Evaluating Hygiene Behaviour Change Within Community Health Clubs in the Rusizi District of Rwanda

Master's thesis submitted to the Charité – Universitätsmedizin Berlin in partial fulfilment of the requirements for the award of a Master of Science degree in International Health

Julia Pantoglou

February, 2018

Thesis Supervisor and First Marker

Dr Richard Carter

Second Marker

NN

Table of Contents

1. Abstract

This project focuses on identifying whether significant hygiene behaviour change took place among Community Health Clubs in the Rusizi district of Rwanda. The project also focuses on identifying whether common factors exist among top performing and low performing Community Health Clubs. The project utilises monitoring and evaluation data that was collected between 2013 and 2017 by Africa AHEAD for 50 villages that were selected to be part of a cluster Randomised Control Trial conducted by Innovations for Poverty Action. Methods used included a variety of quantitative methods for assessing statistical relationships between proxy-indicators that could reflect performance and behaviour change at the household level. The overall aim of the project was to develop a new framework for post processing monitoring and evaluation data in order to gain a deeper insight into why certain villages adopt and benefit from the Community Health Club methodology better than others. The findings showed that statistically significant improvement in overall household performance had taken place across sampled Community Health Clubs over time. It was also found that these findings could be confidently extended to all Rusizi District Community Health Clubs. Common traits were identified among top and low performing groups. It was found that low performing groups consistently struggled with the proxy-indicators relating to body hygiene, water source, sanitation, and hand washing, while high performing groups performed more consistently over time and achieved higher indicator scores than the low groups in every dataset. High groups were also found to attend CHC sessions across datasets more consistently than low groups. The primary conclusion of the study was that behaviour change had taken place within Rusizi district Community Health Clubs, however with significant variation between clubs and datasets.

2. List of Abbreviations

AA	Africa AHEAD
AIDS	Acquired Immune Deficiency Syndrome
BCC	Behaviour Change Communication
CBEHPP	Community Based Environmental Health Promotion Programme
CBO	Community Based Organisation
CHC	Community Health Club
CHW	Community Health Worker
cRCT	cluster-Randomised Control Trial
EHO	Environmental Health Worker
HHI	Household Inventory
HIV	Human Immunodeficiency Virus
HSSP-I	Health Sector Strategic Plan 2005-2009
IEC	Information Education Communication
IPA	Innovations for Poverty Action
M&E	Monitoring and Evaluation
MDGs	United Nations Millennium Development Goals
MOH	Ministry of Health
NEPAD	New Partnership for African Development
NGO	Non-Governmental Organisation
ODK	Open Data Kit
PHAST	Participatory Hygiene and Sanitation Transfer
PI	Post Intervention
SPSS	Statistical Package for the Social Sciences
WASH	Water Sanitation and Hygiene

3. Executive Summary

The aim of this project is to assess the effectiveness of the Rwandan Ministry of Health's Community Based Environmental Health Promotion Programme (CBEHPP), in generating positive sustainable health and hygiene behaviour change within the Rusizi district of Rwanda. This is achieved through the analysis of five sets of Monitoring & Evaluation (M&E) data collected between October 2013 and March 2017 from 50 villages that received training as part of the programme. This includes one pre-intervention Baseline dataset, two collected during the intervention (Midline and Endline), and two Post Intervention (PI) datasets.

The programme was jointly developed, and put into policy in 2009, by the Rwandan Ministry of Health (MOH) and the African based international Non-Governmental Organisation (NGO) Africa AHEAD (AA) (Ministry of Health, 2010), who created the Community Health Club (CHC) approach on which the CBEHPP is based.

The CBEHPP implementation and M&E survey campaigns in Rusizi district were initiated in 2012 as part of efforts funded by the Bill and Melinda Gates Foundation to assess whether the programme has any significant effects on the primary health outcomes of child diarrhoea and anthropometry, as well as the secondary outcome of household water quality. The study was conducted by the international NGO Innovations for Poverty Action (IPA), who designed a cluster-Randomised Control Trial (cRCT) that was applied to 150 randomly selected villages within the district. Of these, AA was tasked with overseeing the delivery of CBEHPP training to 100 villages, delivering a 'Classic' version of their CHC training package to 50 villages, and a less intensive 'Lite' version to the other 50. The remaining 50 villages were used as the control group and received no training throughout the intervention period (Sinharoy *et al.*, 2017). The M&E survey campaigns focused mainly on the villages that received Classic training and were conducted in parallel to the IPA study by Africa AHEAD.

This study was motivated by the findings of the IPA study, which indicated that the CBEHPP had little to no effect on the primary and secondary health outcomes that were its focus. It also stated that only mixed results were observed with respect to "health behaviour change", and that while positive effects on several intermediate outcomes were suggested, none of these translated into reductions in care-giver reported child diarrhoea rates or improved water quality measurements (Sinharoy *et al.*, 2017) .

While the IPA study admitted to methodological limitations within their approach and was subsequently criticised on several points by (Cairncross *et al.*, submitted for publication), the implication that the Rusizi intervention had experienced limited success and lack of consistency, presented the opportunity for deeper analysis of the programme's performance via the independent M&E datasets. This acted as the primary motivation for this thesis project.

In particular, the objectives of this project are to:

1. Quantify overall hygiene behaviour change for the surveyed villages by analysing aggregated observational data on key hygiene indicators.
2. Establish the robustness and limits of the data for the purposes of drawing general conclusions about the programme's ability to generate positive hygiene behaviour change at scale.
3. Rank villages based on overall performance in order to establish whether common traits exist among top performing and low performing villages based on a detailed analysis of individual hygiene behaviour change indicators.
4. Understand why such variations in performance might exist in order to suggest improvements for both the training programme, as well as the existing M&E methodology, and to understand what factors might have impacted the Rusizi programme in particular.

These objectives can be consolidated into the following core research questions:

I. Did significant hygiene behaviour change take place within Rusizi district CHCs? If so, to what extent, and with how much consistency did these changes take place across observed CHCs?
II. Do common traits exist among high-performing and low-performing CHCs with respect to hygiene behaviour change? If so, can these be linked to elements of the CBEHPP training and M&E programme implementations?

The following methodological steps were employed in order to achieve the objectives above:

i. The datasets were pre-processed to prepare them for quantitative analysis. This entailed: data cleaning, aligning the datasets to ensure continuity when conducting

inter-dataset comparisons, and designing a system to convert survey answers to unbiased numerical values for statistical analysis.

Standardised statistical tests were applied to the datasets in parallel with the main analyses in order to track errors, understand the robustness of the data, and to apply limits to how generalizable the results and conclusions might be. This included tests for statistical significance, assessment of sample sizes, and evaluating the central tendency of results.

ii. The aggregated means of key hygiene indicators were computed at household and village levels for each dataset. These were compared to Baseline values in order to measure overall behaviour change. This was then used to rank villages based on their level of improvement, which in turn allowed for the categorisation of villages based on performance into low and high performing groups.

iii. A detailed analysis was then possible in order to establish common trends within each group, as well as key differences between groups, by looking at performance based on individual indicator statistics.

iv. Finally, the results from this analysis were compared to literature sources and externally available information, in order to understand why such trends and differences might have existed between groups.

The findings of the study showed that statistically significant improvement in overall household performance had taken place across sampled CHCs over time. It was also found that these results could be confidently extended to all Rusizi District Community Health Clubs. Common traits were identified among top and low performing CHCs. It was found that households from low performing CHCs consistently struggled with the proxy-indicators relating to body hygiene, water source, sanitation, and hand washing, while those from high performing CHCs performed more consistently over time and achieved higher indicator scores than the low groups in every category and dataset. The primary conclusions of the study were that significant behaviour change had taken place within Rusizi district CHCs, however with significant variation between both clubs and datasets. This implied that a lack of consistent behaviour change had taken place across Rusizi district CHCs. The performance comparison study concluded that the indicators that low performing households struggled with could be used to identify them at an early stage, in order to provide additional support.

4. Introduction and Overview

This section provides the essential background information required to place this project within both the specific context of the Rusizi intervention, including the IPA study and AA's M&E programme, as well as the broader context of public health strategy in Rwanda at the time of the CBEHPP implementation.

This is achieved by first giving a more detailed overview of the IPA study, including its primary structure, objectives, and main findings. AA's role is explained, before the development and implementation of AA's M&E programme, as well as the various data collection rounds are described. This is followed by an introduction of the raw datasets available for this study, and the development of the primary research questions, with the overall aim of laying the groundwork for the development of the study's methodology.

Following this, a more in-depth description of the state of public health in Rwanda is given, in order to outline the role of the CBEHPP, including its primary goals. This is followed by a description of the development and implementation of the CBEHPP in Rwanda, including its relation to the CHC approach. A brief literature review is then carried out in order to give an overview of existing CHC evaluation studies, and to explain the underlying social theories and mechanisms by which the CHC approach attempts to generate positive hygiene behaviour change.

4.1. Overview of the Rusizi Intervention

A formal evaluation of the CBEHPP's impact on health was initiated in 2012, two years after the nationwide launch of the programme in Rwanda. IPA designed a four-year study, funded by the Bill and Melinda Gates Foundation, to assess whether the programme showed any significant effects on a number of health outcomes. In order to accomplish this, Sinharoy et al. (2017) applied a cRCT study, designed to evaluate the effects of the CBHEPP (and therefore the CHC model's implementation in Rwanda) on health outcomes relating to child diarrhoea, child anthropometry and household water quality. These included the primary outcomes of caregiver-reported diarrhoea in children under five years of age, as well as child stunting and wasting, as measured by child height and weight versus age. Household water quality, as measured by counting the number of colony forming units of thermotolerant coliforms per 100 mL of water, was considered a secondary health outcome. Intermediate outcomes included observations relating to improvements in drinking water sources,

household water treatment, nutrition, and food security, in addition to the presence of improved sanitation and handwashing facilities (including structural improvements), and the sanitary disposal of children's faeces.

The Rusizi district, located in Rwanda's Western province, and bordering the Democratic Republic of the Congo and Burundi, was chosen as the location for the IPA study due to its high burden of disease, and lack of previous CBEHPP implementations.

Figure 1 – Map of Rwanda, showing its four provinces. Rusizi district is outlined in red.

IPA randomly selected a total of 150 villages and divided them into three separate branches. 50 villages were to receive the full 'Classic' CHC treatment, while another 50 were to receive a more compact 'Lite' version of the CHC training programme. The remaining 50 villages were to act as a control group, receiving no intervention. The 150 cRCT villages are shown in Figure 2 and were selected so as to minimise the number of shared borders between each other in order to minimise possible contamination through spill-over.

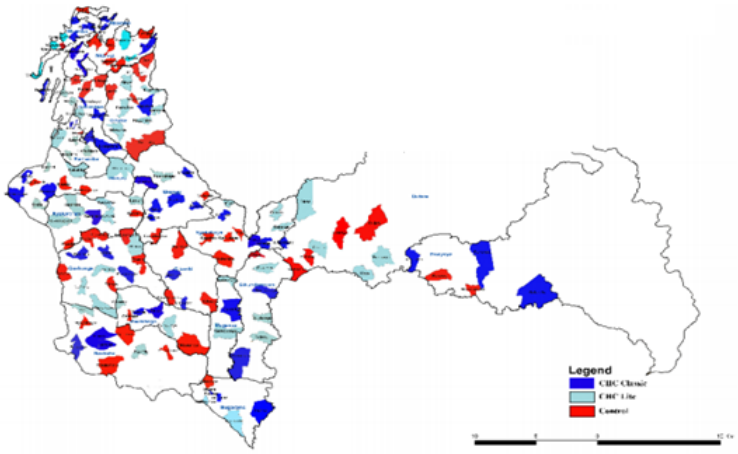

Figure 2 - Distribution of Classic (dark blue), Lite (light blue) and Control (red) villages, as randomly selected by IPA (Africa AHEAD, 2015a).

