **American journal of epidemiology**. 2012;176(5):387-395.
4. Bonkoungou IJO, Haukka K, Österblad M, et al. *Bacterial and viral etiology of childhood diarrhea in Ouagadougou, Burkina Faso*. **BMC pediatrics**. 2013;13(1):36.
5. Valentini D, Vittucci AC, Grandin A, et al. *Coinfection in acute gastroenteritis predicts a more severe clinical course in children*. **European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology**. 2013;32.
6. Praharaj I, Platts-Mills JA, Taneja S, et al. *Diarrheal Etiology and Impact of Coinfections on Rotavirus Vaccine Efficacy Estimates in a Clinical Trial of a Monovalent Human-Bovine (116E) Oral Rotavirus Vaccine, Rotavac, India.* **Clinical infectious diseases : an official publication of the Infectious Diseases Society of America**. 2019;69(2):243-250.
7. Leung DT, Rahman MA, Mohasin M, et al. *Memory B cell and other immune responses in children receiving two doses of an oral killed cholera vaccine*

compared to responses following natural cholera infection in Bangladesh. **Clin Vaccine Immunol**. 2012;19(5):690-698.

8. Shaikh H, Lynch J, Kim J, Excler J-L. *Current and future cholera vaccines*. **Vaccine**. 2020;38:A118-A126.
9. Kopec JA, Esdaile JM. *Bias in case-control studies. A review*. **J Epidemiol Community Health**. 1990;44(3):179-186.
10. Sedgwick P. *Bias in observational study designs: case-control studies*. **BMJ : British Medical Journal**. 2015;350:h560.
11. Azman AS, Lessler J, Luquero FJ, et al. *Estimating cholera incidence with cross-sectional serology*. **Sci Transl Med**. 2019;11(480).
12. Azman AS, Luquero FJ, Rodrigues A, et al. *Urban cholera transmission hotspots and their implications for reactive vaccination: evidence from Bissau city, Guinea bissau*. **PLoS Negl Trop Dis**. 2012;6(11):e1901.
13. Azman AS, Luquero FJ, Salje H, et al. *Micro-Hotspots of Risk in Urban Cholera Epidemics*. **J Infect Dis**. 2018;218(7):1164-1168.
14. Bi Q, Azman AS, Satter SM, et al. *Micro-scale Spatial Clustering of Cholera Risk Factors in Urban Bangladesh*. **PLoS Negl Trop Dis**. 2016;10(2):e0004400.
15. Debes AK, Ali M, Azman AS, Yunus M, Sack DA. *Cholera cases cluster in time and space in Matlab, Bangladesh: implications for targeted preventive interventions*. **International journal of epidemiology**. 2016;45(6):2134-2139.
16. Finger F, Bertuzzo E, Luquero FJ, et al. *The potential impact of case-area targeted interventions in response to cholera outbreaks: A modeling study*. **PLoS Med**. 2018;15(2):e1002509.
17. Idrovo AJ. *Three criteria for ecological fallacy*. **Environ Health Perspect**. 2011;119(8):A332-A332.
18. Sedgwick P. *Understanding the ecological fallacy*. **BMJ : British Medical Journal**. 2015;351:h4773.
19. Shafran-Nathan R, Levy I, Levin N, Broday DM. *Ecological bias in environmental health studies: the problem of aggregation of multiple data sources*. **Air Quality, Atmosphere & Health**. 2017;10(4):411-420.
20. Lancaster GA, Green M, Lane S. *Reducing bias in ecological studies: an evaluation of different methodologies*. **Journal of the Royal Statistical Society: Series A (Statistics in Society)**. 2006;169(4):681-700.
21. Salway R, Wakefield J. *A hybrid model for reducing ecological bias*. **Biostatistics**. 2008;9(1):1-17.
22. Allan M, Grandesso F, Pierre R, et al. *High-resolution spatial analysis of cholera patients reported in Artibonite department, Haiti in 2010-2011*. **Epidemics**. 2016;14:1-10.
23. Blake A, Keita VS, Sauvageot D, et al. *Temporo-spatial dynamics and behavioural patterns of 2012 cholera epidemic in the African mega-city of Conakry, Guinea*. **Infectious diseases of poverty**. 2018;7(1):13.
24. Curtis A, Squires R, Rouzier V, et al. *Micro-Space Complexity and Context in the Space-Time Variation in Enteric Disease Risk for Three Informal Settlements of Port au Prince, Haiti*. **Int J Environ Res Public Health**. 2019;16(5).
25. Mukandavire Z, Liao S, Wang J, Gaff H, Smith DL, Morris JG. *Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe*. **Proceedings of the National Academy of Sciences**. 2011;108(21):8767-8772.
26. Nilima, Kamath A, Shetty K, Unnikrishnan B, Kaushik S, Rai SN. *Prevalence, patterns, and predictors of diarrhea: a spatial-temporal comprehensive evaluation in India*. **BMC public health**. 2018;18(1):1288.

27. Olanrewaju OE, Adepoju KA. *Geospatial Assessment of Cholera in a Rapidly Urbanizing Environment*. **Journal of environmental and public health**. 2017;2017:6847376.

28. Osei FB, Stein A. *Temporal trend and spatial clustering of cholera epidemic in Kumasi-Ghana*. **Sci Rep**. 2018;8(1):17848.

29. Bwire G, Ali M, Sack DA, et al. *Identifying cholera "hotspots" in Uganda: An analysis of cholera surveillance data from 2011 to 2016*. **PLoS Negl Trop Dis**. 2017;11(12):e0006118.

30. Cowman G, Otipo S, Njeru I, et al. *Factors associated with cholera in Kenya, 2008-2013*. **The Pan African medical journal**. 2017;28:101.

31. Ingelbeen B, Hendrickx D, Miwanda B, et al. *Recurrent Cholera Outbreaks, Democratic Republic of the Congo, 2008-2017*. **Emerg Infect Dis**. 2019;25(5):856-864.

32. Lessler J, Moore SM, Luquero FJ, et al. *Mapping the burden of cholera in sub-Saharan Africa and implications for control: an analysis of data across geographical scales*. **Lancet**. 2018;391(10133):1908-1915.

33. Moore SM, Azman AS, Zaitchik BF, et al. *El Nino and the shifting geography of cholera in Africa*. **Proceedings of the National Academy of Sciences of the United States of America**. 2017;114(17):4436-4441.

34. Ngwa MC, Liang S, Kracalik IT, et al. *Cholera in Cameroon, 2000-2012: Spatial and Temporal Analysis at the Operational (Health District) and Sub Climate Levels*. **PLoS Negl Trop Dis**. 2016;10(11):e0005105.

35. Rebaudet S, Sudre B, Faucher B, Piarroux R. *Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes*. **J Infect Dis**. 2013;208 Suppl 1:S46-54.

36. Rieckmann A, Tamason CC, Gurley ES, Rod NH, Jensen PKM. *Exploring Droughts and Floods and Their Association with Cholera Outbreaks in Sub-Saharan Africa: A Register-Based Ecological Study from 1990 to 2010*. **Am J Trop Med Hyg**. 2018;98(5):1269-1274.

37. Childs L, Francois J, Choudhury A, et al. *Evaluation of Knowledge and Practices Regarding Cholera, Water Treatment, Hygiene, and Sanitation Before and After an Oral Cholera Vaccination Campaign-Haiti, 2013-2014*. **Am J Trop Med Hyg**. 2016;95(6):1305-1313.

38. Ciglenecki I, Azman AS, Jamet C, Serafini M, Luquero FJ, Cabrol JC. *Progress and Challenges in Using Oral Cholera Vaccines to Control Outbreaks: The Medecins Sans Frontieres Experience*. **J Infect Dis**. 2018;218(suppl_3):S165-s166.

39. Dureab F, Jahn A, Krisam J, et al. *Risk factors associated with the recent cholera outbreak in Yemen: a case-control study*. **Epidemiology and health**. 2019;41:e2019015.

40. Gorham TJ, Yoo J, Garabed R, Mouhaman A, Lee J. *Water Access, Sanitation, and Hygiene Conditions and Health Outcomes among Two Settlement Types in Rural Far North Cameroon*. **Int J Environ Res Public Health**. 2017;14(4).

41. Kwesiga B, Pande G, Ario AR, Tumwesigye NM, Matovu JKB, Zhu BP. *A prolonged, community-wide cholera outbreak associated with drinking water contaminated by sewage in Kasese District, western Uganda*. **BMC public health**. 2017;18(1):30.

42. Massing LA, Aboubakar S, Blake A, et al. *Highly targeted cholera vaccination campaigns in urban setting are feasible: The experience in Kalemie, Democratic Republic of Congo*. **PLoS Negl Trop Dis**. 2018;12(5):e0006369.

43. Matias WR, Teng JE, Hilaire IJ, Harris JB, Franke MF, Ivers LC. *Household and Individual Risk Factors for Cholera among Cholera Vaccine Recipients in Rural Haiti*. **Am J Trop Med Hyg**. 2017;97(2):436-442.

44. Rajasingham A, Hardy C, Kamwaga S, et al. *Evaluation of an Emergency Bulk Chlorination Project Targeting Drinking Water Vendors in Cholera-Affected Wards of Dar es Salaam and Morogoro, Tanzania*. **Am J Trop Med Hyg**. 2019;100(6):1335-1341.

45. Rebaudet S, Bulit G, Gaudart J, et al. *The case-area targeted rapid response strategy to control cholera in Haiti: a four-year implementation study*. **PLoS Negl Trop Dis**. 2019;13(4):e0007263.

46. Shannon K, Hast M, Azman AS, Legros D, McKay H, Lessler J. *Cholera prevention and control in refugee settings: Successes and continued challenges*. **PLoS Negl Trop Dis**. 2019;13(6):e0007347.

47. Spiegel P, Ratnayake R, Hellman N, et al. *Responding to epidemics in large-scale humanitarian crises: a case study of the cholera response in Yemen, 2016-2018*. **BMJ Glob Health**. 2019;4(4):e001709.

48. Jones N, Bouzid M, Few R, Hunter P, Lake I. *Water, sanitation and hygiene risk factors for the transmission of cholera in a changing climate: using a systematic review to develop a causal process diagram*. **J Water Health**. 2020;18(2):145-158.

49. Lemaitre J, Pasetto D, Perez-Saez J, Sciarra C, Wamala JF, Rinaldo A. *Rainfall as a driver of epidemic cholera: Comparative model assessments of the effect of intra-seasonal precipitation events*. **Acta tropica**. 2019;190:235-243.

50. Lo Iacono G, Armstrong B, Fleming LE, et al. *Challenges in developing methods for quantifying the effects of weather and climate on water-associated diseases: A systematic review*. **PLoS Negl Trop Dis**. 2017;11(6):e0005659.

51. Luby SP, Davis J, Brown RR, Gorelick SM, Wong THF. *Broad approaches to cholera control in Asia: Water, sanitation and handwashing*. **Vaccine**. 2020;38 Suppl 1:A110-A117.

52. Mavian C, Paisie TK, Alam MT, et al. *Toxigenic Vibrio cholerae evolution and establishment of reservoirs in aquatic ecosystems*. **Proceedings of the National Academy of Sciences of the United States of America**. 2020;117(14):7897-7904.

53. Haque F, Ball RL, Khatun S, et al. *Evaluation of a Smartphone Decision-Support Tool for Diarrheal Disease Management in a Resource-Limited Setting*. **PLoS Negl Trop Dis**. 2017;11(1):e0005290.

54. Pickering AJ, Crider Y, Sultana S, et al. *Effect of in-line drinking water chlorination at the point of collection on child diarrhoea in urban Bangladesh: a double-blind, cluster-randomised controlled trial*. **Lancet Glob Health**. 2019;7(9):e1247-e1256.

8

Appendices

Appendix 2-1 Maps of health divisions, neighbourhoods and streets

Quartier Rugenge / AS Kavimvira

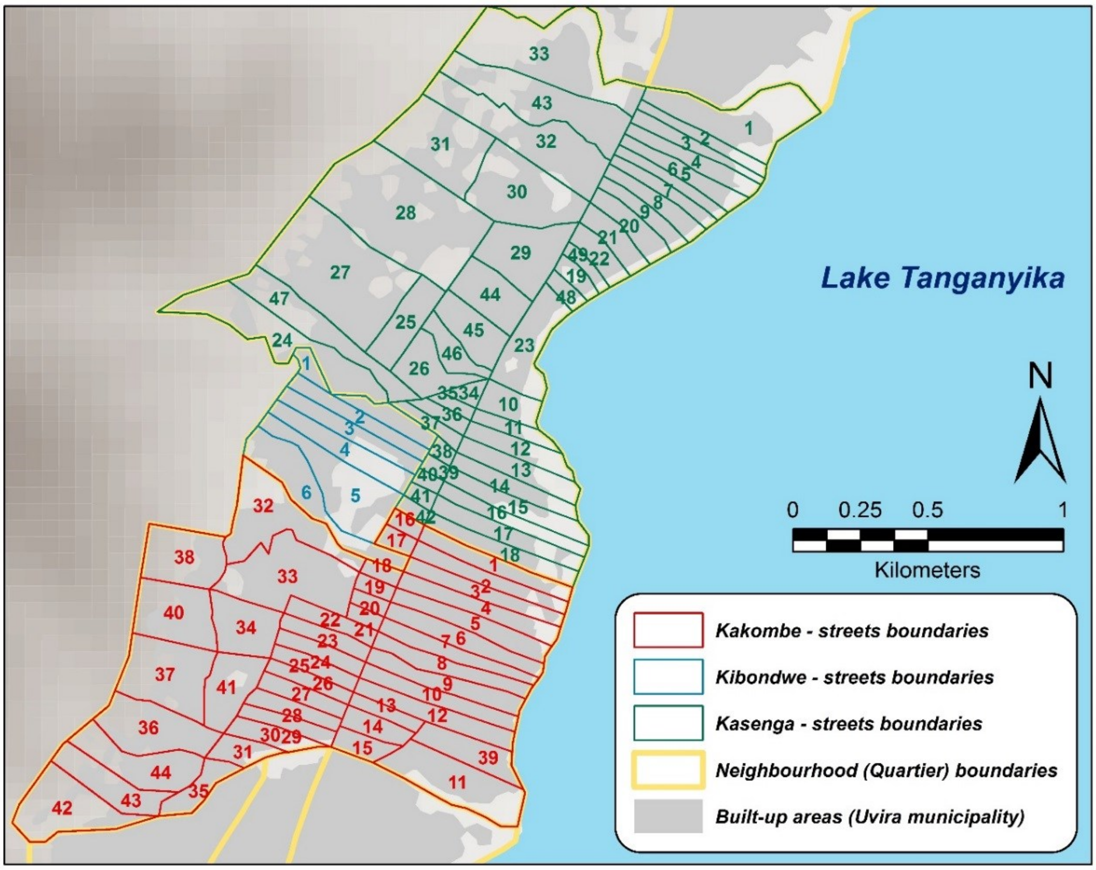
1 Av. Makarunga
2 Av. Rugenge nord
3 Av. Rugenge sud
4 Av. De la paroisse
5 Av. Kinaga
6 Av. Maendeleo
7 Av. Ushirika
8 Av. Petrocongo

Quartier Kavimvira / AS Kilomoni

1 Av. Ndava
2 Av. CPGL
3 Av. Mapendo
4 Av. Tanganyika
5 Av. Lumumba
6 Av. De la plage
7 Av. Nyangara
8 Av. Rubenga
9 Av. Du marche
10 Av. Kamanyola
11 Av. Mobutu
12 Av. Kimbangu
13 Av. Du projet
14 Av. Maiyamoto
15 Av. Du lac

Quartier Kavimvira / AS Kavimvira

16 Av. Maiyamoto
17 Av. Du projet
18 Av. Kimbangu
19 Av. Kalembelembe
20 Av. Kasavubu
21 Av. Rond point
22 Av. Tupendane



Quartier Kakombe / AS Kasenga Etat
1 Av. Nyatutwa
2 Av. Kamanyola
3 Av. Reboisement
4 Av. Ubwari
5 Av. Membo
6 Av. Musulmane
7 Av. Goma
8 Av. Shaba
9 Av. Virunga
10 Av. Kitundu
11 Av. Lenghe III
12 Av. Kivu sud
13 Av. Bavi
14 Av. Kasavubu
15 Av. Simba
39 Av. Maendeleo

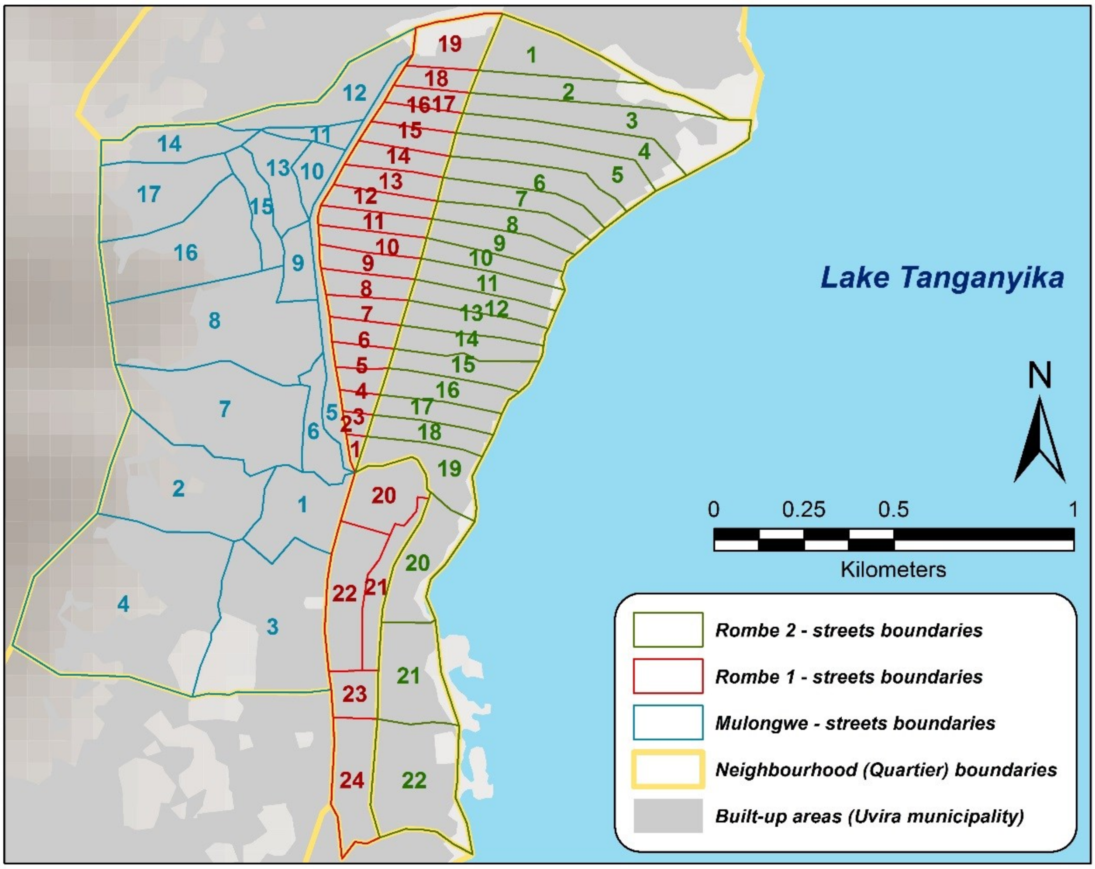
Quartier Kakombe / AS Kasenga Cepac
16 Av. Nyatutwa
17 Av. Kamanyola
18 Av. Ubwari
19 Av. Membo
20 Av. Musulmane
21 Av. Goma
22 Av. Shaba
23 Av. Virunga
24 Av. Kitundu
25 Av. Kivu
26 Av. Bavi
27 Av. Kasavubu
28 Av. Kalundu
29 Av. Simba
30 Av. Likasi
31 Av. Mukulima
32 Av. Nakaziba
33 Av. Ujuzi
34 Av. Du peuple
35 Av. Cinq chantiers
36 Av. Nyakyoya
37 Av. Lukula
38 Av. Kakombe
40 Av. Kalungwe
41 Av. Kigongo
42 Av. Sange
43 Av. Montngaliema
44 Av. Kavuye