AA was given the responsibility of overseeing the implementation of the CHC programme in Rusizi. This included the training of local government and volunteer CHC facilitators, who would deliver CHC training to the villages.

Intervention Data Collection Rounds

In addition to the data collection rounds conducted as part of the IPA cRCT study, M&E data was collected in parallel by AA in a series of five data collection rounds that focused primarily on the 50 Classic CHC villages. A timeline of all data collection rounds relating to the Rusizi intervention can be seen in Table 1 below.

Period of data collection	Survey type	Team
May – Aug 2013	IPA Baseline	IPA
Oct – Nov 2013	Baseline	AA
Apr – May 2014	Midline	AA
Dec 2014	Endline	AA
Sep – Dec 2015	IPA Endline	IPA
Apr – May 2016	Post Intervention I	AA
Feb – Mar 2017	Post Intervention II	AA

Table 1 –Timeline of data collection rounds carried out by IPA and AA.

The M&E surveys were designed by AA to monitor the progress of the clubs both during and after the intervention, relative to their pre-intervention hygiene behaviour levels. This was achieved through the use of observational surveys known as 'Household Inventories' (HHIs), that were composed of questions based on approximately 50 'proxy-indicators' that could indicate the level of positive behaviour change being practised by a given household. The surveys were conducted on randomly selected households from selected CHCs by senior facilitators known as Environmental Health Officers.

Proxy indicators are defined as the observable outcomes of positive health practices, such as the presence of well-maintained hand-washing and sanitation facilities, or the level of cleanliness of cooking and sleeping areas. These are different from the health outcomes measured by studies such as the one conducted by IPA, in that they are more reliable for large scale operational M&E efforts (Waterkeyn and Cairncross, 2005), but do not provide direct information on the health benefits of a given practice. They only indicate whether a recommended behaviour is being practised by a given household. The 50 or so proxy indicators (also known as sub-indicators) can then be grouped into approximately ten main categories that give an overall view of whether specific positive behaviours are being

practiced. These categories are based on core elements of the CHC training programme and can be seen in Table 2, along with the number of sub-indicators per category per dataset. These numbers differed between datasets due to incremental improvements that were made to the corresponding survey designs between data collection rounds.

	Number of sub - indicators			
Main Indicators	Baseline	Midline/ Endline	PI – I/II	Thesis
Housing	-	-	4	-
Compound	8	7	5	4
Water Source	6	5	5	4
Drinking Water Storage	5	5	5	3
Handwashing	6	5	5	3
Sanitation	5	5	6	6
Body Hygiene	6	5	5	2
Malaria	-	-	5	-
Nutrition	-	-	5	-
Cooking/Kitchen	5	5	6	5
Child Care	8	7	5	2
Total	**49**	**44**	**55**	**29**

Table 2 - Table showing main indicators and the corresponding number of sub-indicators for each dataset, as well as those used for this thesis. Main indicators in red were not included in this study.

Evolution of the M&E Surveys and Differences Between Datasets

At different stages of the intervention, the M&E methodology was reviewed and incrementally improved, in order to overcome difficulties with the data collection process and to improve the quality of the collected data. This resulted in three different versions of the HHI being used between the first and last data collection rounds.

The first data collection round (Baseline) used a paper-based form of the HHI. This mode of data collection was upgraded for Midline and Endline to an electronic system using the mobile phone-based survey toolkit known as 'Mobenzi'. This facilitated the data collection process and eliminated errors caused by having to manually transfer data from paper to an electronic format. The Mobenzi system was a 'pay-per-question' system however, that resulted in limitations on the number of survey questions that could be encoded, due to budgetary restrictions. The HHI was subsequently modified a second time for the two PI surveys, in order to take advantage of another mobile survey system known as 'Open Data Kit Collect' (ODK Collect). As this software was open source, no budget restrictions were applied to the length of the questionnaire. This encouraged a rewrite of the survey, resulting in a more logical structure and clearer question and answer formats. Efforts were made to ensure continuity between surveys, however slight differences were introduced that made direct comparisons between datasets challenging.

Furthermore, each data collection round focused on different subsets of the 50 CHCs and corresponded to different numbers of both CHCs and households surveyed per round. An overview of the main differences between datasets and survey types can be seen in Table 3.

Dataset	Baseline	Midline	Endline	Post Intervention I	Post Intervention II
Mode of Collection	Paper Based Survey	Mobenzi	Mobenzi	ODK Collect	ODK Collect
Total Households Surveyed	5745	772	475	502	677
Total CHCs Surveyed	47	30	24	51	25
Average No. of Households Surveyed per CHC	122	26	20	10	27

Table 3 – Overview of the numbers of household and CHCs sampled for each dataset in addition to the survey types employed.

Development of the Primary Research Questions

Upon completion of the cRCTstudy, IPA reported that their findings detected no significant effect of the CBEHPP on the study's primary and secondary health outcomes, while only mixed results were found for its intermediate outcomes. Postive effects were found however for reported household water treament ($p = 0.003$), and the presence of both imporved ($p = 0.017$), as well as structurally complete ($p = 0.046$) sanitation facilities (Sinharoy *et al.*, 2017). This suggested that the CBEHPP intervention in Rusizi had failed to achieve its aims of improving public health through hygiene behaviour change, and that any changes that did take place suffered from a lack of consistency or were either small or statistically undedectable .

In contrast, preliminary analyses carried out by AA after the collection of the Midline and Endline datasets, showed steady increases in the uptake of CHC recommended hygiene practices (Africa AHEAD, 2015a, 2015b). The analytical methods used for this study were relatively basic however and did not include a thorough comparison between datasets over time.

The differences between the two sets of results, combined with the availability of five M&E datasets, spanning the length of the intervention, provided an opportunity for further investigation. By aiming to establish whether or not significant hygiene behaviour change had taken place over the course of the intervention and across the majority of CHCs, a more detailed understanding of the strengths and weaknesses of the CBEHPP's ability to bring about long term consistent change could be achieved.

Building upon the availability of the datasets, it was decided that the best way to achieve this would be to attempt to quantify overall hygiene behaviour change through statistical analyses over time. Furthermore, the question of whether consistent change had taken place could be answered by quantifying the degree of variation between individual CHCs. If both significant change and a significant degree of variation would be detected, further analysis could be conducted to understand the primary differences between CHCs that performed very well and those that showed lower levels of hygiene behaviour change. The results of such a study could then be used to suggest improvements to the CBEHPP training programme as well as AA's M&E programme.

These ideas acted as the basis for the development of the core research questions for this study, which can be summarised as:

I. Did significant hygiene behaviour change take place within Rusizi district CHCs? If so, to what extent, and with how much consistency did these changes take place across observed CHCs?
II. Do common traits exist among high-performing and low-performing CHCs with respect to hygiene behaviour change? If so, can these be linked to elements of the CBEHPP training and M&E programme implementations?

The core research objectives relating to the research questions above, can be found in the executive summary.

Ethical Approval for the Use of the M&E Datasets

Ethical approval was given to Africa AHEAD for the use of all five M&E datasets on the 23rd of August 2017 by the Rwandan National Ethics Committee, for the purpose of conducting process evaluations of the CBEHPP in Rusizi district. AA has given permission to the author of this study as part of its process evaluation efforts.

4.2. Situation Analysis – Public Health in Rwanda

In the year 2000, the Rwandan government's Ministry of Finance and Economic Planning published the 'Rwanda Vision 2020', a development plan that aimed to transform the country into a middle-income economy while lifting its people out of poverty within 20 years. This included a major emphasis on improving the state of the country's health, particularly through cost effective health policies that would target the poorest members of society, while limiting overall population growth. The plan also highlighted the need for

increased gender equality as well as improved clean water access and waste management. A key element for its implementation was the need for good governance through a "small but effective, flexible public sector" that would be capable of "deploying scarce resources" efficiently. This was to be coupled with a decentralised approach, where local communities would play a key role in the making of decisions, as part of a grass-roots level strategy (Ministry of Finance and Economic Planning, 2000).

According to the Vision 2020 document, at the time of its publication in the year 2000, Rwanda had a population of 8.2 million, which was expected to double to 16 million by 2020 without improvements in family planning. It also had a high prevalence of malaria, which was responsible for 40% of hospital consultations, as well as HIV/AIDS, which afflicted 11.2% of the total population. Ambitious goals were set to improve the state of health over the next 20 years, including reducing infant mortality from 107 to 50 deaths per 1000 live-births, maternal mortality from 1070 to 200 deaths per 100,000 live-births, the HIV/AIDS prevalence rate to 8%, and the population growth rate from 3.2% to 2.2% per annum (Ministry of Finance and Economic Planning, 2000).

The First Health Sector Strategic Plan – The Role of Community Based Efforts

These broad goals, were operationalised in 2004 by the MOH's Health Sector Strategic Plan 2005-2009 (HSSP-I) (Ministry of Health, 2004). In particular, the plan outlined the specific challenges and obstacles that needed to be addressed in order to transform the health system and significantly improve the state of the country's health. It also gave shape to the prioritised objectives and strategic interventions required to achieve such a transformation. This included commitments to several international policies such as the United Nations Millennium Development Goals (MDGs) in Health, the New Partnership for African Development (NEPAD) Health Strategy, the Lusaka Declaration on Decentralisation and District Health Systems, and the Abuja Declaration for commitment to health sector expenditure. The MDGs in health consisted of Goal 1 – To eradicate extreme poverty and hunger (malnutrition), Goal 4 – to reduce child mortality, Goal 5 – to improve maternal health, and Goal 6 – to combat AIDS, malaria and other diseases (United Nations, 2003). These commitments would all provide direction, as well as a foundation for the development of the strategic plan.

The key health related challenges identified as part of the HSSP-I health sector performance review included the following:

- That large disparities existed between the poorest and richest quintiles of the population with regards to key health indicators in 2000, particularly with respect to under five mortality (225 versus 120 deaths per 1,000 children respectively), infant mortality (121 versus 70 deaths per 1,000 live-births), under-five mosquito net coverage (0.5 versus 17.9 percent of children), prevalence of underweight children (31.5 versus 13.7 percent of children), and health-professional assisted deliveries (12.1 versus 57.7 percent of births).

- That overall infant, under-five, and maternal mortality rates were reducing at too slow a rate to meet the relevant MDG targets by 2015, while immunisation coverage had recovered to acceptable levels for sub-Saharan Africa since the genocide.

- That most Rwandans were dying from poverty related preventable diseases, particularly transmittable diseases that could be "largely avoided through improved hygiene and behavioural change", and that HIV/AIDS and malaria placed the greatest burden on families.

- That the following diseases were the leading causes of under-five (non-infant) mortality in 2000: Pneumonia - 24%, Diarrhoea - 21%, AIDS - 17%, and Malaria - 8%, and that the leading causes for hospital related morbidity for both adults and children under-five in 2003 were Malaria (contributing to more than 41%), followed by respiratory infections, and then intestinal parasites.

The primary obstacles to addressing these challenges were identified as:

- The health sector receiving only 6.1% of the government's disposable Gross Domestic Product (GDP) at the time, compared to its 2001 Abuja declaration commitment of 15% (Ministry of Health, 2009).

- A debilitating lack of qualified health professionals, particularly doctors, with only one doctor available per 50,000 of the population on average, despite significant improvements to health infrastructure since the genocide. This was seen as a major

contributor to the high maternal mortality rates, due to the lack of obstetric care and relevant expertise.

- Disparities in access to healthcare between geographical regions, as well as between urban and rural populations. Large population exchanges between these regions due to the 1994 Rwandan genocide had also increased HIV contraction rates nationwide. This was compounded by a very low modern contraceptive prevalence rate of 4.3% in 2000, rising to only 10% by 2005 (Ministry of Health, 2009), that also impacted family planning and population control efforts.

In order to overcome these, the HSSP-I developed an approach that was aligned with the MDGs in Health, as well as core aspects of the NEPAD health strategy and the Lusaka Declaration. This included the NEPAD 'strategic axes' of promoting disease control programmes for transmittable and non-transmittable diseases, in particular HIV/AIDS, tuberculosis, malaria, diarrhoea, child pneumonia, and malnutrition, as well as the reinforcement of community-based Information Education Communication (IEC) and Behaviour Change Communication (BCC) efforts. Important aspects taken from the Lusaka Declaration included the promotion of health system decentralisation through a shift in focus to district level health services, and the enhancement of community participation in the management and financing of such services.

This culminated in a strategy centred on seven national programmes that aimed to: (1) increase the availability of human resources, as well as (2) drugs, vaccines, and consumables; (3) improve geographical and (4) financial accessibility to health services; (5) improve the quality of and public demand for disease control services; and (6) improve health infrastructure including national referral hospitals, research facilities, and treatment institutions, while (7) increasing institutional capacity.

In order to address the preventable disease burden and geographic disparities, attempt to cope with the human resource and budget gaps, and promote the decentralisation process, a high priority was given to low-cost, district level community-based initiatives throughout these programmes. This applied in particular to the disease control efforts of programme five, but also to a lesser extent to programmes one, three, and four. Programmes one and three would primarily rely on the 'Community Health Workers' (CHWs) initiative, its aim being to act as the lowest level of community-based support, with thousands of CHWs

providing health sector contact to all districts at the village level. Programme four on the other hand would try to achieve some of its aims through community-based health insurance schemes that would form part of the government's "Mutuelle de santé" initiative (Ministry of Health, 2015).

However, it was through programme five that the majority of community-based efforts would be concentrated. Of the 12 sub-programme components, all but two relied heavily on community-based interventions and support. This included initiatives for the control of transmittable diseases, management of childhood illnesses, nutrition, environmental health, immunisation outreach, response to epidemics, reproductive health, as well as IEC and BCC. The key community-based elements of these initiatives included the large-scale distribution of mosquito nets and anti- malaria, TB, as well as HIV/AIDS IEC training by CHWs; IEC training of communities for family planning, home based care systems, childhood nutrition and disease prevention, as well as the monitoring of epidemics; and an integrated environmental health sub-programme that would focus on improving community hygiene, clean water access, and waste management.

These separate initiatives would go on to act as the precursors of future integrated community-based programmes such as the CBEHPP, with the environmental health and IEC initiatives forming much of the basis. The overall success of the five-year strategic plan would be measured by the reduction of the following key impact indicators: maternal mortality from 1071 to 600 deaths per 100,000 live-births, under-five mortality from 196 to 110 deaths per 1,000 children, infant mortality from 107 to 61 deaths per 1,000 live-births, child malnutrition from 24.3% to 18%, and the HIV prevalence rate to less than 5.1% by 2009.

Birth of the Community Based Environment Health Promotion Programme

By 2008, most of the HSSP-I programmes had been fully implemented and many of its targets already reached. This resulted in the development of a second Health Sector Strategic Plan for 2009 to 2012 (HSSP-II) (Ministry of Health, 2009). According to this, the targets that had already been reached included a decreased infant mortality rate of 62 deaths per 1,000 live-births, an under-five mortality rate of 103 deaths per 1,000 children, a maternal mortality rate of 590 deaths per 100,000 births, and an improved modern contraceptive prevalence rate of 27%. Significant progress had also been made towards

combatting HIV, with a prevalence rate of only 2.8% in 2008, as well as malaria and tuberculosis. This meant that Rwanda was on track to achieve several of the MDGs in health, particularly goals 4 and 6 – relating to child mortality, and combatting HIV, malaria and other diseases respectively. Significant human resource accomplishments had also been made including reducing the doctor to population ratio to 1 in 20,500 and increasing the number of CHWs from 12,000 to 45,000.

However, while the maternal mortality rate had been reduced significantly, it still remained too high to meet either the Vision 2020 target, or the respective MDG by 2015. Additional problem areas still included significant child malnutrition, with 7% of under-fives showing signs of wasting, 24% of being underweight, and 43% of being stunted. The preventable disease burden also remained high, with malaria contributing to 18.7% of recorded under-five mortality, followed by pulmonary infections at 12.8%, and diarrhoea at 11.3%. This was accompanied by worrying indications that the majority of mothers did not seek appropriate medical attention for children suffering from diarrhoea or pulmonary infections, implying that better IEC efforts were required. For the rest of the population, the leading causes of hospital related morbidity were pulmonary infections at 34.1%, followed by malaria at 11.3% and intestinal worms at 10.5%.

While preventable diseases were still the leading causes for morbidity and mortality in Rwanda, these changes presented a significantly different disease and health sector 'landscape' from that which existed in 2004 when the HSSP-I was drafted. Combined with the need to keep up with developments in other sectors and a much evolved decentralised government structure, this resulted in a new set of priorities and the need for the development of a second HSSP. This included the top priorities of improving family health, while improving disease prevention and the promotion of health.

As a result of the successes of the individual community-based initiatives of HSSP-I, and in order to meet these priorities, the HSSP-II proposed the development of a unified Community Based Environmental Health Promotion Programme – the CBEHPP. This would bring all IEC and BCC efforts, along with most environmental health initiatives, under the umbrella of a single standardised programme. It would rely heavily on the newly recruited CHWs, who would be trained by specialised Environmental Health Officers (EHOs) stationed at health centres, with the primary goal being to provide unified health and hygiene training

for all districts of Rwanda. The programme would emphasise in particular the need for combined hygiene behaviour change training and improved Water, Sanitation & Hygiene (WASH) practices. These would include the increased use of hygienic latrines, as well as hand washing with soap, Zero Open Defecation, improved safe drinking water access, and the establishment of Community Health Clubs in every village (Ministry of Health, 2010).

4.3. The Community Based Environmental Health Promotion Programme

The Community Based Environmental Health Promotion Programme (CBEHPP) is an ongoing initiative that according to the MOH's 'Roadmap to CBEHPP' (2010), was launched in December 2009 as part of the HSSP-II to "reduce the national disease burden through community-based hygiene behaviour change and improved sanitation".

The CBEHPP follows the holistic approach of the Community Health Club (CHC) methodology for obtaining positive hygiene behaviour change and mobilising communities at village level. The CHC approach was developed by Dr Juliet Waterkeyn of the INGO Africa AHEAD (AA) in 1995 through her work with community centred development initiatives in Zimbabwe (Waterkeyn and Cairncross, 2005). The CHC model is an approach to community-based health, hygiene and sanitation improvement through the formation of voluntary village level health clubs. It is based on a participatory approach for transferring health knowledge and generating positive behaviour change. The primary goal is to create a 'culture of health' within the community, which is reinforced through positive peer pressure (Waterkeyn and Waterkeyn, 2005).

CBEHPP Mission and Objectives

With an original three-year rollout plan, consisting of three refinement and expansion phases, the CBEHPP's mission was to deliver behaviour change training to villages across all of Rwanda's 30 districts by the end of 2012 (Ministry of Health, 2010), and to have reached at least 15,000 by 2016 (Waterkeyn, 2011b). The first of the rollout phases would consist of adjusting the base CHC approach to the Rwandan specific setting and needs. This would include streamlining the transition from and building on top of existing HSSP-I community-based initiatives, developing culturally appropriate training materials, selecting four 'start-up' districts for initial testing, and training the future Rwandan CHC facilitators. The second phase would use lessons learnt from the first by expanding to a further four districts, while the third phase would involve the gradual expansion to the remaining 22 districts.

At the time, the programme's objectives were aligned with those of the HSSP-II and consisted of the following 'Seven Golden Indicators', that were to be achieved by 2012 by all participating villages (Ministry of Health, 2010):

1. To increase the proportion of hygienic latrines in schools and homes from 28% to 80%.
2. To increase hand washing with soap at critical times from 34% to 80%.
3. To improve safe drinking water access and handling in schools and homes to 80%.
4. To establish Community Health Clubs (CHCs) in 100% of villages.
5. To achieve 100% Zero Open Defecation in all villages.
6. To increase the safe disposal of children's faeces in every household from 28% to 100%.
7. To increase the proportion of households with bath shelters, rubbish pits, pot drying racks, and clean yards to at least 80%.

Available Human Resources and Club Structure

CBEHPP training is delivered to CHCs by a three-level hierarchy of government supported personnel. This is made up of Environmental Health Officers (EHOs), who are based at district health centres and receive the highest level of training in CHC dissemination techniques, followed by social affair mentors known as 'Affaires Sociales' or ASOCs, who are elected from within their communities and are employed by the government for 5-year terms. The ASOCs are in turn assisted by volunteer CHWs, who also come from the same communities, but are unpaid. By 2014 there were approximately 60,000 CHWs nationwide, comprising three CHWs per village on average (Condo *et al.*, 2014).

In Rwanda, CHCs are a type of independent Community-Based Organisations (CBOs) which are formed by their members following initial social marketing and recruitment by EHOs, ASOCS, and CHWs. They are open for both men and women to join voluntarily and are non-religious, non-political, and free of charge. The clubs usually consist of 50 to a maximum of 100 members, with the possibility of multiple clubs being formed within larger villages. Each club has an executive committee consisting of a CHW, as well as an annually elected Chairperson, Secretary, and Treasurer. This helps to classify the club as an official CBO, which can open up the possibility of further government support and micro-credit schemes in the future (Ministry of Health, 2010). Women are particularly encouraged to take leadership

roles within CHCs, especially because they are often the head of their households and the primary caregivers for their families (Waterkeyn, 2011a).

Knowledge Transfer Training Programme and Syllabus

The clubs meet on a weekly basis to learn about, discuss, and debate the implementation of various safe hygiene practices within the community. The initial training programme consists of 20 sessions that are delivered regularly over a minimum period of six months. This length of time is considered necessary to allow for a concrete shift in attitudes towards health and for long-term group cohesion to take place (Waterkeyn, 2011b). The sessions are usually delivered by an ASOC, but experienced CHWs are expected to take over when necessary. EHOs attend CHC sessions from time to time for the purposes of quality control and to provide supervision. Each session covers a different health or hygiene topic and is accompanied by a homework assignment. CHC members are encouraged to share the information from each session and to carry out the homework projects with family members at home. It is the responsibility of the relevant CHC facilitator to monitor attendance and to check that homework assignments have been completed. This is recorded via a membership card that is given to every club member. An overview of the relevant topics and associated homework assignments can be seen below in Table 4.