Quartier Kibondwe / AS Kasenga Cepac
1 Av. Mushule
2 Av. Kitunya
3 Av. Bajoba
4 Av. Bushoga
5 Av. Kabego
6 Av. Reboisement

Quartier Kasenga / AS Kasenga Etat
1 Av. Muranvya
2 Av. Kitumaini
3 Av. Umoja
4 Av. Maendeleo
5 Av. Hewa bora
6 Av. Muungano
7 Av. Majengo
8 Av. Lala salama
9 Av. Mwangaza
10 Av. De la mosquee
11 Av. Tanganyika
12 Av. Salongo
13 Av. Kikula
14 Av. Ginki
15 Av. De la gombe
16 Av. Regezamwendo
17 Av. Bukavu
18 Av. Azuhuri
19 Av. Uwezo
20 Av. Kivu nord
21 Av. Kigobe
22 Av. Tupendane
23 Av. Du port
48 Av. Kasablanca
49 Av. Amani

Quartier Kasenga / AS Kiyaya
24 Av. Kisangani
25 Av. Du general
26 Av. Hebroni
27 Av. Kiyaya ouest
28 Av. Universite
29 Av. Kabomboza
30 Av. Conforti
31 Av. Budota
32 Av. Mangondo
33 Av. Kilima hewa
43 Av. Musheru
44 Av. Kasenga centre
45 Av. Mapendo
46 Av. Bondogolo
47 Av. Du petit pont

Quartier Kasenga / AS Kasenga Cepac
34 Av. De la mosquee
35 Av. Tanganyika
36 Av. Salongo
37 Av. Kikula
38 Av. Ginki
39 Av. De la gombe
40 Av. Regezamwendo
41 Av. Bukavu
42 Av. Azuhuri



Quartier Rombe 2 / AS Mulongwe

1 Av. Munanira
2 Av. Mulongwe
3 Av. Kabare
4 Av. Kalehe
5 Av. Du 24 novembre
6 Av. Uvira
7 Av. Fizi
8 Av. Haut congo
9 Av. Fac
10 Av. Mwenga
11 Av. Shabunda
12 Av. De l authenticite
13 Av. Du 30 juin
14 Av. Walungu
15 Av. Du 04 janvier
16 Av. Bas congo
17 Av. Du 15 decembre
18 Av. Du 27 octobre
19 Av. Major vangu

Quartier Rombe 2 / AS Tanganyika

20 Av. Idjwi 1
21 Av. Idjwi 2
22 Av. Idjwi 3

Quartier Rombe 1 / AS Rombe 1

1 Av. Major vangu
2 Av. Du 27 octobre
3 Av. Du 15 decembre
4 Av. Bas congo
5 Av. Du 04 janvier
6 Av. Walungu
7 Av. Du 30 juin
8 Av. De l authenticite
9 Av. Shabunda
10 Av. Mwenga
11 Av. Fac
12 Av. Haut congo
13 Av. Fizi
14 Av. Uvira
15 Av. Du 24 novembre
16 Av. Kalehe
17 Av. Kabare
18 Av. Mulongwe
19 Av. Munanira

Quartier Rombe 1 / AS Tanganyika

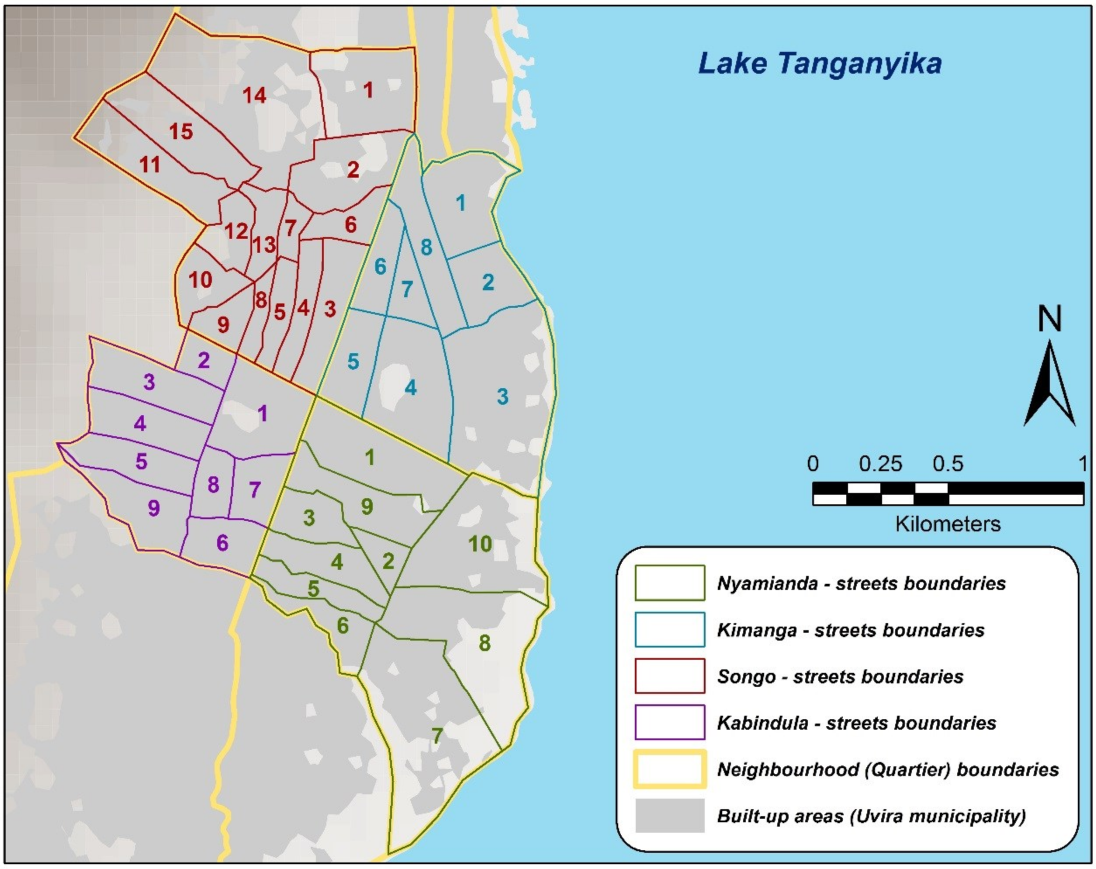
20 Av. Kakungwe 1
21 Av. Kakungwe 3
22 Av. Kakungwe 2
23 Av. Bralima 1
24 Av. Bralima 2

Quartier Mulongwe / AS Tanganyika

1 Av. Shishi 1
2 Av. Shishi 3
3 Av. Shishi 2
4 Av. Shishi 4

Quartier Mulongwe / AS Mitumba

5 Av. Lumumba
6 Av. Kasavubu
7 Av. Kayaja 1 2 3 4
8 Av. Mitumba 1 2 3
9 Av. Matadi 2
10 Av. Makobola
11 Av. Rwegereza
12 Av. De la cite
13 Av. Matadi 1
14 Av. Yohana
15 Av. Kitunge
16 Av. Apollo 1
17 Av. Apollo 2



Quartier Nyamianda / AS Nyamianda

1 Av. Alliance
2 Av. Goma
3 Av. Munanira
4 Av. Mundi
5 Av. Musumba
6 Av. Embouchure
7 Av. Plage d or
8 Av. Lumbulumbu
9 Av. Isiro
10 Av. Kivu

Quartier Kimanga / AS Kimanga

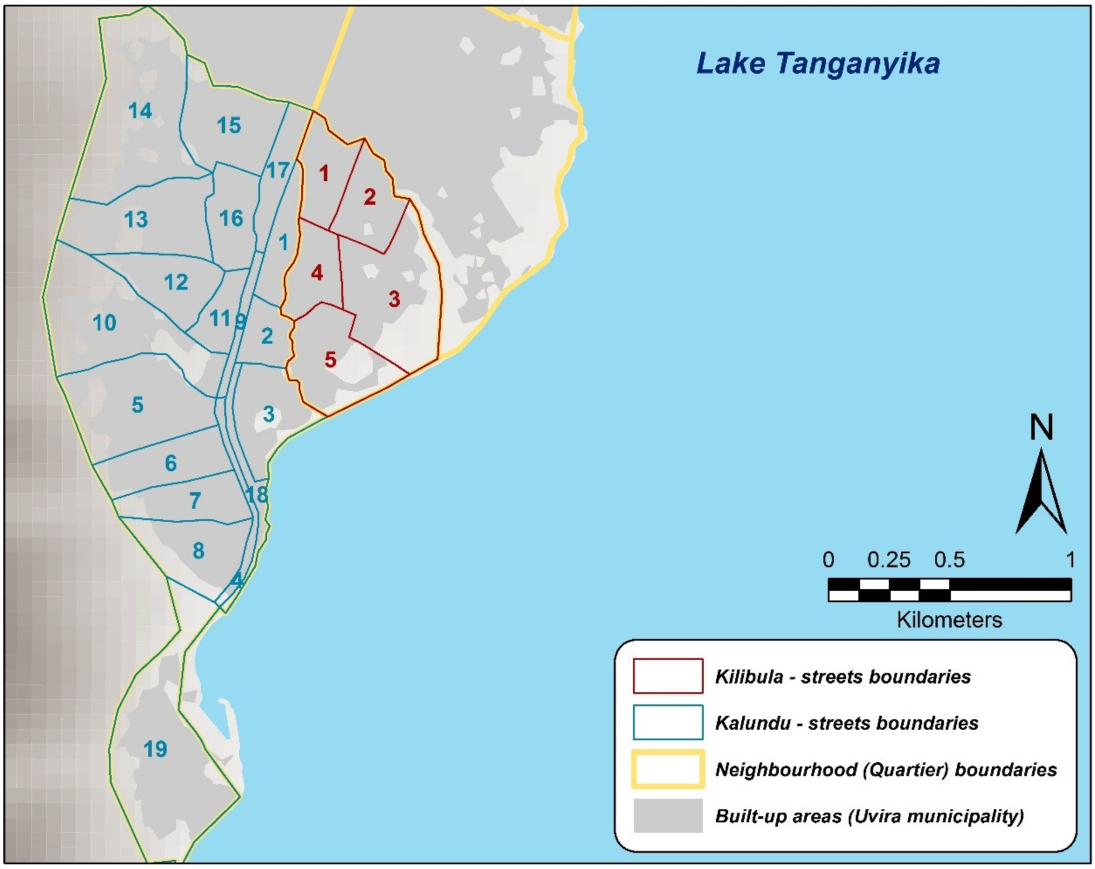
1 Av. Kabungulu 2
2 Av. Kabungulu 1
3 Av. Nyamianda 1
4 Av. Du stade
5 Av. De la paix
6 Av. Du pionnier
7 Av. Kimanga
8 Av. Nyamianda 2

Quartier Songo / AS Saint-Paul

1 Av. Mapinduzi
2 Av. Du stade
3 Av. Alpha
4 Av. Democratie
5 Av. Alliance
6 Av. Mapendano
7 Av. Du marche
8 Av. Maendeleo
9 Av. Lumumba
10 Av. Matumaini
11 Av. Du 17 mai
12 Av. Umoja
13 Av. De la paix
14 Av. De la mission
15 Av. Du progres

Quartier Kabindula / AS Kabindula

1 Av. Kimbangu
2 Av. Elimu
3 Av. Musohoko
4 Av. Musabwa
5 Av. Kashekebwe
6 Av. Kalimabenge
7 Av. Kirambo
8 Av. Maombi
9 Av. Kijaga



Quartier Kilibula / AS Kalundu Etat

1 Av. Kyonga
2 Av. Shaba
3 Av. Mombasa
4 Av. Maendeleo
5 Av. Lenghe III

Quartier Kalundu / AS Kalundu Etat

1 Av. Solange
2 Av. Kinogono 1
3 Av. Kinogono 2
18 Av. Centre commercial

Quartier Kalundu / AS Kalundu Catholique

4 Av. Centre commercial
5 Av. Kagenge
6 Av. Mutarure
7 Av. Ngovi mgja
8 Av. Kamongola
19 Av. Du port

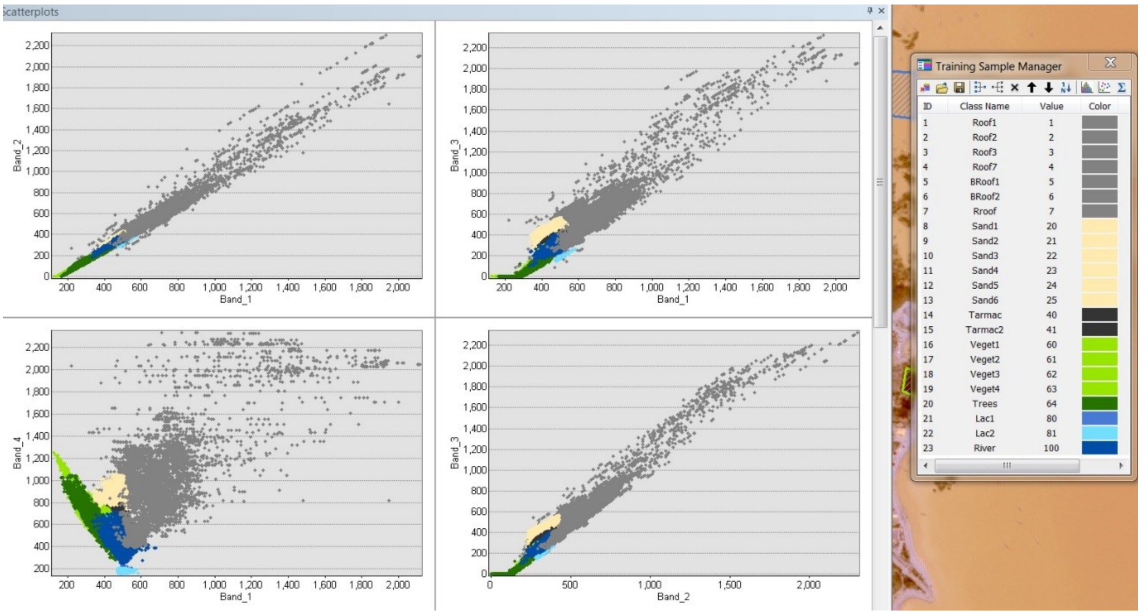
Quartier Kalundu / AS Kalundu Cepac

9 Av. Centre commercial
10 Av. Kasia
11 Av. Nyoroka 1
12 Av. Nyoroka 2
13 Av. Karigo
14 Av. Umoja
15 Av. Rugembe
16 Av. Kakamba
17 Av. Bongisa

Appendix 2-2 Buildings classification process

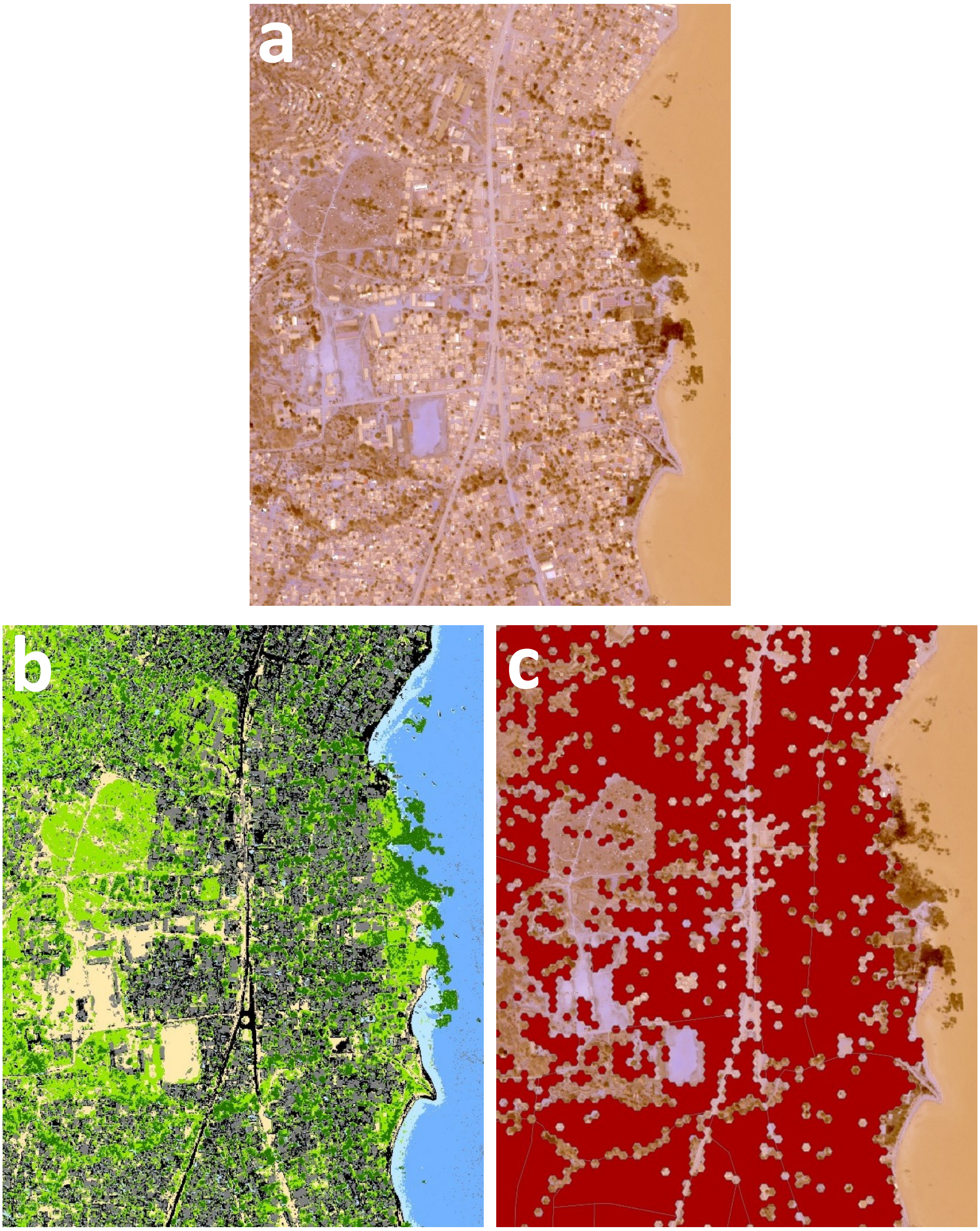
Stage 1: Manual delimitation of training samples (23 classes)

Stage 2: Classes grouping with minimal spectral signature overlapping over the 4 bands (7 classes)



Stages 3 and 4: Maximum likelihood classification of the entire raster and majority classification of 10x10m hexagons based on the "roof" class

a) Original satellite image; b) classified image into 7 classes, with roof surfaces in grey; c)10x10m hexagons tessellation classified into built / not built according to the majority of raster cells



Appendix 2-3 Population census for streets and neighbourhoods in November 2017

Neighbourhood / Street	Census Nov. 2017	Analysis base unit (ABU)	Proportion of neighbourhood population
Kabindula			
Av.Elimu	821	av.elimu.kab_a	5.1%
Av.Kalimabenge	1729	av.kalimabenge.kab_a	10.8%
Av.Kashekebwe	2132	av.kashekebwe.kab_a	13.3%
Av.Kijaga	2272	av.kijaga.kab_a	14.2%
Av.Kimbangu	2568	av.kimbangu.kab_a	16.1%
Av.Kirambo	1686	av.kirambo.kab_a	10.6%
Av.Maombi	1616	av.maombi.kab_a	10.1%
Av.Musabwa	1814	av.musabwa.kab_a	11.4%
Av.Musohoko	1343	av.musohoko.kab_a	8.4%
Kakombe			
Av.Bavi*	1374	av.bavi.kak_a	4.7%
Av.Cinq_chantiers	852	av.cinq_chantiers.kak_a	2.9%
Av.Du_peuple	613	av.du_peuple.kak_a	4.1%
Av.Kalungwe	583		
Av.Goma*	964	av.goma.kak_a	3.3%
Av.Kakombe	673	av.kakombe.kak_a	2.3%
Av.Kalundu	882	av.kalundu.kak_a	3.0%
Av.Kamanyola*	805	av.kamanyola.kak_a	2.7%
Av.Kasavubu*	1280	av.kasavubu.kak_a	4.3%
Av.Kavuye	702	av.nyakyoya.kak_a	8.9%
Av.Montngaliema	691		
Av.Nyakyoya	798		
Av.Sange	437		
Av.Kigongo	746	av.lukula.kak_a	4.3%
Av.Lukula	533		
Av.Kitundu*	1615	av.kitundu.kak_a	5.5%
Av.Kivu*	1669	av.kivu.kak_a	5.7%
Av.Lenghe_III	1883	av.lenghe_III.kak_a	9.2%
Av.Maendeleo	823		
Av.Likasi	608	av.likasi.kak_a	2.1%
Av.Membo*	710	av.membo.kak_a	2.4%
Av.Mukulima	620	av.mukulima.kak_a	2.1%
Av.Musulmane*	1010	av.musulmane.kak_a	3.4%
Av.Nakaziba	1113	av.nakaziba.kak_a	3.8%
Av.Nyatutwa*	1056	av.nyatutwa.kak_a	3.6%
Av.Reboisement	773	av.reboisement.kak_a	2.6%
Av.Shaba*	1626	av.shaba.kak_a	5.5%
Av.Simba*	933	av.simba.kak_a	3.2%
Av.Ubwari*	746	av.ubwari.kak_a	2.5%
Av.Ujuzi	876	av.ujuzi.kak_a	3.0%
Av.Virunga*	1493	av.virunga.kak_a	5.1%
Kalundu			
Av.Bongisa	1462	av.bongisa.kal_a	6.0%
Av.Centre_commercial*	1502	av.centre_commercial.kal_a	6.2%
Av.Du_port	955	av.du_port.kal_a	3.9%
Av.Kagenge	2308	av.kagenge.kal_a	9.5%
Av.Kakamba	2124	av.kakamba.kal_a	8.7%
Av.Kamongola	1703	av.kamongola.kal_a	7.0%
Av.Karigo	1137	av.karigo.kal_a	4.7%
Av.Kasia	969	av.kasia.kal_a	4.0%
Av.Kinogono_1	1324	av.kinogono.kal_a	11.1%
Av.Kinogono_2	1385		
Av.Mutarure	1519	av.mutarure.kal_a	6.2%
Av.Ngovi_mgja	1578	av.ngovi_mgja.kal_a	6.5%
Av.Nyoroka_1	1236	av.nyoroka.kal_a	11.1%
Av.Nyoroka_2	1472		
Av.Rugembe	1572	av.rugembe.kal_a	6.4%

Neighbourhood / Street	Census Nov. 2017	Analysis base unit (ABU)	Proportion of neighbourhood population
Av.Solange	1017	av.solange.kal_a	4.2%
Av.Umoja	1145	av.umoja.kal_a	4.7%
Kasenga			
Av.Amani	223	av.tupendane.kas_a	3.4%
Av.Kasablanca	180		
Av.Tupendane	315		
Av.Uwezo	205		
Av.Azuhuri*	944	av.azuhuri.kas_a	3.5%
Av.Bondogolo	740	av.kasenga_centre.kas_a	18.8%
Av.Hebroni	916		
Av.Kabomboza	1162		
Av.Kasenga_centre	850		
Av.Mapendo	1440		
Av.Budota	779	av.budota.kas_a	2.9%
Av.Bukavu*	822	av.bukavu.kas_a	3.0%
Av.Conforti	1323	av.mangondo.kas_a	10.3%
Av.Mangondo	1468		
Av.De_la_gombe*	558	av.de_la_gombe.kas_a	2.1%
Av.De_la_mosquee*	640	av.de_la_mosquee.kas_a	2.4%
Av.Du_general	628	av.du_general.kas_a	2.3%
Av.Du_petit_pont	355	av.kisangani.kas_a	5.1%
Av.Kisangani	1030		
Av.Du_port	874	av.du_port.kas_a	3.2%
Av.Ginki*	265	av.ginki.kas_a	1.0%
Av.Hewa_bora	678	av.hewa_bora.kas_a	2.5%
Av.Kigobe	230	av.kigobe.kas_a	0.8%
Av.Kikula*	732	av.kikula.kas_a	2.7%
Av.Kilima_hewa	718	av.musheru.kas_a	5.2%
Av.Musheru	684		
Av.Kitumaini	331	av.kitumaini.kas_a	1.2%
Av.Kivu_nord	551	av.kivu.kas_a	2.0%
Av.Kiyaya_ouest	843	av.kiyaya_ouest.kas_a	3.1%
Av.Lala_salama	686	av.lala_salama.kas_a	2.5%
Av.Maendeleo	501	av.maendeleo.kas_a	1.8%
Av.Majengo	479	av.majengo.kas_a	1.8%
Av.Muranvya	823	av.muranvya.kas_a	3.0%
Av.Muungano	478	av.muungano.kas_a	1.8%
Av.Mwangaza	559	av.mwangaza.kas_a	2.1%
Av.Regezamwendo*	541	av.regezamwendo.kas_a	2.0%
Av.Salongo*	507	av.salongo.kas_a	1.9%
Av.Tanganyika*	448	av.tanganyika.kas_a	1.7%
Av.Umoja	460	av.umoja.kas_a	1.7%
Av.Universite	1159	av.universite.kas_a	4.3%
Kavimvira			
Av.CPGL	1054	av.CPGL.kav_a	5.2%
Av.De_la_plage	978	av.de_la_plage.kav_a	4.8%
Av.Du_lac	1053	av.du_lac.kav_a	5.2%
Av.Du_marche	993	av.du_marche.kav_a	4.9%
Av.Du_projet*	1048	av.du_projet.kav_a	5.1%
Av.Kalembelembe	1137	av.kalembelembe.kav_a	5.6%
Av.Kamanyola	953	av.kamanyola.kav_a	4.7%
Av.Kasavubu	1400	av.kasavubu.kav_a	6.9%
Av.Kimbangu*	1349	av.kimbangu.kav_a	6.6%
Av.Lumumba	1038	av.lumumba.kav_a	5.1%
Av.Maiyamoto*	964	av.maiyamoto.kav_a	4.7%
Av.Mapendo	1203	av.mapendo.kav_a	5.9%
Av.Mobutu	1016	av.mobutu.kav_a	5.0%
Av.Ndava	1035	av.ndava.kav_a	5.1%
Av.Nyangara	936	av.nyangara.kav_a	4.6%
Av.Rond_point	1181	av.rond_point.kav_a	5.8%
Av.Rubenga	1019	av.rubenga.kav_a	5.0%
Av.Tanganyika	977	av.tanganyika.kav_a	4.8%
Av.Tupendane	1024	av.tupendane.kav_a	5.0%

Neighbourhood / Street	Census Nov. 2017	Analysis base unit (ABU)	Proportion of neighbourhood population
Kibondwe			
Av.Bajoba	833	av.bajoba.kib_a	15.4%
Av.Bushoga	677	av.bushoga.kib_a	12.5%
Av.Kabego	761	av.kabego.kib_a	14.0%
Av.Kitunya	1145	av.kitunya.kib_a	21.1%
Av.Mushule	1136	av.mushule.kib_a	21.0%
Av.Reboisement	866	av.reboisement.kib_a	16.0%
Kilibula			
Av.Kyonga	2704	av.kyonga.kil_a	27.6%
Av.Lenghe_III	1647	av.lenghe_III.kil_a	16.8%
Av.Maendeleo	1997	av.maendeleo.kil_a	20.4%
Av.Mombasa	1706	av.mombasa.kil_a	17.4%
Av.Shaba	1754	av.shaba.kil_a	17.9%
Kimanga			
Av.De_la_paix	2005	av.de_la_paix.kim_a	14.0%
Av.Du_pionnier	2090	av.du_pionnier.kim_a	14.6%
Av.Du_stade	2651	av.du_stade.kim_a	18.5%
Av.Kabungulu_1	2592	av.kabungulu.kim_a	30.7%
Av.Kabungulu_2	1815		
Av.Kimanga	1118	av.kimanga.kim_a	7.8%
Av.Nyamianda_1	895	av.nyamianda.kim_a	14.5%
Av.Nyamianda_2	1186		
Mulongwe			
Av.Apollo_1	1323	av.apollo.mul_a	9.7%
Av.Apollo_2	1088		
Av.De_la_cite	1314	av.de_la_cite.mul_a	5.3%
Av.Kasavubu	822	av.kasavubu.mul_a	3.3%
Av.Kayaja_1_2_3_4	4333	av.kayaja_1_2_3_4.mul_a	17.3%
Av.Kitunge	892	av.kitunge.mul_a	3.6%
Av.Lumumba	1336	av.lumumba.mul_a	5.3%
Av.Makobola	1151	av.makobola.mul_a	4.6%
Av.Matadi_1	981	av.matadi.mul_a	6.8%
Av.Matadi_2	717		
Av.Mitumba_1_2_3	3327	av.mitumba_1_2_3.mul_a	13.3%
Av.Rwegereza	963	av.rwegereza.mul_a	3.9%
Av.Shishi_1	1359	av.shishi.mul_a	22.9%
Av.Shishi_3	1243		
Av.Shishi_2	1410		
Av.Shishi_4	1699		
Av.Yohana	1020	av.yohana.mul_a	4.1%
Nyamianda			
Av.Alliance	1497	av.alliance.nya_a	10.5%
Av.Embouchure	1273	av.embouchure.nya_a	8.9%
Av.Goma	1308	av.goma.nya_a	9.1%
Av.Isiro	1712	av.isiro.nya_a	12.0%
Av.Kivu	1261	av.kivu.nya_a	8.8%
Av.Lumbulumbu	1383	av.lumbulumbu.nya_a	9.7%
Av.Munanira	1563	av.munanira.nya_a	10.9%
Av.Mundi	1563	av.mundi.nya_a	10.9%
Av.Musumba	1438	av.musumba.nya_a	10.0%
Av.Plage_d_or	1327	av.plage_d_or.nya_a	9.3%
Rombe 1			
Av.Bas_congo	1060	av.bas_congo.rb1_a	4.3%
Av.Bralima_1	865	av.bralima.rb1_a	11.3%
Av.Bralima_2	1887		
Av.De_l_authenticite	969	av.de_l_authenticite.rb1_a	4.0%
Av.Du_04_janvier	797	av.du_04_janvier.rb1_a	3.3%
Av.Du_15_decembre	1043	av.du_15_decembre.rb1_a	4.3%
Av.Du_24_novembre	831	av.du_24_novembre.rb1_a	3.4%
Av.Du_27_octobre	1007	av.du_27_octobre.rb1_a	4.1%
Av.Du_30_juin	1030	av.du_30_juin.rb1_a	4.2%
Av.Fac	900	av.fac.rb1_a	3.7%
Av.Fizi	1168	av.fizi.rb1_a	4.8%

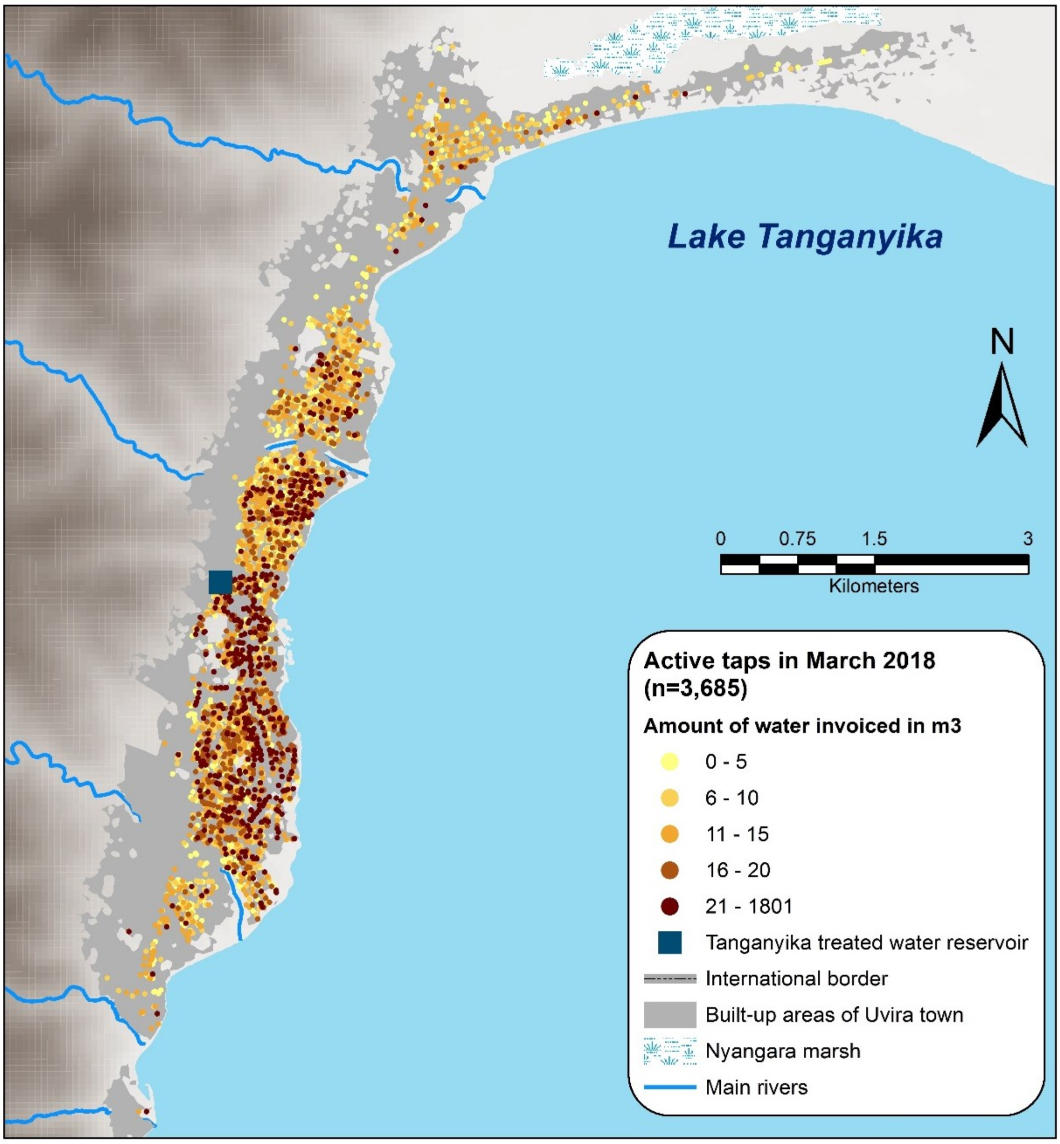
Neighbourhood / Street	Census Nov. 2017	Analysis base unit (ABU)	Proportion of neighbourhood population
Av.Haut_congo	1218	av.haut_congo.rb1_a	5.0%
Av.Kabare	783	av.kabare.rb1_a	3.2%
Av.Kakungwe_1	1423		
Av.Kakungwe_3	1589	av.kakungwe.rb1_a	18.3%
Av.Kakungwe_2	1457		
Av.Kalehe	775	av.kalehe.rb1_a	3.2%
Av.Major_vangu	857	av.major_vangu.rb1_a	3.5%
Av.Mulongwe	648	av.mulongwe.rb1_a	2.7%
Av.Munanira	298	av.munanira.rb1_a	1.2%
Av.Mwenga	968	av.mwenga.rb1_a	4.0%
Av.Shabunda	1029	av.shabunda.rb1_a	4.2%
Av.Uvira	900	av.uvira.rb1_a	3.7%
Av.Walungu	944	av.walungu.rb1_a	3.9%
Rombe 2			
Av.Bas_congo	462	av.bas_congo.rb2_a	2.7%
Av.De_l_authenticite	459	av.de_l_authenticite.rb2_a	2.7%
Av.Du_04_janvier	489	av.du_04_janvier.rb2_a	2.9%
Av.Du_15_decembre	701	av.du_15_decembre.rb2_a	4.1%
Av.Du_24_novembre	1189	av.du_24_novembre.rb2_a	6.9%
Av.Du_27_octobre	350	av.du_27_octobre.rb2_a	2.0%
Av.Du_30_juin	488	av.du_30_juin.rb2_a	2.8%
Av.Fac	598	av.fac.rb2_a	3.5%
Av.Fizi	645	av.fizi.rb2_a	3.8%
Av.Haut_congo	567	av.haut_congo.rb2_a	3.3%
Av.Idjwi_1	1224		
Av.Idjwi_2	990	av.idjwi.rb2_a	19.3%
Av.Idjwi_3	1091		
Av.Kabare	1123	av.kabare.rb2_a	6.5%
Av.Kalehe	1057	av.kalehe.rb2_a	6.2%
Av.Major_vangu	918	av.major_vangu.rb2_a	5.4%
Av.Mulongwe	837	av.mulongwe.rb2_a	4.9%
Av.Munanira	928	av.munanira.rb2_a	5.4%
Av.Mwenga	581	av.mwenga.rb2_a	3.4%
Av.Shabunda	516	av.shabunda.rb2_a	3.0%
Av.Uvira	1186	av.uvira.rb2_a	6.9%
Av.Walungu	751	av.walungu.rb2_a	4.4%
Rugenge			
Av.De_la_paroisse	856	av.de_la_paroisse.rug_a	12.2%
Av.Kinaga	716	av.kinaga.rug_a	10.2%
Av.Maendeleo	1241	av.maendeleo.rug_a	17.6%
Av.Makarunga	424	av.makarunga.rug_a	6.0%
Av.Petrocongo	829	av.petrocongo.rug_a	11.8%
Av.Rugenge_nord	1060	av.rugenge.rug_a	28.4%
Av.Rugenge_sud	939		
Av.Ushirika	970	av.ushirika.rug_a	13.8%
Songo			
Av.Alliance	1450	av.alliance.sg_a	7.4%
Av.Alpha	1645	av.alpha.sg_a	8.4%
Av.De_la_mission	1372	av.de_la_mission.sg_a	7.0%
Av.De_la_paix	1040	av.de_la_paix.sg_a	5.3%
Av.Democratie	1561	av.democratie.sg_a	8.0%
Av.Du_17_mai	1351	av.du_17_mai.sg_a	6.9%
Av.Du_marche	944	av.du_marche.sg_a	4.8%
Av.Du_progres	1132	av.du_progres.sg_a	5.8%
Av.Du_stade	1297	av.du_stade.sg_a	6.6%
Av.Lumumba	1363	av.lumumba.sg_a	7.0%
Av.Maendeleo	1469	av.maendeleo.sg_a	7.5%
Av.Mapendano	1203	av.mapendano.sg_a	6.1%
Av.Mapinduzi	1261	av.mapinduzi.sg_a	6.4%
Av.Matumaini	1479	av.matumaini.sg_a	7.6%
Av.Umoja	1000	av.umoja.sg_a	5.1%