No.	Topic	Homework Topics
1	Introduction	Bring friends and family. Group mapping of the village.
2	Common Diseases	Demonstrate knowledge of causes.
3	Personal Hygiene	Construction of a family wash shelter.
4	Hand Washing	Construction of a hand washing facility, use of soap.
5	Skin Diseases	Check if children are afflicted by skin diseases.
6	Diarrhoea	Use of soap at home, Oral Rehydration Salts.
7	Infant Care	Correct child immunisation and weaning practices.
8	Intestinal Worms	De-worming of children.
9	Food Hygiene	Safe food storage.
10	Nutrition	Demonstrate knowledge of a balanced diet with 'Road to Health Chart.'
11	Food Security	Kitchen gardens and pest control.
12	Water Sources	Village Level Operation, Maintenance and Management of water resources.
13	Safe Drinking Water	Safe storage and usage, individual cups and plates.
14	Improved Sanitation	No open defecation. Latrine improvement & cleanliness.
15	The Model Home	Waste management and greening.
16	Good Parenting	Clean children, children going to school.
17	Respiratory Disease	Adequate household ventilation.
18	Malaria	Use of treated bed nets.

19	Bilharzia	Treatment for bilharzia.
20	HIV/AIDS	Voluntary counselling and testing.

Table 4: CBEHPP Training syllabus for CHC members including homework topics (Ministry of Health, 2010; Waterkeyn, 2011a).

Monitoring & Evaluation of Hygiene Behaviour Change

The level of hygiene behaviour change exhibited by a CHC that has completed the training programme is measured through what is known as a 'Household Inventory' (HHI). This is a set of observable sub-indicators that suggest whether particular positive health related behaviour traits have been adopted by a given CHC member household as a result of the training. Examples of typical indicators include whether hand washing facilities with soap are present, whether a latrine is present including its quality and level of cleanliness, and whether mosquito nets are available, in use, and well maintained. Unannounced spot observations are carried out by a CHW, who on at least three visits must be supervised by an EHO, of at least 30 randomly selected CHC households. The observations are recorded on a paper- or mobile-phone based code-sheet, and at least two survey rounds must be completed, one before the start of training (Baseline), and one after its completion (Endline).

Post CHC Training and Long-Term Sustainability

Upon completion of the training sessions, members that have attended all sessions and completed all assignments are invited to a graduation ceremony and given a certificate. The clubs are then expected to become increasingly independent and self-sufficient, however can still receive support from the ASOCs and EHOs. With CHC expertise being now prevalent within the community, it is hoped that the clubs will recruit new members and build on their newly found social capital by embarking on new initiatives such as income generating and agricultural projects.

5. Literature Review

The aim of this section is to provide an overview of key literature that will be used to compare results and substantiate arguments presented in the discussion. In order to achieve this, two groups of literature are presented. The first group consists of literature that describes the mechanisms through which behaviour change is induced via the CHC approach, and how these compare to those employed by competing programmes. The second group consists of literature relating to existing CHC programme evaluation studies. This will provide a pool of external results with which the results of this study can be compared.

5.1. Social Mechanisms for Hygiene Behaviour Change

While the objectives, structure, training syllabus and historical development of the CBEHPP have already been described, it is also important to review the mechanisms by which the CHC approach aims to generate long-term behaviour change. This is now discussed, while reviewing two other approaches known as PHAST and CLTS. This will allow for the evaluation of whether or not key mechanisms were disrupted during the Rusizi intervention, as part of the discussion.

While the delivery of health knowledge is an important aspect of many behaviour change efforts, it is clear that this alone is not enough to drive and sustain the society-wide changes required to improve the overall health of target communities (Curtis, 2000; Waterkeyn, 2006). Additional 'social mechanisms' are therefore required to translate knowledge into new beliefs, social-norms, and individual action throughout the community.

According to Waterkeyn et al. (2006), most contemporary hygiene behaviour change strategies fall into the following categories when grouped by their primary methods for triggering change:

1. **Regulation and Control:** Where change is imposed by authority via health inspectors and a system of rewards and fines for compliance versus non-compliance.
2. **Health Belief**: Where change is expected to naturally occur once a target community fully understands and believes in the reasons and benefits for doing so.
3. **Status Appeal:** Where social status is used as the basis for marketing hygiene related products and services.

4. **Collective Shame:** Where emotions such as disgust and shame relating to unhygienic practices are used to elicit a simultaneous community-wide resolve to change.
5. **Group Consensus**: Where individuals are expected to change only if everyone around them is doing so as well.

Of these, groups one and three apply to 'top-down' government intensive interventions, and the commercial tactics of so-called 'Public-Private Partnerships', respectively. Both of which differ substantially from 'bottom-up' community-based approaches and are therefore not comparable to CHCs.

Group four however, corresponds to a widely used comparable methodology known as Community Led Total Sanitation (CLTS), which focuses exclusively on sanitation. Its ultimate goal is to achieve Zero Open Defecation, as well as widespread latrine construction and hand-washing with soap throughout its target communities (Whaley and Webster, 2011). It capitalises on the idea of communal 'self-respect' in order to create demand for a 'total sanitation' overhaul by the community. This is achieved through an initial intervention day, where facilitators employ tactics such as the use of crude language to describe human faeces, mapping open defecation sites, and demonstrating the extent of faecal contamination throughout the community, to elicit feelings of disgust and shame. This is meant to 'trigger' the community into action, whereupon group appraisal exercises are used to help the community to take ownership and come up with their own road-map to achieving zero open defecation, and the comprehensive construction of latrines. Change is maintained through negative peer-pressure, which is bolstered through further 'post-triggering' visits after several months, where the desired attitudes towards poor sanitation are reaffirmed, and support is given to on-going initiatives. Village-level entrepreneurship and the communal gathering of funds is encouraged for the procurement of latrine building materials, however subsidies are not given out as part of the programme (Peal, Evans and Van der Voorden, 2010).

In contrast, the CHC approach's primary methods correspond to groups two and five, whereby according to (Waterkeyn and Waterkeyn, 2013), its core mechanisms consist of the comprehensive delivery of knowledge for fostering understanding, the generation of new values and norms through group consensus, and effecting individual change through positive

peer pressure and the support of 'self-efficacy'. Self-efficacy in this context means an individual's confidence in their ability to carry out a planned action and effect change.

Theoretical Basis for Hygiene Behaviour Change through the CHC Approach

These mechanisms are modelled on several psycho-social theories of community development including the Health Belief Model (Janz and Becker, 1984), the Theory of Reasoned Action and Planned Behaviour (Ajzen and Fishbein, 1980), Social Learning Theory (Bandura, 1986), and the concept of Social Capital (Kawachi and Berkman, 2003).

The Health Belief Model states that health related action is only taken by an individual if they believe that such an action would overcome a particular health problem, and thereafter only when they believe that they are capable of taking such action themselves (University of Twente, 2017). This relates to the CHC mechanism of translating knowledge not only into understanding, but also into well ingrained beliefs, while providing enough support for an individual to act on them (self-efficacy).

The Theory of Reasoned Action and Planned Behaviour is a theory that predicts deliberate behaviour. It suggests that this aspect of behaviour is governed by an individual's attitude or beliefs towards a given behaviour, their self-perception of their ability to perform the behaviour (similar to self-efficacy), their intent (or motivation) to practice the behaviour, and their 'subjective norms' - their perception of how other people, whose opinions they care about, would react to them practising the new behaviour. This relates strongly to the concept of behaviour being driven by a combination of personal beliefs, external 'social pressure', and societal acceptance (Neighbors, Foster and Fossos, 2013). This contributes to the CHC mechanisms of positive peer pressure and group consensus for the development and long-term acceptance of new norms, values, and behaviours.

Social Learning theory focuses on how people learn new behaviours from others. It proposes that behaviours are learnt through the continually reinforced mechanisms of observation, imitation, and modelling. Whereby the attitudes and behaviours of others (including outcomes) are first observed, then, if sufficient motivation is present, imitated, and finally 'processed' for the purpose of acting as a model for future individual action. The theory suggests that human behaviour is driven by continuous feedback between cognitive, social, and environmental influences (Bandura, 2017). This suggests that new behaviours are best learnt in a group environment, with regular meetings and high levels of active participation,

interaction, and motivation. This contributes to the design of the CHC weekly sessions and the knowledge transfer methods employed to introduce new behaviours to the community.

Finally, the concept of Social Capital centres on the idea that increasing the number and quality of social relationships within a community can add value through enhanced levels of interpersonal trust, support, and reciprocity, that in turn can aid collective action for the benefit of all (Kawachi, 1999). This is a central tenet behind the idea of forming Community Health Clubs that have a structured membership, which meets regularly over a sufficiently long period of time.

Practical Implementation through the CHC Approach

The theoretical features and mechanisms above are put into practice via the following aspects of the CHC approach:

Motivation: Generating and maintaining motivation is a key element of both the initiation of widespread behaviour change and its long-term sustainability within the community. CHCs achieve this through various means, including satisfying the desire to prevent diseases, creating the opportunity to gain knowledge and achieve recognition upon graduation, and fostering a sense of belonging and community cohesion through commitment to the common goals of the club (Whaley and Webster, 2011). The chance to socialise, have fun, and share ideas, is of course also an important factor. Long-term motivation is maintained through household visits between CHC members and by official M&E facilitators. This generates a 'Hawthorne Effect', whereby positive behaviours are maintained due to the possibility of being observed by others (Waterkeyn and Waterkeyn, 2013). Future CHC phases only available to CHC graduates, such as a 'sanitation phase', where material inputs may be provided for permanent latrine construction, or income-generating phases, act to maintain interest and keep CHCs alive for many years (Waterkeyn and Waterkeyn, 2005).

Transferring Knowledge into Beliefs: Comprehensive health education is delivered through a modified version of another behaviour change methodology known as 'Participatory Hygiene and Sanitation Transformation' (Waterkeyn and Waterkeyn, 2005), which in turn draws on the problem-solving strategies of a participatory planning tool known as SARAR (Self-Esteem, Associative Strengths, Resourcefulness, Action-Planning, and Responsibility) (Peal, Evans and Van der Voorden, 2010). PHAST delivers knowledge through a 'bottom-up'

approach whereby a 'toolkit' consisting of 300 culturally adapted ethnographic illustrations is first used to introduce particular health related topics (Waterkeyn, 2006). This is then followed by a seven-step process that aims to maximise the engagement of community members in the process of identifying, analysing, and planning how to solve local problems relating to each topic (including community run M&E). A variety of exercises is used at each stage to gradually increase the level of individual participation and confidence. This includes the telling of community stories, the anonymous sharing of existing hygiene practices, and designing charts to illustrate disease transmission routes, discuss health improvement plans, and to prioritise activities (Wood, Sawyer and Simpson-Herbert, 1998). The ultimate goal is to arrive at a group consensus regarding the community-wide strategy for solving each problem.

This process is then encapsulated by the CHC's structured membership, regular meetings, and M&E system, in order to provide the optimum conditions for changing community-wide norms and values through group consensus, maximising the potential for individuals to assimilate new behaviours, and supporting self-efficacy as well as the maintenance of positive behaviour change through positive peer-pressure.

5.2. Existing CHC Programme Evaluation Studies

In addition to the IPA paper by Sinharoy et al (2017), a literature search for existing CHC evaluation studies was conducted in order to gather information with which to compare the results of this study. Three primary papers were identified as having study designs and results that most closely matched that of this study. These consisted of the paper by Whaley and Webster (2011), that compared the effectiveness of CHC and CLTS implementations in Zimbabwe, followed by Waterkeyn and Cairncross (2005), which evaluated the uptake of improved hygiene practices within CHCs in two districts of Zimbabwe. A follow up study by Waterkeyn and Waterkeyn (2013), focusing on three districts, including the two previous ones, evaluated the importance of health knowledge as a motivating factor for behaviour change.

Whaley and Webster (2011) conducted a comparative study between two demand driven sanitation and hygiene approaches, namely CLTS and CHCs, in Zimbabwe. The study was carried out across the three districts of Chiredzi, Chipinge and Mutoko. Chiredzi contained both a CHC and a CLTS intervention, while Chipinge only had a CHC intervention, and

Mutoko a CLTS intervention. The purpose of both interventions happening in the same district was to evaluate the effectiveness of each approach. In the districts with separate interventions, the scope was to evaluate the sustainability of the interventions.

The results of the effectiveness study showed that CHCs were more effective than CLTS at encouraging people to dispose of their faeces through cat sanitation, and at having a higher number of individuals owning a handwashing facility. The CLTS respondents however had a greater likelihood of owning or sharing a latrine than the CHCs respondents, who did not share latrines at all. The sustainability study also showed that CHCs respondents owned more handwashing facilities than CLTS respondents. Significantly more CLTS respondents were found to have built their own latrines, however a greater number of them were without a lid in comparison to the CHC respondent's latrines. The study concluded that CHCs overall were more effective than CLTS at encouraging handwashing, but that CLTS was more likely to encourage latrine construction. A proposition was made by the authors for the two approached to be combined as they seem to complement each other (Whaley and Webster, 2011).

Waterkeyn and Cairncross (2005) carried out a comparative analysis between two different districts in Zimbabwe, namely the Tsholotso and Makoni districts, with 32 and 265 CHCs each, respectively. Qualitative and quantitative surveys were conducted between 2000 and 2001 for 25 CHCs from each district, with 15 CHC members selected per CHC for the study. Control groups were formed by selecting non-CHC villages in neighbouring areas in order to ensure similarities among demographic, cultural, and socio-economic factors between the two groups. The results of this study showed large statistically significant differences ($p < 0.01$) between the control and intervention groups for both districts. For example, in Tsholotso (where the biggest differences were observed) the following differences were observed for the following primary hygiene indicators: practice of zero open defecation – 100% versus 2%, latrine construction – 42% versus 3%, use of handwashing facility – 74% versus 39%, and use of handwashing facilities with soap – 39% versus 20%, for CHC members versus control group members respectively. Qualitative results included the observation that CHC members continued to meet throughout the rainy season, despite agricultural pressures, the motivation to attend sessions being derived from the desire to socialise and learn interactively with others, and the observed empowerment of women. The study

concluded that a high level of demand for proper sanitation was generated as a result of the CHCs in both districts (Waterkeyn and Cairncross, 2005).

Waterkeyn and Waterkeyn (2013) is a further extension of the Waterkeyn and Cairncross (2005) study, with the inclusion of a third district (Gutu district) in which CHCs had been formed. The study looked into the uptake of positive hygiene practices by CHC members compared to a control group, and explored the importance of health knowledge as a motivator for behaviour change. The study showed a statistically significant ($p< 0.05$) difference in the uptake of positive hygiene behaviours (as indicated by proxy indicators) for CHC respondents who had attended all CHC sessions - thus having received 'full knowledge', when compared to non-CHC respondents who were categorised as only having 'some knowledge' on factors associated with diarrhoea and disease transmission (Waterkeyn and Waterkeyn, 2013).

6. Methodology

In order to bridge the gap between moving from the available raw data to the generation of results for answering the primary research questions, this section aims to provide a framework for understanding the main steps taken during this study. The methodological steps are aligned with the four research objectives outlined in the executive summary. These are: (1) To quantify overall hygiene behaviour change for the surveyed CHCs; (2) to establish the robustness and limits of the data for the purpose of drawing general conclusions regarding CHC level (population level) change based on observed sample level change; (3) to rank CHCs based on performance in order to establish whether common traits exist among top- and low-performing CHCs; and (4) to understand why such variations in performance might exist between the two groups.

6.1. Data Preparation Process and Indicator Alignment

In order to perform quantitative analyses on all five datasets, a significant portion of the work done for the study was put into the preliminary steps of pre-processing the data for the purpose of converting it into an unbiased numerical format that would also allow for cross dataset comparisons. The process consisted of three main steps, namely data cleaning, aligning datasets and individual hygiene behaviour change indicators across surveys, and designing an appropriate numerical scoring system.

1. The first step involved 'cleaning' each dataset in order to remove any spurious and nonsensical entries that were the result of human error during the data collection or data entry processes. Examples include entries where multiple answers had been given to survey questions requiring only one answer, or where household surveys had been assigned incorrect CHC labels, therefore making it impossible to tell which CHC the associated household belonged to. Such errors were most common for the Baseline dataset, where large numbers of paper-based surveys (corresponding to 5745 observed households) had been manually entered into an electronic database.

   The cleaning process was carried out by first importing the datasets into the SPSS 24.0 Grad Pack (student edition) software package, where individual data types such as sub indicators, CHC names, and attendance levels, were stored as variables. A process of elimination was then carried out using a range of tools for descriptive statistics, such as frequency analysis and cross tabulation, to systematically break the data down by groups of variables, in order to identify gaps and suspicious patterns. These troublesome areas were then inspected manually in order to find and purge individual spurious entries. Such an approach was required in order to efficiently sort through the large numbers of entries per dataset.

2. Next, the challenge of overcoming the differences between each of the three survey types mentioned in Table 3, was met by developing a rigorous filtering process that was applied to each sub-indicator in all five datasets. The process aimed to only allow sub-indicators that appeared in all three survey types, which additionally possessed approximately equivalent question and answer formats, into the study. The sub-indicator questions and answers also had to be capable of clearly indicating whether selected answers corresponded to positive or negative hygiene practices.

   The filtering process was carried out by first using the PI questionnaire as a template for systematically searching through the other surveys in order to find equivalents to each of the PI sub indicators. The PI survey was chosen because it contained the most up to date survey questions, had the most logical structure, and contained the largest number of sub-indicators, as shown in Table 2 in the introduction. As shown in Table 2, it is clear that the lack of Baseline and Midline sub-indicators for the main indicator categories of

Housing, Malaria, and Nutrition, meant that these categories were automatically excluded from the study.

The equivalence of individual sub-indicators across surveys was then evaluated on a case by case basis. Each set of similar sub-indicators had to meet the following requirements:

i. They needed to be similar enough to be represented by a single common question that could be answered in an unbiased and equivalent way by each of the original indicator responses.
ii. They had to be convertible to 'polar questions', such that answers could be clearly split into positive ('yes') and negative ('no') categories in order to clearly state whether a recorded observation was considered to reflect a positive or negative behaviour trait.

This process was highly challenging and could only be achieved by being very strict with the equivalence criteria, while applying an iterative approach to developing the new set of common questions in such a way as to find the best compromise between all sub-indicators. The requirement to balance all questions and multiple answers per indicator in such a way as to not distort any of them resulted in many indicators being filtered out. Despite this, 29 sub-indicators ultimately made it through the selection process and into the study. This included between two and six indicators per main indicator, as shown in Table 5.

Indicator 1: Compound	Indicator 2: Water Source	Indicator 3: Drinking Water Storage	Indicator 4: Handwashing
1a: Is there sufficient drainage?	**2a:** Does the household use a safe primary water source?	**3a:** Is drinking water stored in a sealed container?	**4a:** Is there a handwashing facility available?
1b: Is the compound swept clean?	**2b:** Is the walking distance to the water source 30 minutes or less?	**3b:** Is the drinking water storage container clean?	**4b:** Is there a handwashing facility of good enough design?
1c: Is there no waste seen around the house?	**2c:** Is the waiting time at the water source 30 minutes or less?	**3c:** Is drinking water treated?	**4c:** Is there soap at the handwashing facility?
1d: Is there solid waste management?	**2d:** Are there 15 litres or more of water available per household member each day?		

Indicator 5: Sanitation	Indicator 6: Body Hygiene	Indicator 7: Cooking	Indicator 8: Child Care
5a: Does the household have access to a latrine?	**6a:** Is there a designated area for bathing?	**7a:** Is cooked food stored safely?	**8a:** Are the children wearing clean clothes?
5b: Does the household not share a latrine with other households?	**6b:** Is soap available for bathing?	**7b:** Is the cooking done in a designated kitchen area?	**8b:** Do the children have clean faces?
5c: Does the household have an improved latrine?		**7c:** Is safe fuel used for cooking?	
5d: Is zero open defecation practised?		**7d:** Is the cooking area not contaminated from livestock?	
5e: Is the latrine well covered?		**7e:** Is the kitchen clean?	
5f: Is the latrine clean?			

Table 5 – Detailed overview of the aligned indicators and their sub-indicators.

3. Finally, following the development of the polar question system, where all sub-indicator responses would correspond to either a 'yes' or a 'no' answer, a numerical scoring system was developed that would assign binary values to each answer. A value of 1 was assigned to 'yes' (positive) answers, while a value of 0 was assigned to 'no' (negative) answers. This would make it possible to compute averages across sub-indicator groups in order to calculate overall primary indicator scores for each household. In turn, the primary indicator scores could be averaged to give an overall performance score for each household. This would result in scores that would lie between 0 and 1, which could then be expressed as a percentage. This meant that households with an overall score of 1 for example, would therefore be practising 100% of all 29 recommended positive hygiene practices.

   This system was considered to be the simplest possible scoring system, with the least likelihood of introducing artificial errors and bias into the study. Alternative systems were considered, such as ones where indicators most directly related to primary health outcomes (for example indicators for handwashing and sanitation), would be more heavily weighted than others. However, these were dismissed, as they could introduce artificial bias and would go against the core aim of the CHC methodology to generate holistic hygiene behaviour change across all indicator categories.

   One source of possible error or bias that was recognised for the chosen system, was the inconsistent number of sub-indicators for each of the main indicator groups. By averaging over 6 sub-indicators, as in the case of indicator 5 (sanitation), it was possible to generate a wider range of possible averages using only ones and zeros, than it was with only 2 sub-indicators, as in the case of indicator 6 (body hygiene). This meant that different indicators could correspond to different 'resolutions' of computed averages. This issue was partially overcome by computing overall averages using the 29 sub-indicators combined, and by computing indicator level averages across large numbers of households.

6.2. Quantifying Overall Hygiene Behaviour Change

The analysis of indicator performance scores was carried out using a combination of the SPSS 24.0 statistical analysis software package, and the R programming language. R was utilised in order to overcome the licensing restrictions imposed on SPSS for the student license available to the author. The restrictions included limited access to advanced statistical analysis sub-packages. The use of R also came with challenges however, as significant time was required to learn the language and to debug program scripts.

Computing Performance Scores

Following the data preparation process and numerical scoring system implementation, all five datasets were combined into a single database that was aligned using the new hybrid sub-indicators. A schematic of the data is presented in Figure 3, where an overview of the data's structure can be seen.

Dataset	Household	CHC	Indicator 1a	Indicator 1b	...	Indicator 1 Mean	...	Indicator 8c	Indicator 8d	Indicator 8 Mean	Overall Performance
Baseline	HH1	Kabeza	1	0	...		...	1	1		0.75
Baseline	HH2	Kabeza	1	1	...		....	0	0		0.93
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Baseline	HH23	Munini	0	1	...		...	1	0		
Baseline	HH24	Munini	1	1							
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Midline	HH 5746	Kabeza	0	0							
⋮	⋮	⋮	⋮	⋮							
PI – I	HH 7071	Kabeza	1	1							

Figure 3 - Combined dataset schematic. Each row corresponds to a single household survey. Horizontal dots denote horizontally repeating patterns, while vertical dots represent vertically repeating ones. Blue and red arrows illustrate how primary indicator, and overall performance scores are calculated using sub-indicator means, respectively. Green and purple arrows illustrate the groupings used to compute dataset level and CHC level statistics, respectively.