** Streets divided across several aires de santé (AS)*

Appendix 2-4 Neighbourhoods' population data and estimates over time

	Neighbourhood census date (dd/mm/yyyy) Time index (days)										Forecasted estimates	
Neighbourhood	**01/07/2008** *1*	**01/07/2010** *731*	**15/02/2012** *1325*	**15/05/2012** *1415*	**15/02/2013** *1691*	**15/05/2014** *2145*	**15/11/2014** *2329*	**15/11/2015** *2694*	**15/05/2016** *2876*	**15/08/2017** *3333*	**15/03/2018** *3545*	**16/05/2019** *3972*
Kabindula	10,284	12,692	12,624	12,514	13,834	14,092	14,274	14,290	15,400	15,679	16,204	17,003
		11.7%	*-0.3%*	*-3.5%*	*13.9%*	*1.5%*	*2.6%*	*0.1%*	*15.6%*	*1.4%*		
Kakombe	18,888	24,965	24,702	24,377	24,678	24,661	24,636	24,148	29,399	29,435	31,523	34,380
		16.1%	*-0.6%*	*-5.3%*	*1.6%*	*-0.1%*	*-0.2%*	*-2.0%*	*43.6%*	*0.1%*		
Kalundu	17,205	23,965	23,899	23,966	23,394	23,483	23,577	23,589	23,999	24,341	24,608	25,075
		19.6%	*-0.2%*	*1.1%*	*-3.2%*	*0.3%*	*0.8%*	*0.1%*	*3.5%*	*1.1%*		
Kasenga	13,088	15,344	15,337	15,288	19,206	19,206	19,228	19,239	24,601	24,421	26,487	29,248
		8.6%	*0.0%*	*-1.3%*	*33.9%*	*0.0%*	*0.2%*	*0.1%*	*55.9%*	*-0.6%*		
Kavimvira	15,564	16,382	16,538	18,123	18,442	18,592	18,830	19,340	19,381	20,170	20,428	21,018
		2.6%	*0.6%*	*38.9%*	*2.3%*	*0.7%*	*2.5%*	*2.7%*	*0.4%*	*3.3%*		
Kibondwe	2,810	4,030	3,982	4,058	4,770	5,044	4,866	4,875	5,091	5,369	5,540	5,856
		21.7%	*-0.7%*	*7.7%*	*23.2%*	*4.6%*	*-7.0%*	*0.2%*	*8.9%*	*4.4%*		
Kilibula	7,903	8,678	8,615	8,833	9,064	9,197	9,261	9,272	9,519	9,744	9,911	10,206
		4.9%	*-0.4%*	*10.3%*	*3.5%*	*1.2%*	*1.4%*	*0.1%*	*5.3%*	*1.9%*		
Kimanga	11,149	12,951	12,961	13,110	13,509	13,564	13,620	13,639	14,057	14,328	14,576	14,997
		8.1%	*0.0%*	*4.7%*	*4.0%*	*0.3%*	*0.8%*	*0.1%*	*6.1%*	*1.5%*		
Mulongwe	16,426	23,213	23,213	22,234	22,748	22,748	22,988	22,998	25,668	26,226	27,453	29,294
		20.7%	*0.0%*	*-17.1%*	*3.1%*	*0.0%*	*2.1%*	*0.0%*	*23.3%*	*1.7%*		
Nyamianda	14,550	11,204	11,218	11,243	11,138	11,138	11,198	11,211	11,239	14,423	15,410	17,716
		-11.5%	*0.1%*	*0.9%*	*-1.2%*	*0.0%*	*1.1%*	*0.1%*	*0.5%*	*22.6%*		
Rombe I	18,801	17,455	17,452	18,220	21,167	21,551	20,783	20,793	23,135	24,146	25,382	27,372
		-3.6%	*0.0%*	*17.8%*	*21.4%*	*1.5%*	*-7.1%*	*0.0%*	*22.6%*	*3.5%*		
Rombe II	12,308	14,344	14,392	15,150	15,177	15,912	15,988	16,002	16,639	17,150	17,559	18,270
		8.3%	*0.2%*	*21.4%*	*0.2%*	*3.9%*	*0.9%*	*0.1%*	*8.0%*	*2.5%*		
Rugenge	5,532	6,419	6,404	6,379	6,177	6,198	6,156	6,173	6,355	7,024	7,301	7,881
		8.0%	*-0.1%*	*-1.6%*	*-4.2%*	*0.3%*	*-1.3%*	*0.3%*	*5.9%*	*8.4%*		
Songo	17,050	18,004	18,013	17,992	19,077	19,077	19,095	19,118	19,280	20,224	20,577	21,344
		2.8%	*0.0%*	*-0.5%*	*8.0%*	*0.0%*	*0.2%*	*0.1%*	*1.7%*	*3.9%*		

Appendix 2-5 Map of active Regideso taps in March 2018



Appendix 2-6 Informed consent form for household survey

Impact evaluation of water supply improvements in Uvira, South-Kivu, Democratic Republic of the Congo, on cholera, other diarrhoeal diseases and water-related practices

Information for participants

Households' practices related to collecting, storing, treating and using water, as well as hygiene practices, are key factors in the transmission of water-related diseases, especially cholera. Cholera is a disease that may lead to death that has been affecting Uvira for many years, despite many attempts by health authorities and actors to prevent it.

The London School of Hygiene and Tropical Medicine in the United Kingdom is studying households' practices related to collecting, storing, treating and using water, as well as hygiene practices, during the improvements in tap water supply in Uvira. This project is led by the Regideso and the Congolese government, with the support from the French Development Agency and the Veolia Foundation.

Study results will help understanding whether these tap water supply improvements have favoured households' water and hygiene practices less prone to the transmission of acute diarrhoeal diseases and cholera. Results will also strengthen our understanding of how these diseases are transmitted and will inform further improvements in tap water supply in Uvira and in other communities in the region that are also affected by cholera.

To participate to this study, investigators will ask questions on your household and on water and hygiene practices to an adult household member, and this interview will take about 1h. This interview may be repeated up to 4 times between September 2015 et November 2017. During this visit, the investigator will also collect a small amount of the drinking water stored in your house in order to check its microbiological quality.

Your participation to this study is entirely voluntary and you can decide to withdraw your participation at any time, especially if you do not wish to participate to future interviews. Interrupting or withdrawing your participation will have no adverse consequence on your household in any way.

If you have any concerns about the study, please feel free to ask the interviewer who will do his/her best to answer your questions. If he/she cannot do so, you can contact the study principal investigator (Oliver Cumming, oliver.cumming@lshtm.ac.uk, tel (UK): +44 207 636 8636) or the study coordinator (Aurelie Jeandron, aurelie.jeandron@lshtm.ac.uk, tel (UK): +44 207 927 2417).

Take the time to think about your participation and do not hesitate to ask any question before deciding.

This information is for you to keep. Thank you for considering taking the time to read this sheet.

Informed consent form

I read and/or understood the explanations given in to me about this study by the investigator named below, and he/she answered my questions. I understand that my participation on behalf of my household is entirely voluntary and that I can withdraw my participation at any time.

I agree to participate to the study entitled "Impact evaluation of water supply improvements in Uvira, South-Kivu, Democratic Republic of the Congo, on cholera, other diarrhoeal diseases and water-related practices" led by the London School of Hygiene and Tropical Medicine.

Participant's name:____

Participant's signature: ___________________________

Impartial witness name and signature, if the participant is unable to read and/or sign:

____

Investigator's name and signature: ____Date ______________

Household ID number: __________

Appendix 2-7 Informed consent form for RDT confirmation

Impact evaluation of water supply improvements in Uvira, South-Kivu, Democratic Republic of the Congo, on cholera, other diarrhoeal diseases and water-related practices

Information for participants

As a patient admitted to the Cholera Treatment Centre (CTC) in Uvira, or the parent or guardian of a child under 15 admitted to the CTC, you are invited to take part to a study by researchers of the London School of Hygiene and Tropical Medicine (LSHTM) and implemented by the Bureau Central de Zone (BCZ) of Uvira and the General Hospital of Uvira. This study was approved by the Ethics Committee of the LSHTM and by the School of Public Health at the University of Kinshasa. Before you decide to take part in the study, it is important for you to understand why the research is done and what is involved if you decide to participate. Please ask us if there is anything that is not clear and take the time to decide if you wish to participate or not.

Cholera is a severe diarrhoeal disease present in Uvira and the province of South-Kivu. People coming to the Cholera Treatment Centre (CTC) in Uvira with severe diarrhoea may have be affected by cholera, but also by other diarrhoeal diseases caused by other types of bacteria. In order to better understand the particularities of cholera compared to these other diarrhoeal diseases, it is important to know if the illness that brought you to the CTC is caused by cholera or by another bacteria. This research therefore aims at confirming if the severe diarrhoea in patients admitted to the CTC is due to cholera or to another bacteria. This research will invite all the patients admitted to the CTC in Uvira to participate until December 2018, so that we better understand cholera transmission during different seasons, and so that we can see if cholera is reduced by improving water supply in the town of Uvira.

If you accept to participate to this study, or accept for your child to participate in this study, we will collect a rectal swab from you or your child. This means inserting briefly a clean cotton-tip in the rectum, and then removing it, to collect a small portion of fecal matter. This will be done by a trained laboratory technician. You or your child may feel a little pressure when the swab is inserted but it is painless in most cases.

The swab will then be analysed in the laboratory to see if the diarrhoea was caused by cholera. It will also be shipped to the United Kingdom for further analyses. The results will be anonymous and will not be communicated to anyone. The results will not influence in any way the treatment you or your child will receive from the CTC, because this treatment is appropriate whatever the result of the test.

Along with the swab we will collect information on your age, gender, neighbourhood or village of residence (or those of your child) and if you (or you child) have taken antibiotics prior to the swab being collected. This information will be anonymous.

Your participation to this study is entirely voluntary. If you do not wish to participate, it will not affect in anyway your or your child's treatment at the CTC. It may however help understanding better how cholera is affecting people in Uvira and help decide how to better fight against cholera in Uvira and elsewhere. The results of the research will be communicated to the Ministry of Health in Kinshasa and other parties involved in the fight against cholera in DRC and in the rest of the world. You can decide to withdraw your or your child's participation at any time – if your child refuses to have the rectal swab taken or shows evident distress when prepared for the swab, we will not proceed.

If you have any concerns about the study, please feel free to ask the trained laboratory technician who will do his/her best to answer your questions. If he/she cannot do so, you can contact the study principal investigator (Oliver Cumming, oliver.cumming@lshtm.ac.uk, tel (UK): +44 207 636 8636) or the study coordinator (Aurelie Jeandron, aurelie.jeandron@lshtm.ac.uk, tel (UK): +44 207 927 2417).

This information is for you to keep. Thank you for considering taking the time to read this sheet.

Informed consent form

I read and/or understood the explanations given in to me about this study by the investigator named below, and he/she answered my questions. I understand that my / my child's participation *[highlight the appropriate statement]* **is entirely voluntary and that I can withdraw my / my child's** *[highlight the appropriate statement]* **participation at any time.**

I agree to participate to the study / for my child to participate to the study *[highlight the appropriate statement]* entitled "Impact evaluation of water supply improvements in Uvira, South-Kivu, Democratic Republic of the Congo, on cholera, other diarrhoeal diseases and water-related practices" led by the London School of Hygiene and Tropical Medicine.

Participant's name and study ID number:____

Participant's parent / legal guardian name if the participant is a child under 15:______________________________

Participant's / participant's parent signature *[highlight the appropriate statement]*____________________

Impartial witness name and signature, if the participant or participant's parent is unable to read and/or sign:

____

Investigator's name and signature: ____Date ________________

Appendix 4-1a Questionnaire on water-related practices – survey 2016 (French)

QUESTIONNAIRE SUR LES PRATIQUES LIEES A L'EAU DANS LES MÉNAGES			
ENQUETE A DOMICILE			
Identification du questionnaire et du ménage			
1.1	Identifiant de l'enquêteur	[__|__]	
1.2	Identifiant du point GPS cible	A[__|__|__]	
1.3	Confirmez que vous avez bien obtenu le consentement écrit de l'enquêté(e) de plus de 18 ans, et que vous lui avez remis une copie du consentement signé.	OK ❑	
1.4	Identifiant du WAYPOINT MENAGE	[__|__|__]	
Informations sur l'enquêté(e)			
2.1	Qui répond aux questions ? *(Plusieurs réponses possibles)*	Chef de ménage – femme (ou épouse du chef de ménage) ❑ Chef de ménage - homme ❑ Autre(s) membre(s) du ménage ❑ Domestique / ménager(e) ❑	
2.2	Quelle est votre adresse exacte ? *Enregistrez toutes les informations pertinentes pour retrouver le ménage sélectionné lors d'une visite ultérieure : couleur ou type de porte, type de murs, place du ménage dans la parcelle (par exemple "1er ménage à gauche"), position par rapport à un repère ("derrière la pharmacie xxx"), etc...*	Quartier *(liste)* Avenue *(liste)* Numéro [__|__|__] Informations supplémentaires : ________________________________	
Informations sur la démographie du ménage			
3.1	Combien de personnes vivent dans votre ménage et étaient présentes HIER ? *Un ménage est un groupe de personnes qui dorment sous le même toit et partagent la majorité des repas. Ce groupe de personnes reconnait l'autorité d'une même personne comme étant le ou la chef du ménage. Si le ménage est famille d'accueil pour réfugiés/déplacés, ne pas les compter comme membres du ménage.*	Enfants de moins de 2 ans [__|__] Enfants de 2 à 5 ans [__|__] Enfants de 5 à 15 ans [__|__] Adultes de plus de 15 ans [__|__] *Présent signifie que la personne s'est réveillée dans le logement et est revenue y dormir. Des raisons d'absences peuvent être par exemple d'être parti pêcher plusieurs jours, aux champs, en déplacement professionnel, chez des proches, à l'hôpital, etc...*	
Nombre total de membres du ménage présents HIER : [_] *(calcul automatique)*			
3.2	Accueillez-vous des réfugiés / déplacés dans votre ménage actuellement ? Si oui, combien ? *Seulement les réfugiés/déplacés "officiels". Ils ne doivent pas être comptés dans les membres du ménage à la question précédente.*	Enfants de moins de 2 ans [__|__] Enfants de 2 à 5 ans [__|__] Enfants de 5 à 15 ans [__|__] Adultes de plus de 15 ans [__|__]	
3.3	Combien y-a-t-il de ménages qui vivent dans votre parcelle, y compris le vôtre ? *(un ménage = une cuisine)*	[__|__] ménages *Une parcelle est un groupe d'habitations délimité et partageant le même espace extérieur (cour). S'il s'agit d'une maison isolée / individuelle, mettre 1 (le ménage interrogé). Un individu célibataire compte comme un ménage.*	
3.4	**Seulement si 3.3>1** Combien d'autres personnes au total vivent dans votre parcelle, en dehors de votre ménage ? *Saisissez 99 si ne sait pas.*	[__|__] personnes	