Using the schematic in Figure 3 as a guide, absolute scores for primary indicator and overall performance per household were computed by taking column-wise (horizontal) averages of individual sub-indicator scores. Clustered row-wise (vertical) groupings could then be used for computing dataset and CHC level statistics. In turn, these allowed for relative performance comparisons between Baseline clusters and those in other datasets, for the purpose of quantifying behaviour change over time.

The analysis procedure for quantifying overall hygiene behaviour change consisted of first conducting dataset level analyses of overall household performance scores and establishing whether statistically significant change had taken place between datasets. This was followed by CHC level performance analyses to establish the degree of behaviour change of individual CHCs. Steps to establish the robustness and limits of the data were conducted throughout the process.

Dataset Level Analyses

1. The first dataset level analysis was conducted by clustering overall performance scores by dataset and computing various summary statistics for each group. This consisted of the mean, median, sample standard deviation and standard error (about the mean) of the clustered performance scores within each group. The use of the formula, $s = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$, where s is the sample standard deviation, $\bar{X}$ is the sample mean, X represents each score, and n is the sample size, allowed for the interpretation of how well the sampled data reflected the likely true population mean. This was achieved through analysing the standard error, calculated as $\frac{s}{\sqrt{n}}$. In this way, it was possible to state how well the dataset level statistics (based on observed samples), reflected the likely performance of all CHC households within the surveyed CHCs.
2. Next, the distribution of performance scores was explored for each dataset using a combination of box plots, histograms, and kernel density plots.
3. The question of whether statistically significant hygiene behaviour change had occurred between datasets was answered by conducting 'independent samples t-tests' between the clustered performance scores of each dataset. The t-tests produced estimates for the differences in means, their statistical significance, and confidence intervals for the location of the true differences between the sample populations of each dataset pair.

   The appropriateness of the choice of the t-test as the main tool for establishing the degree of statistically significant change was verified in the following manner: According to Laerd statistics (Lumley *et al.*, 2002), reliable application of the t-test requires several assumptions to be satisfied. These assumptions are:

i. That the dependent variable (overall performance) is measured on a continuous scale – including evenly spaced 'interval' data. This was satisfied by the averaging procedure across all of the 29 sub indicators, with the scale ranging from 0 to 1.

ii. That the independent variables consist of two categorical, independent groups. The groups in this case were the dataset pairs for comparison. Independence was guaranteed by the fact that each data collection round took place during different periods of time.

iii. That observations are independent. This was guaranteed by the fact that households were randomly selected for observation within each CHC, thereby minimising the likelihood that the same household was surveyed twice. The separation over time also added to this.

iv. That the number of significant outliers should be minimised. Boxplot analysis did show that outliers existed, however that they were much fewer in number than the sample sizes of the associated datasets.

v. That the dependent variable should be approximately normally distributed within each group. The distributional analysis in step 2 above, showed that this was the case for Baseline, and partially for Midline, the other three datasets showed distinctly non-normal distributions. According to Lumley et. al (2002) however, due to the Central Limit Theorem, which states that samples with large sample sizes (bigger than 30 as a rule of thumb) can be considered to be approximately normal in behaviour, this would not present a problem for large sample sizes. This assumption was therefore not violated by any of the datasets due to their large sample sizes.

vi. That there needs to be a homogeneity of variances between the two groups. In each case, a Levene's test for the homogeneity of variances was conducted between each group prior to the application of the t-test. Any dataset pairs that failed this test, were then subjected to a modified t-test known as the Welch's t-test, where this assumption is not required. In general, both tests are considered to be roughly equal with regard to their level of statistical power. This procedure was applied throughout the study, whenever t-tests were carried out.

CHC Level Analyses

1. The first step of the CHC level analysis was to assess the robustness of household data, clustered by CHC, for the purpose of conducting statistical comparisons between

clusters. This was done by first computing sample size statistics for each dataset in order to understand quantities such as the average sample size per CHC and the degree of variation of sample sizes about their respective means. The results, indicated that a wide range of variation existed both within and between different datasets. On the whole however, all of the datasets showed positive attributes that balanced out their negative ones. This suggested that reasonable CHC level statistical analyses could be carried out.

2. Next, household data in each dataset was clustered by CHC, and the distribution of absolute overall performance per CHC was investigated using boxplots. This was partially conducted for the purpose of evaluating the appropriateness of conducting t-tests between CHC clusters, following the procedure mentioned in the preceding section. Most clustered datasets met the requirements of being reasonably normally distributed, non- heavily skewed, and possessing relatively few outliers. A few datasets had very narrow distributions owing to very low numbers of households however, and were excluded from subsequent analyses.
3. Relative CHC level behaviour change was assessed by conducting t-tests between Baseline CHC clusters and equivalent CHC clusters in the other datasets. In this way, individual CHCs could be tracked over time, and their relative levels of behaviour change quantified. The same procedure, as mentioned in the preceding section, was conducted to test for an equality of variances before subjecting paired CHC data to either the independent samples- or Welch's t-test. Results relating to differences in mean performance, statistical significance, and 95% confidence intervals for the difference in means were all used to produce dot-plots showing relative overall hygiene behaviour change per CHC.
4. These results were finally ranked by overall behaviour change, before being split into high-, mid-, and low-performing groups based on where they fell based on the calculated quartiles of the behaviour change results. High and low performing CHCs fell into the upper and lower quartiles respectively, while mid-performing ones fell into the Inter Quartile Range.

6.3. Identifying Common Traits and Behavioural Patterns

Identifying Common Traits

1. The final phase of the study was to use the groupings established at the end of the previous section to compare indicator level behaviour change between groups. This was

achieved by using the CHC names in each group to only select corresponding households within both the Baseline dataset and the dataset to which the grouping corresponded.

2. Next, t-tests were conducted for each main indicator category between individual groups in each dataset. Household data was not segregated by CHC within each group and resulted in large sample sizes that would result in reliable application of the t-test.
3. Finally, this procedure resulted in three sets of results (one per performance group) for each of the 8 indicators per dataset. These were then visualised as grouped bar-charts for visual analysis.

Understanding Common Traits and Patterns

In addition to comparing the primary results above to literature sources and externally available information, three sets of secondary results were produced in order to explore the common traits found through the preceding step.

1. The first set of results were produced by carrying out Pearson's correlation tests between each set of main indicator scores within each dataset. This included tests for statistical significance for each pair of indicators. This was carried out by supplying the entire column (all households) for each of the 8 main indicators as an input to SPSS.
   A possible source of error or bias that was identified, was the fact that such correlation tests usually require the supplied data to be truly continuous. By comparing individual indicator columns whose values had only been calculated from their associated sub-indicators (between 2 to 6 sets of ones and zeros), these values could not be considered 'highly' continuous. This was partially offset by the large number of entries per column, which resulted in statistically significant, if not somewhat coarse correlations.
2. The second set of results consisted of investigating whether any pattern could be seen in the CHC session attendance levels of each of the groups. This was produced by computing means and standard deviations for the attendance level values recorded for each household within the different groups.
3. The final set of results corresponded to an inconclusive exercise where it was attempted to track how CHCs moved within and between groups from one dataset to the next. This proved to be difficult as the selection of CHCs within most datasets (except for Baseline and PI-I) were never identical. This meant that only four CHCs ever came up in all five datasets, making it difficult to draw general conclusions in this regard.

7. Results

The findings of this study are organised into two sections in order to address the primary research questions, namely, whether significant hygiene behaviour change took place within Rusizi district CHCs, and whether common traits exist among high- and low-performing CHCs with respect to hygiene behaviour change. The key features of each result are highlighted, particularly those that play an important role in the discussion.

7.1. Quantifying Overall Hygiene Behaviour Change

The first set of results, presented in Table 6, give an overview of key summary statistics for overall household performance per dataset. The values shown correspond to the proportion of recommended hygiene practices that an observed household is performing, as defined by the indicators outlined in Table 5.

Dataset	Sample Size (n)	Mean (%)	Median (%)	Std. Deviation (%)	Std. Error of the Mean
Baseline	5745	63.5	64.0	12.9	0.170
Midline	772	64.9	66.9	15.5	0.556
Endline	475	80.2	82.4	10.3	0.475
Post Intervention I	502	79.8	82.3	14.8	0.659
Post Intervention II	677	85.8	90.2	13.8	0.529

Table 6 – Statistical summary on overall household performance for each dataset.

The results show a general increase in mean household performance over time, with modest improvement from Baseline to Midline, followed by a substantial increase of 15.3% from Midline to Endline, then a slight decrease of 0.4% from Endline to PI-I, and finally a 6% increase by PI-II. Median performance increases in a similar fashion, however an ever-increasing gap between median and mean values can be observed for each dataset as time progresses. The standard error of the mean is calculated by dividing the standard deviation by the square root of the sample size. This is a measure of dispersion of the sample means around the population mean for each dataset. The results show that all datasets have a relatively small standard error, therefore indicating that dataset level means of observed CHC households most likely reflect the means of all CHC households (the population mean) fairly well. It is also important to note the relatively high Baseline mean of 63.5% and associated low standard error of 0.17%.

These results are complemented by those in Figure 4, which shows the distribution of overall household performance per dataset through a notched box and whisker plot. The shaded

region within each box corresponds to the Inter Quartile Range (IQR), where the majority of results lie, while the whiskers show the $IQR \pm 1.5 \times IQR$, and contain 99.3% of results. Any outliers that fall outside of this range are represented by individual black dots, while the vertical width of the notches correspond to the 95% confidence interval about the sample median, shown as a horizontal black line found at the apexes of the notches. The confidence interval shows the region where there is a 95% chance of finding the true population median based on the observations. Vertical overlap between notches can be used as an informal comparison of whether or not samples are statistically similar to each other.

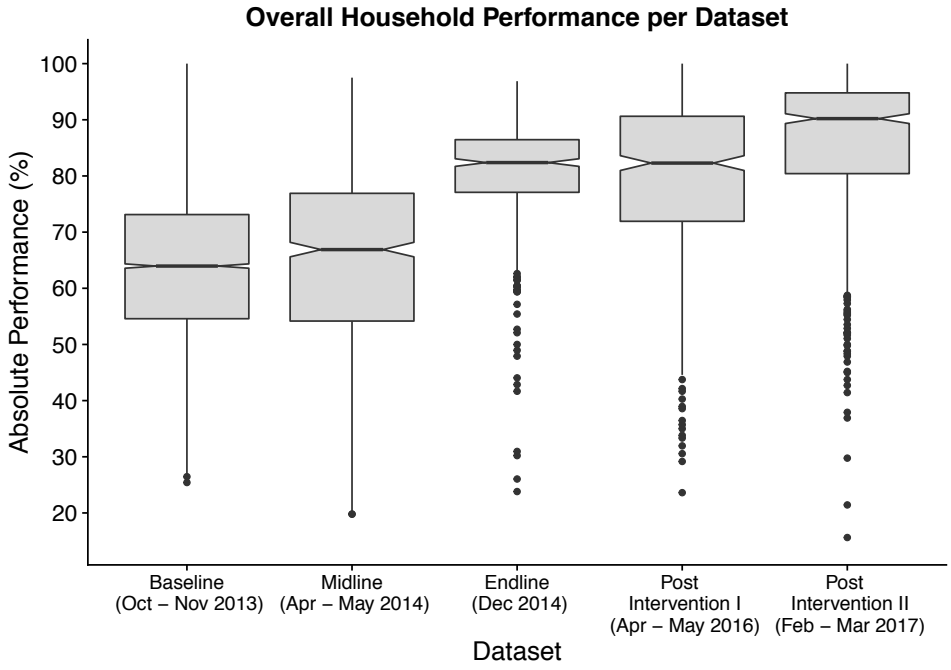

Figure 4 – Notched Box and Whiskers plot showing distribution of overall household performance per dataset.

Beginning with the boxed regions, an upwards shift of the IQRs over time can be seen. While the Midline region mostly overlaps with that of Baseline's, the shifted upper half and non-overlapping notches imply that statistical change is starting to be seen within most households. This is particularly distinct for the Endline result, where the bottom of the IQR is above the tops of those for both Baseline and Midline, implying that major behaviour change has taken place for the majority of observed households. Vertical overlap between the IQRs and notches of Endline and PI-I, however suggest a lack of statistical change, more than one year after the end of the training programme. A further upwards shift of the entire IQR can then be seen for PI-II, which also shows a non-uniform. Furthermore, with an IQR that ranges from approximately 55 to 75%, the Baseline result suggests that the majority of households were already practising over half of the recommended practices prior to CHC

intervention. Finally, the whiskers show that all datasets contain high performing households, and that from Endline onwards, all lower whiskers have moved upwards, leaving a finite 'tail' of outliers.

More formal results relating to whether statistically significant change in overall performance took place between datasets are presented in Table 7. These were produced by conducting 'independent-samples t-tests' between datasets. Comparisons were made between Baseline and all other datasets, in addition to sequential comparisons between neighbouring datasets. The results shown include the difference in means between datasets, the t-value, which is a measure of how strong the difference between groups is in relation to the variability (or noise) within each of the groups (calculated as a 'signal to noise' ratio), and the p-value, which indicates whether the difference is statistically significant. The confidence interval has the same units as overall performance and is computed using a combination of the standard error, t-value, and sample sizes of each group (to give degrees of freedom). Typically, a p-value that is less than or equal to 0.05 is considered significant and a t-value much larger than 1.0 indicates a strong difference between groups.

	Difference in	Std. Error of			95% Confidence Interval of the Difference (%)	
Dataset Comparison	**Means (%)**	**the Difference**	**t-value**	**p-value**	**Lower Limit**	**Upper Limit**
Baseline – Midline	1.46	0.582	2.52	0.0120	0.323	2.61
Baseline – Endline	16.8	0.505	33.2	p<<0	15.8	17.8
Baseline – Post Intervention I	16.3	0.681	23.9	p<<0	14.9	17.6
Baseline – Post Intervention II	22.3	0.528	42.2	p<<0	21.3	23.4
Midline – Endline	15.3	0.731	20.9	p<<0	13.9	16.7
Endline – Post Intervention I	- 0.484	0.813	-0.596	0.552	-2.08	1.11
Post Intervention I – Post Intervention II	6.04	0.836	7.22	p<<0	4.40	7.68

Table 7 - Independent t-test results on overall household performance between datasets.

The results confirm that statistically significant change took place between Baseline and all other datasets, however that the change between Baseline and Midline is modest, as shown by the relatively small effect size (t-value). While both Endline and PI-I show a similar difference in relation to Baseline (~16%), the Endline t-value is much stronger, indicating less spread and more consistent performance overall when compared to PI-I. Results in the lower half of the table confirm that the strongest sequential change happened from Midline to Endline and that statistically significant change could not be detected between Endline and PI – I.

A further exploration of the variation of overall performance per dataset is made possible via the Kernel density plot in Figure 5. The plot is similar to a frequency histogram, however instead of plotting a 'count' of discretised performance scores, the continuous probability density of finding scores within particular 'bandwidths' of performance are plotted instead. The plots are superior for showing distribution shapes, by smoothing out sharp variations, while making it easy to compare distributions composed of very different sample sizes.

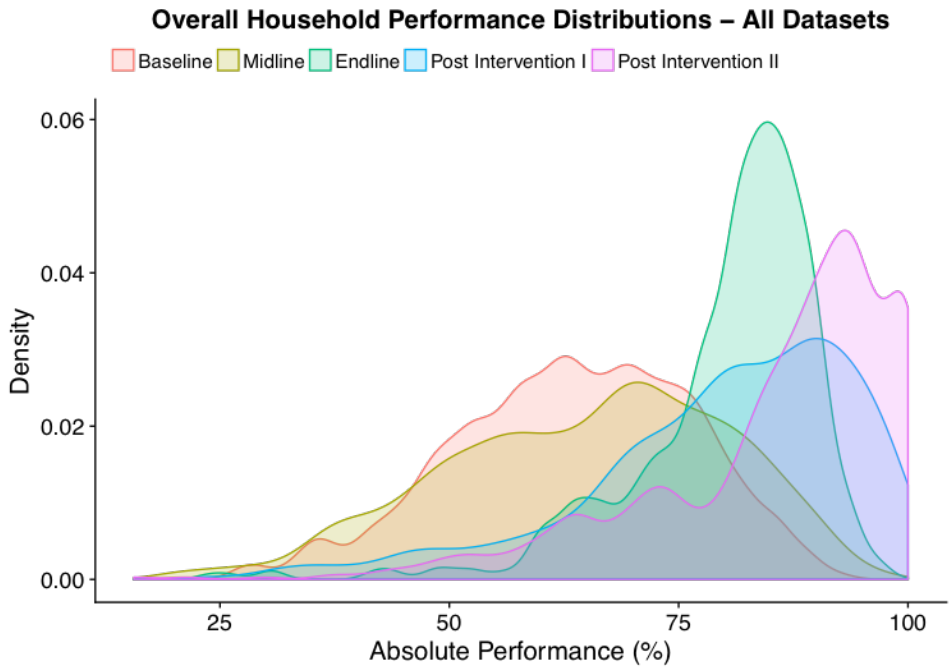

Figure 5 – Kernel density plot of overall performance (%) of all households surveyed all five datasets.

The plot clearly shows a gradual rightward shift of the post-Baseline distribution peaks over time. Beginning with a classically symmetric normal distribution shape in Baseline, the Midline dataset shows an equally broad rightward-leaning distribution, before the striking consolidation of the Endline dataset about the location of its peak, past the 75% mark. It is clear that most of the body of the Endline distribution has traversed rightwards, resulting in a relatively 'thin' tail and a narrow, tall peak. While a further rightward increment of the PI-I peak can be seen, this distribution is broader, however with most its area horizontally overlapping with that of Endline's. Finally, the PI-II distribution shows a similar peak to that of Endline, however, with substantial saturation at 100%.

Clustered Results - CHC Level Behaviour Change

Table 8 presents summary statistics relating to the analysis of CHC level sample sizes. The results show that there is a large amount of variation between datasets with respect to the number of CHCs surveyed, the mean and median number of households surveyed per CHC, and the standard deviations of households surveyed about their respective means.

		Households surveyed per CHC		
Dataset	**Total CHCs surveyed**	**Mean**	**Median**	**Std. Deviation**
Baseline	47	122	123	57.4
Midline	30	25.7	16.5	25.0
Endline	24	19.8	18.5	14.7
Post Intervention I	51	9.84	10.0	1.29
Post Intervention II	25	27.1	25.0	6.49

Table 8 - CHC level sample size statistics

Comparing means between datasets, it is clear that Baseline has the largest number of households surveyed per CHC, while PI-I has the lowest. This stands in contrast to the standard deviation results, where Baseline shows the largest spread, while PI-I shows the smallest. Comparing means with medians within datasets, it is clear that most datasets have similar values indicating symmetric distributions. This is not the case for Midline however, where the median is substantially lower than the mean, indicating a skewed distribution. Midline also has a standard deviation that is of the same size as the mean, indicating a large spread. On the whole, most of these characteristics act to balance each other out within each dataset (for example the low mean of PI-I being balanced by the surveying of 51 CHCs). The results suggest that PI-II will yield the most 'statistically sound' results for its observed CHCs, while Midline will possibly yield the least reliable results, followed by PI-I, due to the low number of households surveyed per CHC.

In order to answer questions relating to CHC level performance, boxplots of overall household performance, clustered by CHC are presented in Figure 6 and Figure 7 for each dataset. The CHCs and their associated results are ordered by increasing mean household performance from left to right. These plots help to explore both the robustness of CHC level data for the purpose of intra-CHC comparisons across time, and the variation of household performance within individual CHCs.

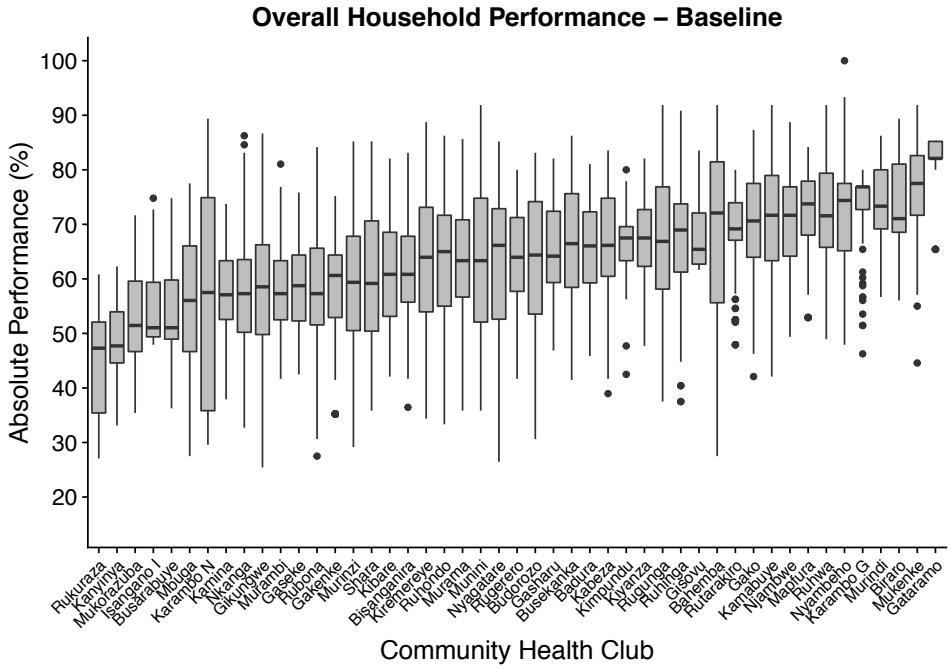

Figure 6 – Box and Whiskers plot showing distribution of overall household performance per CHC in the Baseline dataset. The CHCs have been ordered by mean CHC performance.

Beginning with Figure 6, it is clear that most of the clustered CHC level results in Baseline show the characteristics of reasonably symmetric distributions, free of heavy skewing, as shown by the relatively central locations of the medians within the individual IQRs. Only three of the datasets, found towards the right of the plot, show heavy skewness or a significant number of outliers. As a whole, the results also show a wide range for the IQRs, ranging from approximately 35% to 80%, however the vertical symmetry of the overall distribution about the dataset mean of 63.5% (as many CHCs above this level as below) suggests that the clustered CHC level means are also normally distributed.

Figure 7 shows a similar picture for the other datasets with regards to most containing non-heavily skewed distributions and relatively few outliers per CHC. A few CHCs have extremely narrow distributions however, suggesting low numbers of sampled households for those clusters. Moving sequentially from Midline to PI-II, the gradual upwards shift and narrowing of the IQR 'band' is apparent, along with the development of the upper half of an 's-shaped' profile, with CHCs on the left side seemingly being left behind by those on the right. It is also clear that after Midline, higher performing CHCs to the right of the plots show considerably narrower ranges than lower performing ones to the left.