3.5	Depuis combien de temps est-ce que votre ménage habite ce logement ? *(Une seule réponse)*	Moins d'1 mois ---------- 1 Entre 1 et 6 mois ---------- 2 Entre 6 mois et 1 an ---------- 3 Entre 1 et 2 ans ---------- 4 Depuis plus de 2 ans ---------- 5 Ne sait pas ---------- 98 Ne souhaite pas repondre ---------- 99	
Informations sur la quantité d'eau puisée et utilisée dans le ménage			
Je vais maintenant vous poser des questions sur l'eau que vous puisez et utilisez dans votre ménage.			
4.1	Combien de fois par jour votre ménage puise-t-il l'eau habituellement ? *(Une seule réponse)*	Plus de 2 fois par jour ---------- 0 2 fois par jour ---------- 1 1 fois par jour ---------- 2 1 fois tous les 2 jours ---------- 3 1 fois tous les 3 jours ---------- 4 Moins d'une fois tous les 3 jours ---------- 5 Ne sait pas ---------- 98	
4.2	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour le LAVAGE DES ADULTES pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 ---------- [___] litres Volume 2 ---------- [___] litres Volume 3 ---------- [___] litres Volume 4 ---------- [___] litres Volume 5 ---------- [___] litres Volume 6 ---------- [___] litres Volume 7 ---------- [___] litres	
4.3	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour le LAVAGE DES ENFANTS (de 0 à 15 ANS) pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 ---------- [___] litres Volume 2 ---------- [___] litres Volume 3 ---------- [___] litres Volume 4 ---------- [___] litres Volume 5 ---------- [___] litres Volume 6 ---------- [___] litres Volume 7 ---------- [___] litres	
4.4	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour le NETTOYAGE DU LOGEMENT (Y COMPRIS LATRINE) pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 ---------- [___] litres Volume 2 ---------- [___] litres Volume 3 ---------- [___] litres Volume 4 ---------- [___] litres Volume 5 ---------- [___] litres Volume 6 ---------- [___] litres Volume 7 ---------- [___] litres	
4.5	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour la LESSIVE DES VÊTEMENTS DES ENFANTS /ADULTES pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 ---------- [___] litres Volume 2 ---------- [___] litres Volume 3 ---------- [___] litres Volume 4 ---------- [___] litres Volume 5 ---------- [___] litres Volume 6 ---------- [___] litres Volume 7 ---------- [___] litres	
4.6	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour LA BOISSON ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 ---------- [___] litres Volume 2 ---------- [___] litres Volume 3 ---------- [___] litres Volume 4 ---------- [___] litres Volume 5 ---------- [___] litres Volume 6 ---------- [___] litres Volume 7 ---------- [___] litres	
4.7	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour la PRÉPARATION DE LA NOURRITURE OU LE LAVAGE DES PROVISIONS pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 ---------- [___] litres Volume 2 ---------- [___] litres Volume 3 ---------- [___] litres Volume 4 ---------- [___] litres Volume 5 ---------- [___] litres Volume 6 ---------- [___] litres Volume 7 ---------- [___] litres	

4.8	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour la VAISSELLE pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 [___] litres Volume 2 [___] litres Volume 3 [___] litres Volume 4 [___] litres Volume 5 [___] litres Volume 6 [___] litres Volume 7 [___] litres	
4.9	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour le LAVAGE DES MAINS pour votre ménage ? *Utiliser l'outil visuel pour aider à l'estimation – bien enregistrer l'eau utilisée au domicile, et non celle utilisée à la source, PAR TOUS LES MEMBRES DU MENAGE.*	Volume 1 [___] litres Volume 2 [___] litres Volume 3 [___] litres Volume 4 [___] litres Volume 5 [___] litres Volume 6 [___] litres Volume 7 [___] litres	
4.10	Combien d'eau « neuve » (puisée, pas ré-utilisée après une autre activité) estimez-vous avoir utilisé HIER dans votre ménage, à votre domicile, pour une ACTIVITÉ ÉCONOMIQUE (POUR UN REVENU) pour votre ménage ? *Exemples d'activités économiques :* *Préparation de nourriture/boissons/eau/glace pour vendre, arrosage d'un jardin potager/des cultures, abreuvage des animaux d'élevage, lavage du poisson/ légumes / fruits pour vendre, lessive/nettoyage pour d'autres ménages, préparation de matériaux de construction*	Volume 1 [___] litres Volume 2 [___] litres Volume 3 [___] litres Volume 4 [___] litres Volume 5 [___] litres Volume 6 [___] litres Volume 7 [___] litres	
4.11	Je vais résumer vos réponses sur les quantités d'eau neuve utilisée pour les différentes activités. Est-ce que cela vous semble correct ? *Si la quantité parait trop faible ou trop élevée, insistez auprès de l'enquêté(e) pour vérifier les chiffres. Vous pouvez corriger en revenant en arrière.* *Le nombre de litres enregistrés pour les questions précédentes seront automatiquement calculés par la tablette ou le téléphone.*	Lavage des adultes [ ] litres Lavage des enfants de 0 à 15 ans [ ] litres Nettoyage du logement (y compris latrine) [ ] litres Lavage des vêtements des adultes et des enfants [ ] litres Boisson [ ] litres Préparation de la nourriture/lavage des provisions [ ] litres Vaisselle [ ] litres Lavage des mains [ ] litres Activité économique [ ] litres TOTAL d'eau neuve utilisée HIER [ ] litres Soit en jerricans de 20L [ ] litres	
4.12	Avez-vous (ou les autres membres de votre ménage) ré-utilisé HIER de l'eau pour les activités suivantes ? *(Plusieurs réponses possibles)*	1. Préparation de la nourriture / Lavage des provisions pour votre ménage ❑ 2. Vaisselle ❑ 3. Lavage des adultes ❑ 4. Lavage des enfants (entre 0 et 15 ans) ❑ 5. Nettoyage du logement (y compris latrine) ❑ 6. Lessive des vêtements des enfants /adultes ❑ 7. Lavage des mains ❑ 8. Activité économique (pour un revenu) ❑ 9. Pas de ré-utilisation de l'eau ❑	
4.13	Avez-vous (ou les autres membres de votre ménage) effectué HIER les activités suivantes à une source située hors de votre parcelle ? *(Plusieurs réponses possibles)*	1. Préparation de la nourriture / Lavage des provisions pour votre ménage ❑ 2. Vaisselle ❑ 3. Lavage des adultes ❑ 4. Lavage des enfants (0 à 15 ans) ❑ 5. Lessive des vêtements des enfants /adultes ❑ 6. Activité économique (pour un revenu) ❑ 7. Pas d'utilisation à la source ❑	
Informations sur les lieux d'approvisionnement en eau			
Je vais maintenant vous poser des questions sur les endroits où vous puisez l'eau que vous utilisez dans votre ménage.			
Informations sur le lieu principal d'approvisionnement en eau			

5.1	Quel est le lieu où vous avez le plus puisé l'eau en quantité pour votre ménage, pour tous les usages, les DEUX DERNIERES SEMAINES ? *Attention, il s'agit D'UN LIEU DE PUISAGE et non de SOURCE : un ménage peut puiser son eau dans la rivière Kavimvira (source) à différents endroits (av. Musheru ou av. Muranvya par exemple)*	Lac Tanganyika ---------- 1 Etang Nyangara ---------- 2 Rivière Kavimvira ---------- 3 Rivière Mulongwe ---------- 4 Rivière Kalimabenge ---------- 5 Rivière Mai ya moto ---------- 6 Rivière Kamongola ---------- 7 Rivière Karigo ---------- 8 Autre cours d'eau ---------- 9 Robinet hors de votre parcelle ---------- 10 Robinet dans votre parcelle à l'extérieur de la maison ---------- 11 Robinet dans votre parcelle à l'intérieur de la maison ---------- 12 Canalisation Regideso percée ---------- 13 Eau de pluie ---------- 14	
5.2	Dans quel quartier se trouve ce lieu de puisage ?	Liste	
5.3	Dans quelle avenue se trouve ce lieu de puisage ?	Liste	
Utilisations du lieu principal d'approvisionnement en eau			
Je vais maintenant vous poser des questions sur ce lieu de puisage que vous avez le plus utilisé ces 2 dernières semaines.			
6.1	Depuis combien de temps utilisez-vous cet endroit pour vous approvisionner en eau dans votre ménage ?	Moins d'1 mois ---------- 1 Entre 1 et 6 mois ---------- 2 Entre 6 mois et 1 an ---------- 3 Entre 1 et 2 ans ---------- 4 Depuis plus de 2 ans ---------- 5 Depuis que nous habitons ici ---------- 6 Ne sait pas ---------- 98 Ne souhaite pas répondre ---------- 99	
6.2	Quand avez-vous utilisé ce lieu de puisage la dernière fois ?	Hier / Aujourd'hui ---------- 1 Il y a moins de 3 jours ---------- 2 Il y a moins d'une semaine ---------- 3 Il y a plus d'une semaine ---------- 4 Ne sait pas ---------- 98	
6.3	La dernière fois que vous avez puisé de l'eau à cette source, est-ce que votre ménage a utilisé cette eau pour boire ?	❑ OUI ❑ NON ❑ Ne sait pas	
6.4	La dernière fois que vous avez puisé de l'eau à cette source, est-ce que votre ménage a utilisé cette eau pour une activité économique (génératrice de revenus) ? *Exemples d'activités économiques :* *Préparation de nourriture/boissons/eau/glace pour vendre, arrosage d'un jardin potager/des cultures, abreuvage des animaux d'élevage, lavage du poisson/ légumes / fruits pour vendre, lessive/nettoyage pour d'autres ménages, préparation de matériaux de construction*	❑ OUI ❑ NON ❑ Ne sait pas	
Accessibilité du lieu principal d'approvisionnement en eau			
SI LIEU PRINCIPAL HORS DE LA PARCELLE (toutes options sauf 11-12-14 en question 5.1) :			
6.5	La dernière fois que vous avez-utilisé l'eau de cet endroit pour votre ménage, qui est-allé la puiser ? *Les questions suivantes se réfèrent au lieu de puisage PRINCIPAL. Plusieurs réponses possibles.*	Chef de ménage homme ---------- ❑ Chef de ménage femme ---------- ❑ Autre membre adulte homme du ménage ---------- ❑ Autre membre adulte femme du ménage ---------- ❑ Garçon entre 10 et 15 ans ---------- ❑ Fille entre 10 et 15 ans ---------- ❑ Garçon de moins de 10 ans ---------- ❑ Fille de moins de 10 ans ---------- ❑ Adulte ou enfant qui n'est pas membre du ménage ---------- ❑ Puiseur d'eau ---------- ❑ Ne sait pas ---------- ❑ Ne souhaite pas répondre ---------- ❑	

6.6	***Sauf si utilisation puiseur/ne sait pas/ne souhaite pas répondre (options 9-10-11 en 6.5)*** Combien de temps cela prend-il pour se rendre à cette source depuis votre logement ? Et pour en revenir avec l'eau puisée ? *(Bien préciser «sans escale » - sans s'arrêter en chemin pour se promener, faire une course, etc)*	Aller [__|__|__] min Retour [__|__|__] min *Saisissez 999 si ne sait pas.*	
6.7	***Sauf si utilisation puiseur/ne sait pas/ne souhaite pas répondre (options 9-10-11 en 6.5)*** Comment transportez-vous l'eau puisée ?	A pied ---------- 0 A bicyclette ---------- 1 Avec un véhicule motorisé ---------- 2	
SI LIEU PRINCIPAL EST UN ROBINET HORS DE LA PARCELLE (option 10 en question 5.1) :			
6.8	***Sauf si utilisation puiseur/ne sait pas/ne souhaite pas répondre (options 9-10-11 en 6.5)*** Combien de temps la personne qui va puiser l'eau pour votre ménage fait-elle habituellement la queue à cette source ?	Temps d'attente en moyenne ---------- [__|__|__] min	
SI LIEU PRINCIPAL N'EST PAS UN ROBINET (toutes options sauf 10-11-12 en question 5.1) :			
6.9	Est-ce que vous avez rencontré les problèmes suivants en puisant l'eau à cet endroit dans LES DEUX SEMAINES PASSEES ?	L'eau puisée est colorée ou turbide/sale ---------- OUI / NON L'eau puisée a un mauvais goût ---------- OUI / NON L'eau puisée a une mauvaise odeur ---------- OUI / NON	
SI LIEU PRINCIPAL EST UN ROBINET (options 10-11-12 en question 5.1) :			
6.10	Est-ce que vous avez rencontré les problèmes suivants en puisant l'eau à cet endroit dans LES DEUX SEMAINES PASSEES ?	L'eau puisée est colorée ou turbide/sale ---------- OUI / NON L'eau puisée a un mauvais goût ---------- OUI / NON L'eau puisée a une mauvaise odeur ---------- OUI / NON Manque de débit / de pression ---------- OUI / NON	
6.11	LA SEMAINE DERNIERE combien de jours y-a-t-il eu de l'eau disponible à ce robinet DANS LA JOURNEE (de 6h à 21h) ?	[__] jours	
6.12	LA SEMAINE DERNIERE combien de jours y-a-t-il eu de l'eau disponible à ce robinet DANS LA NUIT (de 21h à 6h) ?	[__] jours	
6.13	Dans les 2 semaines passées, est-ce qu'il vous est arrivé d'aller à ce robinet pour puiser, et il n'y avait pas d'eau ?	❑ OUI ❑ NON ❑ Ne sait pas	
6.14	Dans les deux semaines passées, avez-vous (ou quelqu'un de votre ménage) puisé l'eau à ce robinet entre 21h et 6h du matin ?	❑ OUI ❑ NON ❑ Ne sait pas	
Coût de l'eau puisée à la SOURCE PRINCIPALE			
SI LIEU PRINCIPAL EST UN ROBINET (options 10-11-12 en question 5.1) :			
6.15	Qui est le propriétaire ou l'abonné Regideso de ce robinet ?	Votre propre ménage ---------- 1 Le propriétaire de votre logement ---------- 2 Un membre de la famille dans un autre logement ---------- 3 Un autre ménage ---------- 4 Robinet public / communautaire / ONG ---------- 5 L'employeur d'un membre de votre ménage ---------- 6 Ne sait pas ---------- 0	
6.16	Depuis combien de temps ce robinet a-t-il été installé ?	Moins d'1 mois ---------- 1 Entre 1 et 6 mois ---------- 2 Entre 6 mois et 1 an ---------- 3 Entre 1 et 2 ans ---------- 4 Depuis plus de 2 ans ---------- 5 Depuis que nous habitons ici ---------- 6 Ne sait pas ---------- 98	
6.17	A votre connaissance, est-ce qu'il y a un compteur fonctionnel sur ce robinet ?	❑ OUI ❑ NON ❑ Ne sait pas	
SI LIEU PRINCIPAL EST UN ROBINET DONT LE MENAGE N'EST PAS L'ABONNE REGIDESO (options 10-11-12 en question 5.1 et 1-2-3-4 en question 6.15) :			
6.18	Est-ce que votre ménage paie pour l'approvisionnement en eau à cette source, pour le puisage et/ou le transport ? Payez-vous par mois ou par quantité ?	Ne paie pas pour l'eau de cette source ---------- 0 Troc / paiement en nature ---------- 1 Paiement par mois ---------- 2 Paiement par bidon de 20L ---------- 3	

6.19	***Seulement si paiement par mois (option 1 en 6.18)*** La dernière fois que vous avez payé, combien avez-vous payé pour le mois pour l'approvisionnement en eau à ce robinet pour votre ménage ?	[__|__|__|__|__|__|__] FC/mois *Si ne sait pas, mettre 999 999.*	
6.20	***Seulement si paiement par mois (option 1 en 6.18)*** Dans l'année passée, vous est-il arrivé d'avoir à étaler le paiement de l'eau utilisée sur plusieurs mois ou à retarder votre paiement de l'eau utilisée car vous n'aviez pas assez d'argent ?	❑ OUI ❑ NON ❑ Ne sait pas	
6.21	***Seulement si paiement par 20L (option 2 en 6.18)*** Combien avez-vous payé la dernière fois pour un container de 20 L pour l'approvisionnement en eau à ce robinet pour votre ménage ?	[__|__|__|__|__] FC / 20 L *Si ne sait pas mettre 99 999.*	
6.22	***Seulement si paiement par 20L (option 2 en 6.18)*** Le mois passé, vous est-il arrivé de manquer d'argent pour acheter un container de 20L d'eau et d'utiliser une source d'eau gratuite à la place ?	❑ OUI ❑ NON ❑ Ne sait pas	
SI LIEU PRINCIPAL EST UN ROBINET DANS LA PARCELLE DU MENAGE (options 11-12 en question 5.1) :			
6.23	Combien d'autres ménages que le vôtre se sont approvisionnés à ce robinet dans votre parcelle le mois passé ?	[__|__] *Si le robinet est uniquement utilisé par le ménage interrogé, mettre 0. Si ne sait pas, mettre 99.*	
SI LIEU PRINCIPAL EST UN ROBINET DANS LA PARCELLE DU MENAGE ET LE MENAGE EST L'ABONNE DE LA REGIDESO (options 11-12 en question 5.1 ET 0 en question 6.15) :			
6.24	Le mois passé, quel type de tarif la Regideso a-t-elle appliquée à votre abonnement ?	Au forfait ---------- 0 Au compteur ---------- 1	
6.25	Le mois passé, à combien s'est monté votre facture de la Regideso au total ?	[__|__|__|__|__|__|__] FC/mois *Si ne sait pas mettre 999 999.*	
6.26	Le mois passé, combien de ménages vous ont-payé pour s'approvisionner à votre robinet ?	[__|__] *Attention, bien compter les ménages qui PAIENT, pas ceux qui utilisent sans payer. Si aucun ménage ne paie, mettre 0, si ne sait pas mettre 99.*	
6.27	***Seulement si d'autres ménages ont payé pour s'approvisionner au robinet (>0 en question 6.26)*** Le mois passé, comment avez-vous facturé la consommation des autres ménages s'approvisionnant à votre robinet ?	6.27.1 Par utilisation ---------- OUI / NON 6.27.2 Par semaine ---------- OUI / NON 6.27.3 Par mois ---------- OUI / NON 6.27.4 Par bidon de 20 L ---------- OUI / NON 6.27.5 En troc / en nature ---------- OUI / NON	
6.28	***Seulement si d'autres ménages ont payé PAR UTILISATION pour s'approvisionner au robinet (OUI en question 6.27.1)*** Le mois passé, combien vous a payé chaque ménage payant par utilisation pour l'eau puisée à votre robinet ?	[__|__|__|__|__|__] FC/ménage par utilisation	
6.29	***Seulement si d'autres ménages ont payé PAR SEMAINE pour s'approvisionner au robinet (OUI en question 6.27.2)*** Le mois passé, combien vous a payé chaque ménage payant par semaine pour l'eau puisée à votre robinet ?	[__|__|__|__|__|__] FC/ménage par semaine	
6.30	***Seulement si d'autres ménages ont payé PAR MOIS pour s'approvisionner au robinet (OUI en question 6.27.3)*** Le mois passé, combien vous a payé chaque ménage payant par mois pour l'eau puisée à votre robinet ?	[__|__|__|__|__|__] FC/ménage par mois	
6.31	***Seulement si d'autres ménages ont payé PAR BIDON DE 20L pour s'approvisionner au robinet (OUI en question 6.27.4)*** Le mois passé, combien vous a payé chaque ménage payant par bidon de 20L pour l'eau puisée à votre robinet ?	[__|__|__|__|__|__] FC/ménage par bidon de 20L	

6.32	Est-ce que votre ménage a déjà fait des réclamations auprès de la Regideso sur votre facturation ?	❑ OUI ❑ NON ❑ Ne sait pas	
6.33	***Seulement si le ménage a déjà fait une réclamation (OUI en 6.32)*** Quand votre ménage a-t-il fait la dernière réclamation ?	Moins d'1 mois_ _ _ 1 Entre 1 et 6 mois_ _ _ 2 Entre 6 mois et 1 an_ _ _ 3 Entre 1 et 2 ans_ _ _ 4 Depuis plus de 2 ans_ _ _ 5 Ne sait pas_ _ _ 99	
6.34	***Seulement si le ménage a déjà fait une réclamation (OUI en 6.32)*** Auprès de qui exactement votre ménage a-t-il fait la dernière réclamation ?	Le releveur de compteur_ _ _ 1 Le bureau de la Regideso_ _ _ 2 Le chef d'avenue / de quartier_ _ _ 3 Ne sait pas_ _ _ 9 Autre (préciser)_ _ _ 4 ____________________	
SI LIEU PRINCIPAL N'EST PAS UN ROBINET DANS LA PARCELLE DU MENAGE (toutes options sauf 11-12 en question 5.1) :			
6.35	Il n'y a pas de robinet fonctionnel de la Regideso dans votre parcelle ou vous n'utilisez pas ce robinet comme lieu principal de puisage. Quelle est la raison principale pour celà d'après vous ? *Ne pas suggérer mais sélectionner la réponse qui correspond dans la liste ou choisir autre et préciser. (Une seule réponse)*	Il n'y a pas eu d'eau à ce robinet dans la parcelle depuis deux semaines_ _ _ 12 L'installation du robinet/ le raccordement coûte trop cher_ _ _ 1 Les factures d'eau pour l'abonné sont trop chères_ _ _ 2 Le propriétaire de la maison / de la parcelle ne souhaite pas installer de robinet (ménage locataire)_ _ _ 3 La Regideso ne peut pas installer de robinet ici pour des raisons techniques_ _ _ 4 La disponibilité de l'eau aux robinets n'est pas assez bonne_ _ _ 5 Notre ménage n'a pas besoin d'un robinet dans la parcelle_ _ _ 6 Nous venons d'arriver, nous n'avons pas eu le temps de faire la demande_ _ _ 7 Nous ne comptons pas rester dans cette habitation_ _ _ 8 Nous n'arrivons pas à nous mettre d'accord avec les voisins _ 9 Autre raison pas dans la liste ci-dessus (préciser)_ _ _ 10 ____________________ Pas de raison particulière / Ne sait pas_ _ _ 11	
6.36	Est-ce qu'il y a une deuxième raison ? *Ne pas suggérer mais sélectionner la réponse qui correspond dans la liste ou choisir autre et préciser. (Une seule réponse)*	L'installation du robinet/ le raccordement coûte trop cher_ _ 1 Les factures d'eau pour l'abonné sont trop chères_ _ _ 2 Le propriétaire de la maison / de la parcelle ne souhaite pas installer de robinet (ménage locataire)_ _ _ 3 La Regideso ne peut pas installer de robinet ici pour des raisons techniques_ _ _ 4 La disponibilité de l'eau aux robinets n'est pas assez bonne_ 5 Notre ménage n'a pas besoin d'un robinet dans la parcelle_ _ 6 Nous venons d'arriver, nous n'avons pas eu le temps de faire la demande_ _ _ 7 Nous ne comptons pas rester dans cette habitation_ _ _ 8 Nous n'arrivons pas à nous mettre d'accord avec les voisins 9 Autre raison pas dans la liste ci-dessus _ _ _ 10 ____________________ Pas de 2ème raison_ _ _ 11	
6.37	Est-ce qu'il y a une troisième raison ? *Ne pas suggérer mais sélectionner la réponse qui correspond dans la liste ou choisir autre et préciser. (Une seule réponse)*	L'installation du robinet/ le raccordement coûte trop cher_ _ 1 Les factures d'eau pour l'abonné sont trop chères_ _ _ 2 Le propriétaire de la maison / de la parcelle ne souhaite pas installer de robinet (ménage locataire)_ _ _ 3 La Regideso ne peut pas installer de robinet ici pour des raisons techniques_ _ _ 4 La disponibilité de l'eau aux robinets n'est pas assez bonne_ 5 Notre ménage n'a pas besoin d'un robinet dans la parcelle_ _ 6 Nous venons d'arriver, nous n'avons pas eu le temps de faire la demande_ _ _ 7 Nous ne comptons pas rester dans cette habitation_ _ _ 8 Nous n'arrivons pas à nous mettre d'accord avec les voisins 9 Autre raison pas dans la liste ci-dessus _ _ _ 10 ____________________ Pas de 3ème raison_ _ _ 11	
SI LIEU PRINCIPAL N'EST PAS UN ROBINET (toutes options sauf 10-11-12 en question 5.1) :			
6.38	Est-ce que votre ménage paie pour l'approvisionnement en eau à cette source, pour le puisage et/ou le transport ? Payez-vous par mois ou par quantité ?	Ne paie pas pour l'eau de cette source_ _ _ 0 Troc / paiement en nature _ _ _ 1 Paiement par mois _ _ _ 2 Paiement par bidon de 20L_ _ _ 3	

Informations sur le LIEU ALTERNATIF 1 d'approvisionnement en eau			
Merci beaucoup pour ces informations sur le lieu de puisage que vous avez le plus utilisé dans les 2 dernières semaines.			
7.1	Quel est le deuxième lieu où vous avez le plus puisé l'eau en quantité pour votre ménage, pour tous les usages, les DEUX DERNIERES SEMAINES ? *Attention, il s'agit D'UN LIEU DE PUISAGE et non de SOURCE : un ménage peut puiser son eau dans la rivière Kavimvira (source) à différents endroits (av. Musheru ou av. Muranvya par exemple)*	Pas de source alternative 0 Lac Tanganyika 1 Etang Nyangara 2 Rivière Kavimvira 3 Rivière Mulongwe 4 Rivière Kalimabenge 5 Rivière Mai ya moto 6 Rivière Kamongola 7 Rivière Karigo 8 Autre cours d'eau 9 Robinet hors de votre parcelle 10 Robinet dans votre parcelle à l'extérieur de la maison 11 Robinet dans votre parcelle à l'intérieur de la maison 12 Canalisation Regideso percée 13 Eau de pluie 14	
7.2	Dans quel quartier se trouve ce lieu de puisage ?	Liste	
7.3	Dans quelle avenue se trouve ce lieu de puisage ?	Liste	
REPETITION DES QUESTIONS 6.1 à 6.38 POUR LIEU ALTERNATIF 1 d'approvisionnement en eau			
REPETITION DES QUESTIONS 6.1 à 6.38 POUR LIEU ALTERNATIF 2 d'approvisionnement en eau			
8.1	En dehors des sources dont nous avons déjà parlé, est-ce qu'il y a d'autres sources d'eau que vous avez utilisées LE MOIS PASSE ? *(Plusieurs réponses possibles - Bien préciser que cette source peut-être pour tous les usages – boisson et/ou usages domestiques)*	Pas de d'autre source utilisée le mois passé ❑ Lac Tanganyika ❑ Etang Nyangara ❑ Rivière Kavimvira ❑ Rivière Mulongwe ❑ Rivière Kalimabenge ❑ Rivière Mai ya moto ❑ Rivière Kamongola ❑ Rivière Karigo ❑ Autre cours d'eau ❑ Robinet hors de votre parcelle ❑ Robinet dans votre parcelle à l'extérieur de la maison ❑ Robinet dans votre parcelle à l'intérieur de la maison ❑ Canalisation Regideso percée ❑ Eau de pluie ❑	
Nous avons bientôt terminé le questionnaire, merci beaucoup pour toutes vos réponses sur les lieux où vous puisé l'eau pour votre ménage. Je vais maintenant vous poser des questions sur les méthodes d'assainissement de l'eau et le stockage de l'eau dans votre ménage.			
Informations sur le traitement de l'eau de boisson			
9.1	Avez-vous déjà utilisé les méthodes suivantes pour assainir l'eau de boisson dans votre ménage ?	Faire reposer l'eau OUI / NON Faire bouillir et laisser refroidir OUI / NON Filtrer avec un linge propre OUI / NON Filtrer avec un filtre céramique OUI / NON Chlore au point de chloration OUI / NON Chlore acheté dans le commerce OUI / NON Autre méthode (précisez) OUI / NON ____________________	
9.2	Pouvez-vous citer le nom d'un ou plusieurs produits de traitement de l'eau que l'on peut trouver dans le commerce à Uvira ? *Ne pas suggérer mais sélectionner la ou les réponses dans la liste ou choisir autre et préciser. (Plusieurs réponses possibles)*	Ne connait pas de produit de traitement de l'eau ❑ Uzima ❑ Aquatabs ❑ Pur ❑ Autre (précisez) ❑ ____________________	
9.3	***(Seulement si au moins 1 OUI à la question 9.1)*** Quand avez-vous assaini votre eau de boisson la dernière fois ?	Hier / Aujourd'hui 1 Il y a moins de 3 jours 2 Il y a moins d'une semaine 3 Il y a plus d'une semaine 4 Ne sait pas 98	

9.4	*(Seulement si au moins 1 OUI à la question 9.1)* La dernière fois que vous avez traité l'eau de boisson pour votre ménage, était-ce dans les situations suivantes ? *(Plusieurs réponses possibles)*	Eau puisée sale/turbide ❑ Eau puisée à une source inhabituelle ❑ Adulte malade dans le ménage ❑ Enfant malade dans le ménage ❑ Choléra dans le voisinage/épidémie ❑ Distribution gratuite de produit ❑ Point de chloration au lieu de puisage ❑ Par habitude / parce que je le fais toujours ❑ Pas de raison particulière / ne sait pas ❑	
9.5	*(Seulement si au moins 1 OUI à la question 9.1)* Si vous assainissez un container d'eau pour boire, est-ce vous utilisez aussi cette eau assainie pour d'autres activités ? *(Plusieurs réponses possibles)*	1. Préparation de la nourriture / Lavage des provisions pour votre ménage ❑ 2. Vaisselle ❑ 3. Lavage des adultes ❑ 4. Lavage des enfants (entre 0 et 15 ans) ❑ 5. Nettoyage du logement (y compris latrine) ❑ 6. Lessive des vêtements des enfants /adultes ❑ 7. Lavage des mains ❑ 8. Activité économique (pour un revenu) ❑ 9. Pas d'autre utilisation de l'eau assainie ❑	
Informations sur le stockage de l'eau			
10.1	Si vous remplissez tous les containers que vous possédez, combien d'eau au total pouvez-vous avoir à la maison ? *Eau pour tous les usages*	Nombre de bidons/containers de 30L [___] Nombre de bidons/containers de 20L [___] Nombre de bidons/containers de 10L [___] Nombre de bidons/containers de 5L [___] Volume dans d'autres containers (en litres) [___] litres Volume dans d'autres containers (en litres) [___] litres Volume dans d'autres containers (en litres) [___] litres	
10.2 Quantité totale maximale de stockage [___] litres *(automatiquement calculé)*			
Prélèvement d'un échantillon d'eau de boisson			
11.0	Est-ce que je peux voir comment vous stockez votre eau de boisson dans votre ménage ?	❑ OUI ❑ NON ❑ Déclare ne pas avoir d'eau de boisson dans le ménage	
11.0a	*seulement si 11.0=Oui* OBSERVATION Comment est stockée l'eau de boisson dans le plus grand de ces containers ?	Le récipient est fermé / recouvert VRAI/FAUX Le récipient est surélevé VRAI/FAUX Il n'y a pas d'ouverture assez grande pour que la main touche l'eau VRAI/FAUX	
11.1	Acceptez-vous de me servir un (deux) verres d'eau de boisson comme vous le feriez pour vous-même ou un membre de votre ménage, pour que nous analysions sa qualité ?	❑ OUI ❑ NON ❑ Déclare ne pas avoir d'eau de boisson dans le ménage	
11.2	*seulement si 11.1=Oui* Vous souvenez-vous quand cette eau a été puisée ?	Aujourd'hui 1 Hier 2 Avant-hier 3 Il y a plus de deux jours 4 Ne sait pas 5	
11.3a	*seulement si 11.1=Oui* Où a-t-elle été puisée ?	Ne sait pas 0 Lac Tanganyika 1 Etang Nyangara 2 Rivière Kavimvira 3 Rivière Mulongwe 4 Rivière Kalimabenge 5 Rivière Mai ya moto 6 Rivière Kamongola 7 Rivière Karigo 8 Autre cours d'eau 9 Robinet hors de votre parcelle 10 Robinet dans votre parcelle à l'extérieur de la maison 11 Robinet dans votre parcelle à l'intérieur de la maison 12 Canalisation Regideso percée 13 Eau de pluie 14	
11.3 b	*seulement si 11.1=Oui* Dans quel quartier ?	Liste	
11.3c	*seulement si 11.1=Oui* Dans quelle avenue ?	Liste	

	seulement si 11.1=Oui OBSERVATION Depuis quel type de récipient l'échantillon a-t-il été prélevé ?	Le récipient est fermé / recouvert _VRAI/FAUX/PAS OBSERVE Le récipient est surélevé ______VRAI/FAUX/PAS OBSERVE Il n'y a pas d'ouverture assez grande pour que la main touche l'eau ________________________VRAI/FAUX/PAS OBSERVE	
	seulement si 11.1=Oui OBSERVATION Comment le répondant a-t-il/elle servi l'eau dans le sac ? *(une seule réponse)*	En plongeant un ustensile (tasse, cruche…)_________________ 1 Avec un robinet installé sur le récipient____________________ 2 En versant par le bec/l'ouverture du récipient______________ 3	
Nous sommes arrivés à la fin du questionnaire, il ne me reste que quelques questions sur votre niveau de vie, votre maison et le niveau d'éducation dans votre ménage. Ce sont des questions que nous posons à tous les ménages interrogés et qui nous permettent de mieux comprendre les conditions de vie qui peuvent influencer votre accès et votre utilisation de l'eau. Comme toutes les réponses précédentes, les réponses à ces questions sont totalement confidentielles.			
Informations sur le statut socio-économique du ménage			
12.1	Est-ce que le ou la chef de ménage homme ou femme dans votre ménage a un emploi salarié contractuel ?	❑ OUI ❑ NON ❑ Ne sait pas ❑Ne souhaite pas répondre	
12.3	Est-ce que votre ménage est propriétaire de son habitation ?	❑ OUI ❑ NON ❑ Ne sait pas ❑Ne souhaite pas répondre	
12.4	Quel est le type de sol dans l'habitation ? *(Plusieurs réponses possibles)*	Terre battue / sable / paille________________________❑ Bois / planches________________________________❑ Ciment_______________________________________❑ Carrelage_____________________________________❑ Autre (Spécifier)_______________________________❑ _______________________________________ Ne souhaite pas répondre____❑	
12.5	Combien de pièces sont utilisées par les membres du ménage pour dormir ?	Une____0 Deux____1 Trois____2 Plus de trois___________________________________3 Ne souhaite pas répondre_________________________4	
12.6	Quel est le mode principal d'éclairage de l'habitation ? *(Plusieurs réponses possibles)*	Branchement SNEL / groupe électrogène_________________❑ Panneaux solaires / lampe solaire_____________________❑ Lampe à pétrole ou à gaz__________________________❑ Lampe à piles__________________________________❑ Bougie / feu de bois______________________________❑ Autre (Spécifier)_______________________________❑ _______________________________________ Ne souhaite pas répondre____❑	
12.7	Quel est le combustible principal utilisé pour cuisiner ? *(Plusieurs réponses possibles)*	Electricité____________________________________❑ Gaz____❑ Kérosène / Pétrole_______________________________❑ Charbon de bois / braise__________________________❑ Bois de chauffe_________________________________❑ Sciure de bois__________________________________❑ Autre (Spécifier)_______________________________❑ _______________________________________ Ne souhaite pas répondre____❑	
12.8	Quel est le lieu d'aisance que les membres de votre ménage utilisent le plus souvent ? *(Plusieurs réponses possibles)*	Plage/champ de défécation/brousse_______________________❑ Trou dans la parcelle____________________________________❑ Latrine partagée / publique (hors de la parcelle)_____________❑ Latrine partagée dans la parcelle________________________❑ Latrine privée à l'extérieur du logement___________________❑ Latrine privée à l'intérieur du logement___________________❑ Ne souhaite pas répondre______________________________❑	
12.9	Est-ce que le ménage possède les équipements <u>fonctionnels</u> suivants ? *(option « ne souhaite pas répondre »possible)*	1. Télévision______________OUI / NON 2. Téléphone portable _____OUI / NON 3. Congélateur ____________OUI / NON 4. Réfrigérateur___________OUI / NON 5. Cuisinière______________OUI / NON 6. Mobylette/moto________OUI / NON	7. Voiture________________OUI / NON 8. Générateur_____________OUI / NON 9. Panneau solaire avec batterie ________________________OUI / NON 10. Ventilateur____________OUI / NON 11. Antenne satellite avec abonnement ________________________OUI / NON
Informations sur le niveau d'éducation dans le ménage			

13.1	Quel est le niveau d'éducation du chef de ménage <u>homme</u> dans votre ménage ?	Aucune instruction ---- 0 Primaire incomplet ---- 1 Primaire complet ---- 2 Secondaire incomplet ---- 3 Secondaire complet ---- 4 Etudes supérieures ---- 5 Pas de chef de ménage homme ---- 6 Ne sait pas ---- 98 Ne souhaite pas répondre ---- 99	
13.2	Quel est le niveau d'éducation du chef de ménage <u>femme</u> dans votre ménage ?	Aucune instruction ---- 0 Primaire incomplet ---- 1 Primaire complet ---- 2 Secondaire incomplet ---- 3 Secondaire complet ---- 4 Etudes supérieures ---- 5 Pas de chef de ménage femme ---- 6 Ne sait pas ---- 98 Ne souhaite pas répondre ---- 99	

Voilà, j'ai terminé les questions que je souhaitais vous poser dans le cadre de cette enquête. Je vous remercie sincèrement pour votre temps et toutes les réponses à ces questions. Il est possible qu'un superviseur de l'enquête revienne vous voir quelques minutes dans les prochains jours pour contrôler la qualité de mon travail. Est-ce que vous souhaitez me poser des questions sur cette enquête avant que je ne parte ?

Appendix 4-1b Questionnaire on water-related practices- (English summary)

1.1 Interviewer ID
1.2 Target GPS point ID
1.3 Consent confirmation
1.4 Household waypoint ID

2.1 Interview participant(s)
2.2 Exact household address and location information

3.1 Household composition (day before the interview)
3.2 Number and age group of people hosted officially as internally displaced
3.3 Number of households sharing the compound
3.4 Number of people sharing the compound
3.5 Time since the household moved into the dwelling

4.1 Usual water collection frequency per day
4.2 to 4.10 Estimated amount of fresh water used at home the day before the interview for adult bathing, children bathing, house cleaning, laundry, drinking, food / produce washing, dishwashing, handwashing, income generating activity.
4.11 Summary of water estimates
4.12 Water re-use for specific activities (day before interview)
4.13 Activities performed directly at a water source outside the compound (day before interview)

5.1 Main water source (in volume, for 2 weeks preceding the interview)
5.2 and 5.3 Neighbourhood and street for main water source

6.1 Duration of main water source use
6.2 Last use of main water source
6.3 Drinking of water collected at main source (last time it was used)
6.4 Using water collected at main source for income generating activities (last time itw as used)
6.5 Water collector(s) at main water source (last time it was used)
6.6 Travel time to and from the main water source
6.7 Transportation mean to main source (last time it was used)
6.8 Average waiting time at the main source
6.9 and 6.10 Issues at the main water source (during the 2 weeks preceding the interview) : turbid water, bad taste, bad smell, lack of pressure
6.11 Number of days water was flowing at tap used as main water (during the day, week preceding interview)
6.12 Number of nights water was flowing at tap used as main water (during the night, week preceding interview)
6.13 Absence of water at the tap when attempting to use it (during the 2 weeks preceding interview)

6.14 Household collecting water at main source during night time
6.15 Registered owner of the tap used as main source
6.16 Time since tap was installed
6.17 Fonctional water meter on the tap
6.18 Payment type for water collected
6.19 Amount paid last month if monthly payment
6.20 Difficulties to pay for water monthly (in the year preceding interview)
6.21 Amount paid for 20L jerrycan if payment per jerrican (last time source was used)
6.22 Difficulties to pay for jerrycan (in the month preceding the interview)
6.23 Number of other households using the tap used as main water source (in the month preceding the interview)
6.24 Regideso invoice based on meter or estimate (last monthly invoice)
6.25 Regideso invoice amount in CDF (last monthly invoice)
6.26 Number of households paying for water at owned tap (in the month preceding interview)
6.27 Type of payment applied to other households for owned tap (in the month preceding the interview)
6.28 to 6.31 Amount paid by households for owned tap, per use, per week, per month or per jerrican (in CDF, in the month preceding the interview)
6.32 to 6.34 Complaint ever made to the Regideso for owned tap, when was the last complaint, to whom the complaint was made
6.35 to 6.37 Reasons for not having or using a tap in the compound as main source of water

7.1 to 7.3 1st alternative water source in the past two weeks : type and location
7.4 to 7.37 Repeat of questions 6.1 to 6.34 for 1st alternative water source

8.1 to 8.3 2nd alternative water source in the past two weeks : type and location
8.4 to 8.37 Repeat of questions 6.1 to 6.34 for 2nd alternative water source

9.1 Water treatment methods ever used for drinking water
9.2 Known products for drinking water treatment
9.3 and 9.4 Time when drinking water was treated last : time since and reason why
9.5 Activities for which treated water is used in the household in addition to drinking

10.1 Total amount of water that can be stored in the household if all containers are filled up

11.0 and 11.0b Consent to observe stored drinking water and storage conditions observations (in the largest container)
11.1 Consent to sample and analyse stored drinking water
11.2 Time the stored drinking water was collected
11.3 Source from which stored drinking water was collected
11.4 Observation of the storage conditions from which sample was collected
11.5 Observation of how the sample was collected

12.1 Male and/or female head of household holding a contractual salaried employment
12.3 Dwelling ownership

12.4 Dwelling floor type
12.5 Number of rooms used for sleeping
12.6 Sources of lighting
12.7 Cooking fuel
12.8 Sanitation facilities used
12.9 Ownership of durable assets

13.1 Highest education level achieved by the male head of household
13.2 Highest education level achieved by the female head of household

Appendix 4-2 Visual aid to estimate water quantity used by households during interviews

Appendix 4-3 Polychoric PCA coefficients for wealth index construction

		N	n (%)	Coef.
Ownership of functional assets				
TV	No			-0.177
	Yes	518	191 (36.9%)	0.291
Mobile phone	No			-0.236
	Yes	518	433 (83.6%)	0.043
Freezer	No			-0.084
	Yes	518	72 (13.9%)	0.483
Fridge	No			-0.032
	Yes	517	28 (5.4%)	0.449
Stove	No			-0.047
	Yes	516	35 (5.8%)	0.525
Moped / motorbike	No			-0.033
	Yes	516	36 (7%)	0.36
Car	No			-0.047
	Yes	518	38 (7.3%)	0.527
Generator	No			-0.03
	Yes	516	19 (3.7%)	0.552
Solar panel with battery	No			-0.054
	Yes	517	73 (14.1%)	0.301
Ventilator	No			-0.066
	Yes	518	54 (10.4%)	0.511
Satellite TV	No			-0.064
	Yes	516	50 (9.7%)	0.515
Dwelling ownership and size				
Household owning their dwelling	No			-0.019
	Yes	516	331 (63.9%)	0.01
3 or more rooms used for sleeping	No			-0.14
	Yes	518	284 (54.8%)	0.112
Source of lighting (ordered)				
	Torch	517	195 (37.7%)	-0.281
	Petrol light	517	38 (7.3%)	-0.064
	Solar panel or solar-charged lights	517	120 (23.2%)	0.042
	Power grid or generator	517	164 (31.7%)	0.306
Cooking fuel (ordered)				
	Sawdust / wood	518	61 (11.8%)	-0.402
	Charcoal	518	435 (84%)	0.022
	Gaz	518	2 (0.4%)	0.38
	Electricity	518	20 (3.9%)	0.482
Dwelling floor material (ordered)				
	Mud	518	232 (44.8%)	-0.215
	Wood	518	32 (6.2%)	-0.016
	Cement	518	231 (44.6%)	0.155
	Tiles	518	23 (4.4%)	0.481
Defecation place often used by household members (ordered)				
	Open defecation / hole in the ground	516	103 (20%)	-0.242
	Public latrines	516	19 (3.7%)	-0.134
	Shared latrine in the compound	516	212 (41.1%)	-0.022
	Private latrine outside the dwelling	516	138 (26.7%)	0.142
	Private latrine inside the dwelling	516	44 (8.5%)	0.311

Appendix 4-4 Results of linear regression models of the reported amount of water used at home by households for specific activities

	Adult bathing		Children bathing[1]		Laundry	
	Coefficient (95% CI)	p-value	Coefficient (95% CI)	p-value	Coefficient (95% CI)	p-value
Intercept (household with 1 adult member)	22.4 (12.9; 31.8)	p<0.001 ***	13.8 (7.8; 19.8)	p<0.001 ***	20.7 (7.2; 34.1)	0.003 **
Number of extra household members aged 15 or more	11 (9.9; 12.2)	p<0.001 ***	0.5 (-0.2; 1.3)	0.16	1.7 (0.1; 3.4)	0.04 *
Number of household members aged between 5 and 14	1.4 (0.1; 2.8)	0.03 *	6.3 (5.4; 7.1)	p<0.001 ***	3.4 (1.5; 5.2)	p<0.001 ***
Number of household members aged between 2 and 4	1.2 (-1.3; 3.8)	0.34	3.5 (1.8; 5.1)	p<0.001 ***	1.3 (-2.4; 4.9)	0.5
Number of household members aged 2 or less	-4.3 (-7.8; -0.8)	0.02 *	0.7 (-1.6; 2.9)	0.55	5.6 (0.7; 10.6)	0.03 *
Activity performed directly at the source	-7.6 (-13.1; -2)	0.008 **	-3.7 (-7.3; 0)	0.05 *	-14.6 (-22.5; -6.6)	p<0.001 ***
Reuse of water at home	-7.1 (-15.9; 1.7)	0.12	0.8 (-4.8; 6.5)	0.77	0.7 (-11.8; 13.2)	0.91
Paying for water	-2.2 (-8.7; 4.4)	0.52	0.9 (-3.3; 5.1)	0.68	-0.3 (-9.6; 9)	0.96
Total time spent to collect water						
2 minutes and less (within the compound)	reference		reference		reference	
Between 3 and 30 minutes	-8.2 (-16.9; 0.4)	0.06	-3.2 (-8.8; 2.4)	0.26	-1.6 (-14; 10.7)	0.8
Between 31 and 90 minutes	-11.8 (-19.9; -3.8)	0.004 **	-1.9 (-7.1; 3.3)	0.47	-1.8 (-13.3; 9.7)	0.76
More than 91 minutes	-15.4 (-23.6; -7.2)	p<0.001 ***	-4 (-9.3; 1.3)	0.14	-8.1 (-19.9; 3.6)	0.18
Wealth quintile						
1 (poorest)	9.3 (1.1; 17.5)	0.03 *	0.1 (-5.2; 5.4)	0.99	-1.2 (-12.9; 10.5)	0.84
2	4.6 (-3.6; 12.8)	0.27	-3.1 (-8.4; 2.2)	0.26	-3.4 (-15; 8.3)	0.57
3	5.9 (-2.2; 13.9)	0.15	-2.2 (-7.4; 3)	0.41	-7.7 (-19.1; 3.8)	0.19
4	10.3 (2.9; 17.6)	0.006 **	-1.1 (-5.8; 3.7)	0.66	-5.5 (-15.9; 5)	0.31
	reference		reference		reference	
Adjusted R-squared		0.54		0.46		0.1
Residual standard error on 362 degrees of freedom		21.5		13.9		30.6

[1] For the quantity used for children washing, the intercept represents a household with 1 adult member and 1 member aged 5 to 14

Significance codes: *** =[p<0.001]; **= [0.001<=p< 0.01]; *=[0.01<=p<0.05]

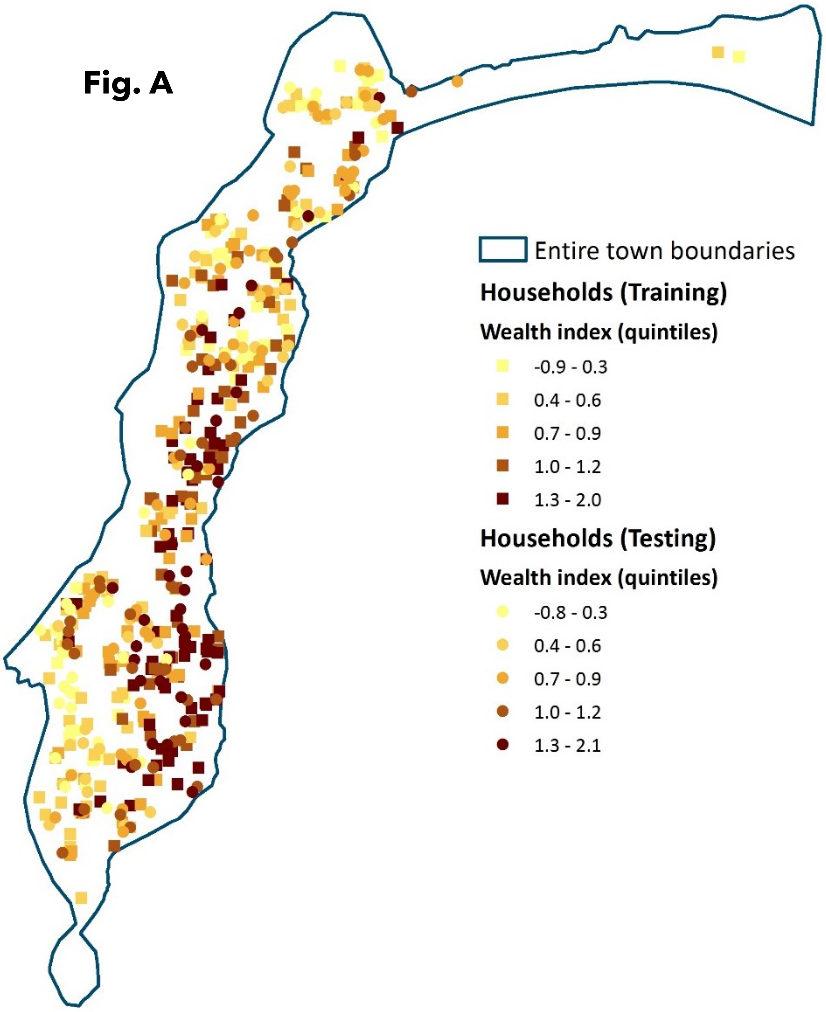
	House cleaning		Dishwashing		Food/produce rinsing		Handwashing	
	Coefficient (95% CI)	p-value	Coefficient (95% CI)	p-value	Coefficient (95% CI)	p-value	Coefficient (95% CI)	p-value
Intercept (household with 1 adult member)	15.2 (11.5; 18.9)	p<0.001 ***	15.3 (12.4; 18.2)	p<0.001 ***	12.6 (9.5; 15.8)	p<0.001 ***	2.7 (1.5; 3.8)	p<0.001 ***
Number of extra household members aged 15 or more	1 (0.5; 1.5)	p<0.001 ***	0.9 (0.5; 1.2)	p<0.001 ***	1 (0.6; 1.4)	p<0.001 ***	0.4 (0.3; 0.5)	p<0.001 ***
Number of household members aged between 5 and 14	0.9 (0.4; 1.4)	0.001 **	0.3 (-0.1; 0.7)	0.12	1.5 (1.1; 1.9)	p<0.001 ***	0.4 (0.3; 0.6)	p<0.001 ***
Number of household members aged between 2 and 4	-0.3 (-1.3; 0.7)	0.59	0.6 (-0.2; 1.3)	0.17	2.1 (1.3; 3)	p<0.001 ***	0.1 (-0.2; 0.4)	0.48
Number of household members aged 2 or less	0.5 (-0.8; 1.9)	0.45	-0.1 (-1.1; 1)	0.89	0 (-1.2; 1.1)	0.93	0 (-0.5; 0.4)	0.88
Activity performed directly at the source	0.6 (-1.6; 2.8)	0.61	-1.9 (-3.6; -0.2)	0.03 *	-0.9 (-2.8; 0.9)	0.33	-0.3 (-1; 0.4)	0.45
Reuse of water at home	-3.3 (-6.7; 0.2)	0.06	0.9 (-1.8; 3.6)	0.52	1.5 (-1.4; 4.4)	0.32	1.2 (0.1; 2.3)	0.04 *
Paying for water	0.3 (-2.3; 2.8)	0.84	-1.7 (-3.7; 0.3)	0.1	-2.8 (-4.9; -0.6)	0.01 *	0.2 (-0.6; 1)	0.61
Total time spent to collect water								
2 minutes and less (within the compound)	reference		reference		reference		reference	
Between 3 and 30 minutes	-5.3 (-8.7; -1.9)	0.002 **	-2.8 (-5.5; -0.1)	0.04 *	-2.3 (-5.1; 0.6)	0.13	0.2 (-0.9; 1.3)	0.77
Between 31 and 90 minutes	-5 (-8.2; -1.8)	0.002 **	-2.4 (-4.9; 0.1)	0.06	-0.1 (-2.8; 2.6)	0.92	-0.4 (-1.4; 0.6)	0.42
More than 91 minutes	-5 (-8.3; -1.8)	0.003 **	-2.5 (-5; 0.1)	0.06	0.7 (-2.1; 3.4)	0.63	-0.5 (-1.6; 0.5)	0.31
Wealth quintile								
1 (poorest)	-6 (-9.2; -2.7)	p<0.001 ***	-3.2 (-5.7; -0.7)	0.01 *	-2.7 (-5.5; 0)	0.05	-0.1 (-1.1; 1)	0.91
2	-4.9 (-8.2; -1.7)	0.003 **	-3 (-5.5; -0.5)	0.02 *	-2.3 (-5; 0.4)	0.1	-0.1 (-1.1; 0.9)	0.83
3	-3.7 (-6.8; -0.5)	0.02 *	-2.2 (-4.7; 0.3)	0.09	-1.7 (-4.4; 1)	0.22	-0.1 (-1.1; 0.9)	0.89
4	-2.2 (-5.1; 0.6)	0.13	-1 (-3.3; 1.3)	0.39	-1.1 (-3.6; 1.3)	0.36	0.1 (-0.8; 1.1)	0.76
	reference		reference		reference		reference	
Adjusted R-squared		0.23		0.14		0.29		0.18
Residual standard error on 362 degrees of freedom		8.5		6.6		7.2		2.7

Significance codes: *** =[p<0.001]; **= [0.001<=p< 0.01]; *=[0.01<=p<0.05]

Appendix 6-1 Empirical Bayesian Kriging (EBK) of households' wealth index

A total of 511 households with wealth index was divided into a training set (n=318; 62.2 %) and a testing set (n=193). Wealth index for all households ranged between -0.9 and 2.1 (mean 0.77; standard deviation 0.54). There was no evidence of a difference in wealth index distribution between the testing and training sets of households (**Figure A**).

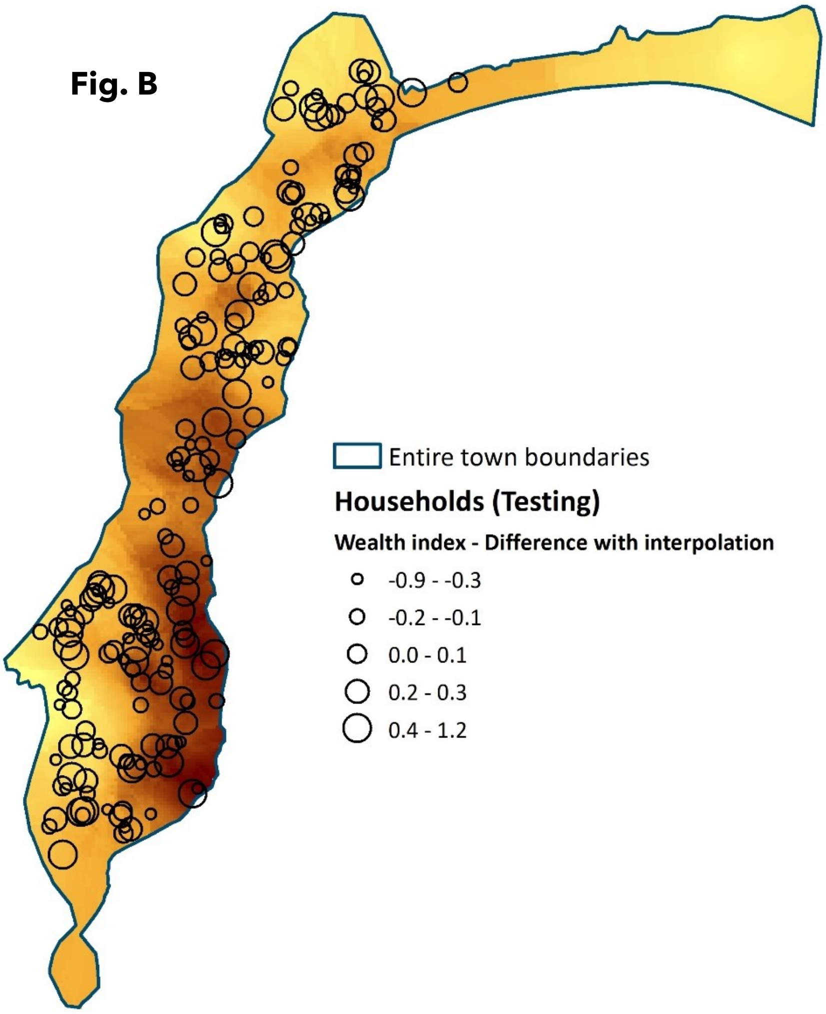
Fig. A



Four different EBK parametrisations were run on the testing set of households for wealth index interpolation over the entire town surface. All were based on 100 simulated semi-variograms assumed to follow a power function. Differences were related to the number of points included in the circular search and the degree of overlap between local models.

The difference between actual wealth index and interpolation values for households in the testing set was calculated and summarised into mean and standard deviation. The interpolation parameters leading to the smallest standard deviation was selected; a circular search with a number of neighbours between 10 and 15 points and an overlap of 1 (average number of local models each point is included into)(**Figure B**). The mean difference between actual wealth index and interpolation was 0.009, with a standard deviation of 0.37.

Fig. B



Appendix 6-2 Spatial smoothing of population estimates

Population numbers for each of the 201 analysis base units (ABU) were originally compiled from the administrative census available in November 2017. They were aggregated/summed up into the 37 neighbourhoods used for time-space modelling and estimates over time were constructed as described in chapter *Data and methods: an overview* and *appendix 2-4*.

Discrete population numbers for ABUs in November 2017 were converted to a constant population surface over ABUs built-up areas (**Figure A**). Distribution of ABUs population density was highly skewed, with most outliers being very small ABUs grouped into the neighbourhood Rombe I. To compensate for the implausible assumption that population density was sharply changing over ABUs delimitation, we therefore smoothed population surfaces by applying a focal mean radius over 100 m distance, effectively averaging population density over ABUs limits (**Figure B**). This reduced noticeably the skewness of population density distribution by ABU and by neighbourhood (**Figure C**).

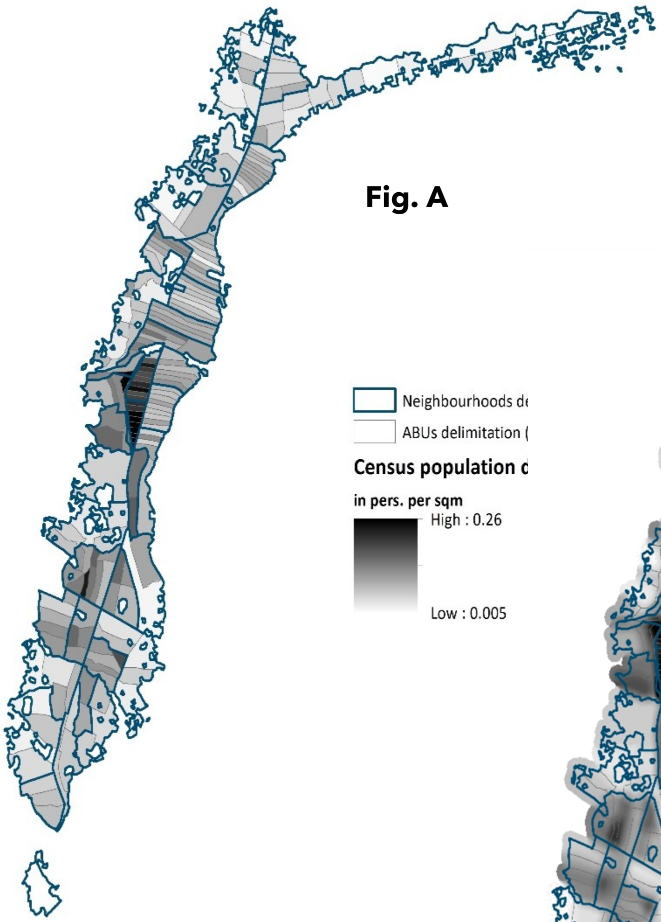


Fig. A

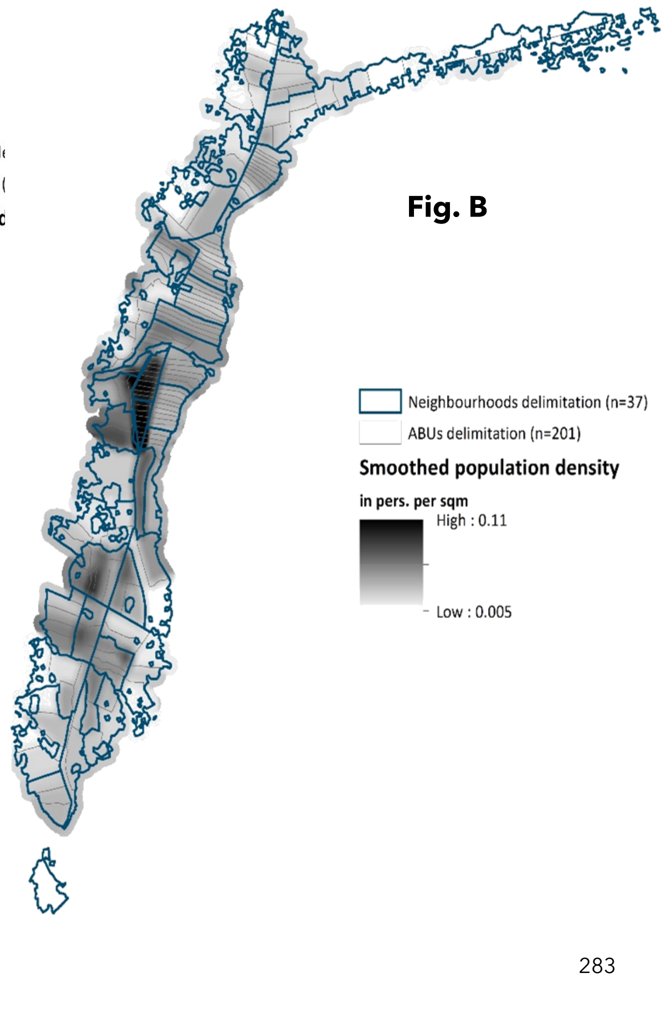
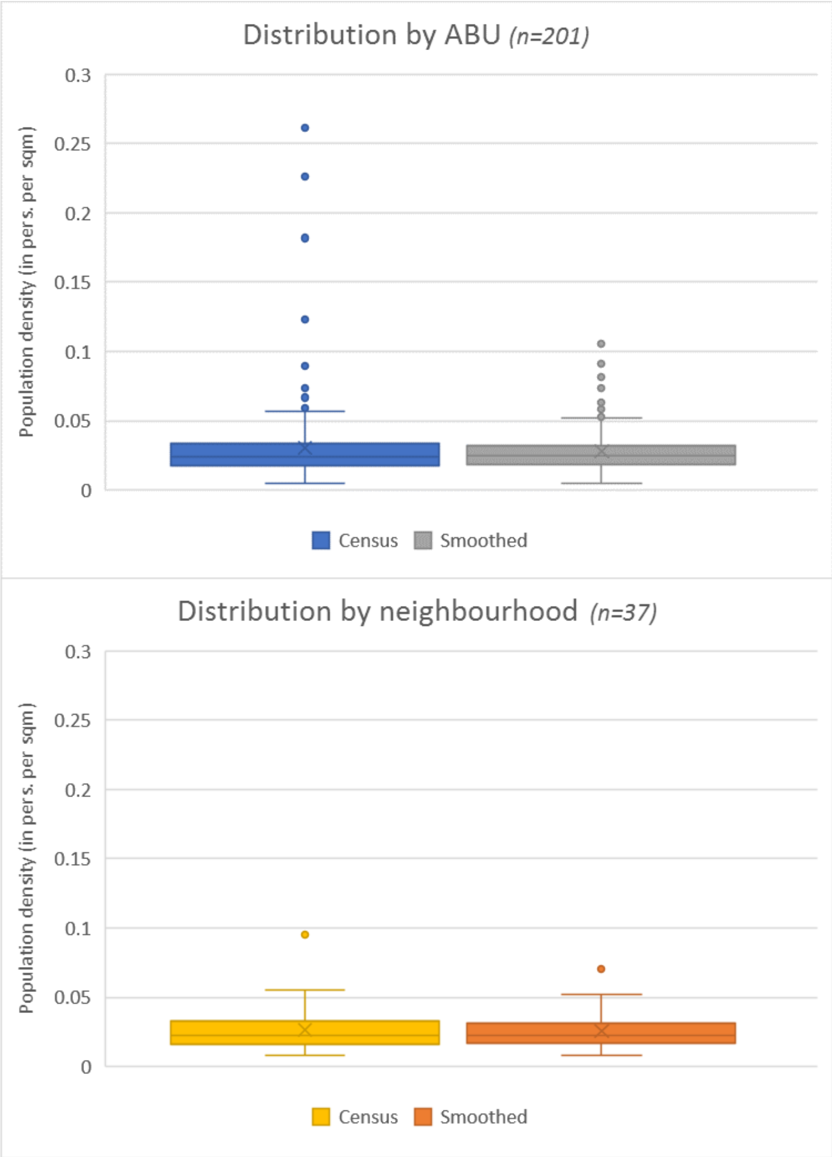


Fig. B

Fig. C



The smoothed population surface was then aggregated/summed up into 37 neighbourhoods. Population change rates over time for each neighbourhood was then applied as for the original estimates, to generate retrospective and forecasted population estimates per quartier and then per neighbourhood from the November 2017 value (**Table next page**).

Neighbourhood	**Neighbourhood census date** (dd/mm/yyyy) *Time index (days)*										**Smoothed spatial estimates**	**Forecasted estimates**	
	01/07/2008 *1*	**01/07/2010** *731*	**15/02/2012** *1325*	**15/05/2012** *1415*	**15/02/2013** *1691*	**15/05/2014** *2145*	**15/11/2014** *2329*	**15/11/2015** *2694*	**15/05/2016** *2876*	**15/08/2017** *3333*	**15/11/2017** *3425*	**15/03/2018** *3545*	**16/05/2019** *3972*
	10,284	*12,692*	*12,624*	*12,514*	*13,834*	*14,092*	*14,274*	*14,290*	*15,400*	*15,679*	*15,981*	*16,204*	*17,003*
Kabindula	**10,109**	**12,517**	**12,449**	**12,339**	**13,659**	**13,917**	**14,099**	**14,115**	**15,225**	**15,504**	**15,806**	**16,029**	**16,828**
	18,888	*24,965*	*24,702*	*24,377*	*24,678*	*24,661*	*24,636*	*24,148*	*29,399*	*29,435*	*29,487*	*31,523*	*34,380*
Kakombe	**18,901**	**24,978**	**24,715**	**24,390**	**24,691**	**24,674**	**24,649**	**24,161**	**29,412**	**29,448**	**29,500**	**31,537**	**34,393**
	17,205	*23,965*	*23,899*	*23,966*	*23,394*	*23,483*	*23,577*	*23,589*	*23,999*	*24,341*	*24,408*	*24,608*	*25,075*
Kalundu	**17,673**	**24,433**	**24,367**	**24,434**	**23,862**	**23,951**	**24,045**	**24,057**	**24,467**	**24,809**	**24,876**	**25,076**	**25,543**
	13,088	*15,344*	*15,337*	*15,288*	*19,206*	*19,206*	*19,228*	*19,239*	*24,601*	*24,421*	*27,125*	*26,487*	*29,248*
Kasenga	**13,457**	**15,713**	**15,706**	**15,657**	**19,575**	**19,575**	**19,597**	**19,608**	**24,970**	**24,790**	**27,494**	**26,856**	**29,617**
	15,564	*16,382*	*16,538*	*18,123*	*18,442*	*18,592*	*18,830*	*19,340*	*19,381*	*20,170*	*20,358*	*20,428*	*21,018*
Kavimvira	**15,388**	**16,206**	**16,362**	**17,947**	**18,266**	**18,416**	**18,654**	**19,164**	**19,205**	**19,994**	**20,182**	**20,252**	**20,842**
	2,810	*4,030*	*3,982*	*4,058*	*4,770*	*5,044*	*4,866*	*4,875*	*5,091*	*5,369*	*5,418*	*5,540*	*5,856*
Kibondwe	**2,542**	**3,762**	**3,714**	**3,790**	**4,502**	**4,776**	**4,598**	**4,607**	**4,823**	**5,101**	**5,150**	**5,272**	**5,588**
	7,903	*8,678*	*8,615*	*8,833*	*9,064*	*9,197*	*9,261*	*9,272*	*9,519*	*9,744*	*9,808*	*9,911*	*10,206*
Kilibula	**7,707**	**8,482**	**8,419**	**8,637**	**8,868**	**9,001**	**9,065**	**9,076**	**9,323**	**9,548**	**9,612**	**9,715**	**10,010**
	11,149	*12,951*	*12,961*	*13,110*	*13,509*	*13,564*	*13,620*	*13,639*	*14,057*	*14,328*	*14,352*	*14,576*	*14,997*
Kimanga	**10,945**	**12,747**	**12,757**	**12,906**	**13,305**	**13,360**	**13,416**	**13,435**	**13,853**	**14,124**	**14,148**	**14,373**	**14,793**
	16,426	*23,213*	*23,213*	*22,234*	*22,748*	*22,748*	*22,988*	*22,998*	*25,668*	*26,226*	*24,978*	*27,453*	*29,294*
Mulongwe	**17,802**	**24,589**	**24,589**	**23,610**	**24,124**	**24,124**	**24,364**	**24,374**	**27,044**	**27,602**	**26,354**	**28,829**	**30,670**
	14,550	*11,204*	*11,218*	*11,243*	*11,138*	*11,138*	*11,198*	*11,211*	*11,239*	*14,423*	*14,325*	*15,410*	*17,716*
Nyamianda	**14,823**	**11,477**	**11,491**	**11,516**	**11,411**	**11,411**	**11,471**	**11,484**	**11,512**	**14,696**	**14,598**	**15,683**	**17,989**
	18,801	*17,455*	*17,452*	*18,220*	*21,167*	*21,551*	*20,783*	*20,793*	*23,135*	*24,146*	*24,446*	*25,382*	*27,372*
Rombe I	**14,934**	**13,588**	**13,585**	**14,353**	**17,300**	**17,684**	**16,916**	**16,926**	**19,268**	**20,279**	**20,579**	**21,516**	**23,506**
	12,308	*14,344*	*14,392*	*15,150*	*15,177*	*15,912*	*15,988*	*16,002*	*16,639*	*17,150*	*17,150*	*17,559*	*18,270*
Rombe II	**14,426**	**16,462**	**16,510**	**17,268**	**17,295**	**18,030**	**18,106**	**18,120**	**18,757**	**19,268**	**19,268**	**19,677**	**20,388**
	5,532	*6,419*	*6,404*	*6,379*	*6,177*	*6,198*	*6,156*	*6,173*	*6,355*	*7,024*	*7,035*	*7,301*	*7,881*
Rugenge	**5,695**	**6,582**	**6,567**	**6,542**	**6,340**	**6,361**	**6,319**	**6,336**	**6,518**	**7,187**	**7,198**	**7,464**	**8,043**
	17,050	*18,004*	*18,013*	*17,992*	*19,077*	*19,077*	*19,095*	*19,118*	*19,280*	*20,224*	*19,567*	*20,577*	*21,344*
Songo	**17,392**	**18,346**	**18,355**	**18,334**	**19,419**	**19,419**	**19,437**	**19,460**	**19,622**	**20,566**	**19,909**	**20,919**	**21,686**
	181,558	*209,646*	*209,350*	*211,487*	*222,381*	*224,463*	*224,500*	*224,687*	*243,763*	*252,680*	*254,438*	*262,957*	*279,658*
TOTAL Uvira	**181,795**	**209,883**	**209,587**	**211,724**	**222,618**	**224,700**	**224,737**	**224,924**	**244,000**	**252,917**	**254,675**	**263,194**	**279,895**

Note: Estimates from census are indicated in small italic font above the smoothed estimate.