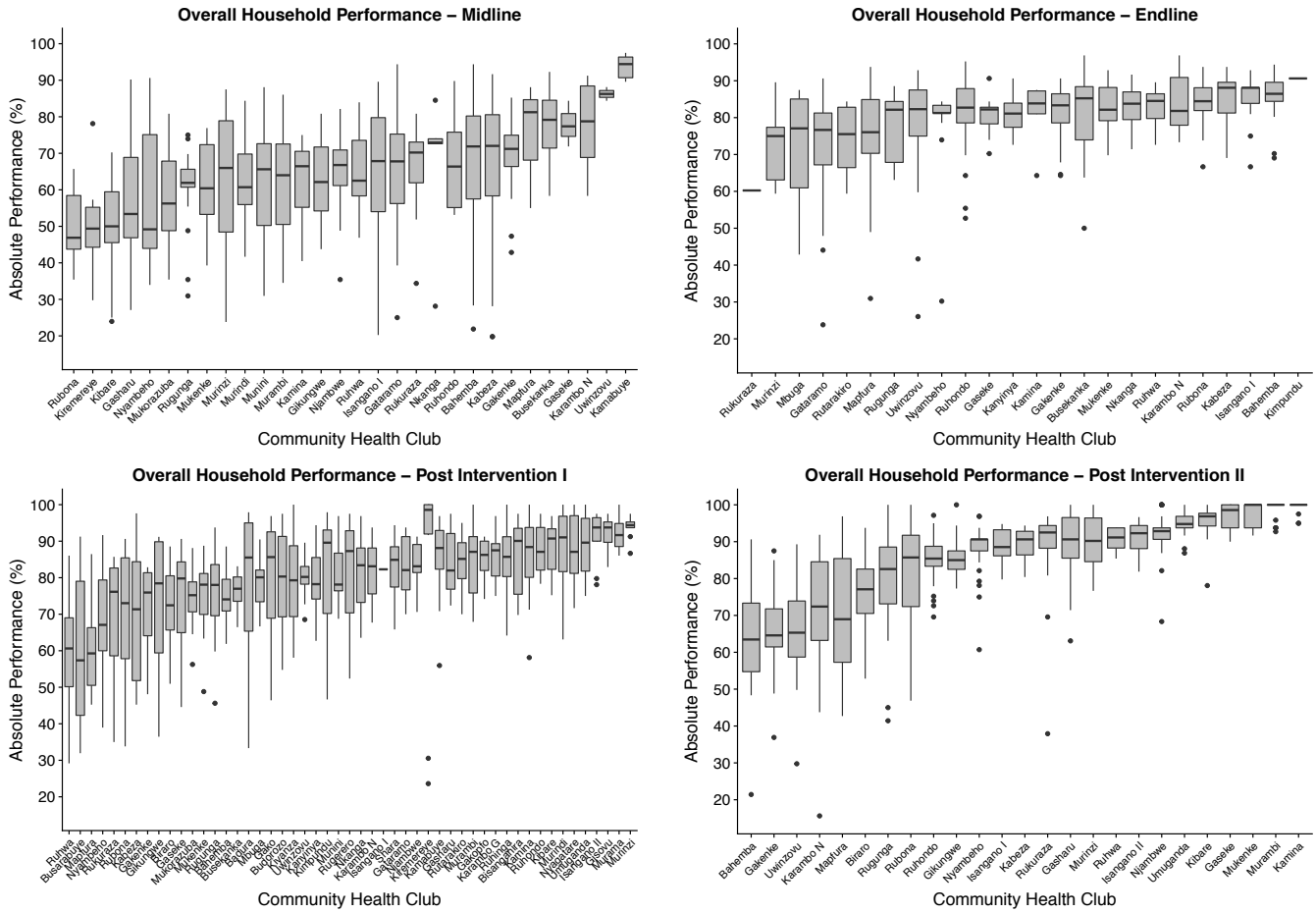

Figure 7 - Box and Whiskers plot showing distribution of overall household performance per CHC in the Midline (top left), Endline (top right), PI- I (bottom left), and PI-II (bottom right) datasets. The CHCs have been ordered by increasing mean CHC performance from left to right.

In order to track and evaluate individual CHC level behaviour change over time, independent samples t-tests were conducted between household observations for CHCs in Baseline and those corresponding to the same CHCs in the other datasets. This yielded results for the differences in means between each pair of observations, the statistical significance of the estimated change ($p \leq 0.05$), and the 95% confidence intervals for the true difference in means between the groups. These results are presented in Figure 8, for the average difference in overall household performance per CHC (relative to their Baseline values). They are shown as dot-plots with error bars corresponding to the 95% confidence intervals, and colouration based on statistical significance. The results have been ordered by increasing positive change relative to Baseline from left to right.

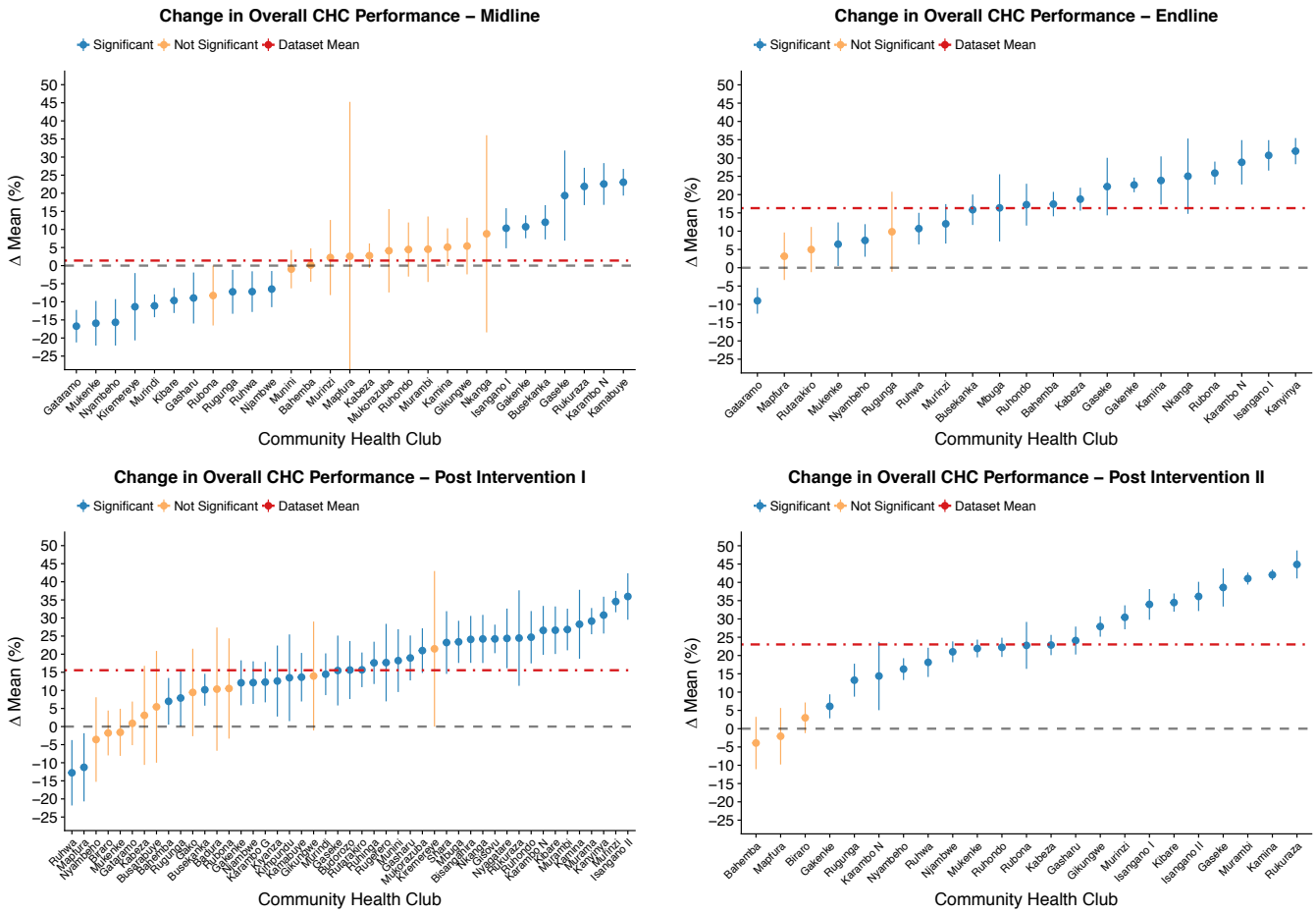

Figure 8 - Dot plots showing 95 % confidence intervals of difference in mean performance per CHC between Baseline and Midline (top left), Endline (top right), PI-I (bottom left) and PI-II (bottom right). Blue points indicate statistically significant results ($p<0.05$), orange points show non-significant results ($p>0.05$). The grey dotted line shows where 0 is on the y-axis and the red dotted line shows the overall mean difference for each dataset. The results are ordered by increasing mean difference from left to right.

The plots confirm that after Midline, positive behaviour change has taken place for most of the observed CHCs over time, relative to their pre-CHC (Baseline) levels. It is also clear that a large range of behaviour change can be seen across CHCs, with Midline showing a 50% difference between top and bottom CHCs (including error bars), followed by 45% for Endline, 65% for PI-I, and 60% for PI-II. Midline contains the most number of negative results, implying that some degree of hygiene behaviour degradation has taken place. Endline and PI-I, contain up to two CHCs that have experienced statistically significant negative change, however all significant CHC results in PI-II show a degree of positive behaviour change.

Regarding the statistically inconclusive results in general, it is clear that Midline and PI-I contain the highest number of inconclusive results, while Endline and PI-II only contain three each. Most inconclusive results lie in the neighbourhood of zero per cent, which is to be

expected. The four highest non-significant results in PI-I seem to violate this, however inspection of the error bars show that they all drop down and meet the zero per cent line. This seems to point to a general rule for all datasets, which is that if error bars significantly overlap the zero per cent line, it is likely that those results will not be statistically significant. Inspection of the size of error bars for significant results shows that PI-II has the most reliable results per CHC, followed by Endline, and that Midline and PI-I contain generally larger and more varying error bars. Finally, looking at the intercept of the dataset mean lines with each of the dot-plots, it seems that most of the results are symmetrically distributed about the mean, and therefore most likely normally distributed.

7.2. Common Trends Within CHCs Grouped by Performance

In order to answer the question of whether common traits exist among high-, mid-, and low-performing CHCs, the CHCs in Figure 8, showing significant change and ranked by overall behaviour change, were split into the respective performance groups based on whether they fell into the upper quartile (top 25%), IQR (middle 50%), or lower quartile (lowest 25%) of the Δ Mean distribution for each dataset. Independent samples t-tests were then conducted between Baseline and the respective dataset for all households from CHCs in the respective groups on a per indicator basis. The resulting per-indicator average change per group is presented in Figure 9 as sets of grouped bars, coloured by group and showing error bars corresponding to the 95% confidence interval for the location of the true difference in means. Non-significant results are coloured red and error bar overlaps should be minimal in order to be confident of statistically significant differences between performance groups.

Beginning with general observations and trends between datasets, it is clear that indicators 7 (cooking) and 8 (childcare) always show consistent positive behaviour change for all three performance groups throughout the four datasets. Apart from the Midline result for cooking, they also show a generally step-wise difference in performance going from low to high groups, with the high group always standing out from the other two when they are similar. The high groups for these indicators tend to consistently stay high or improve slightly, while the other two groups show slower incremental improvement over time in order to ‘catch-up’ with the high group.

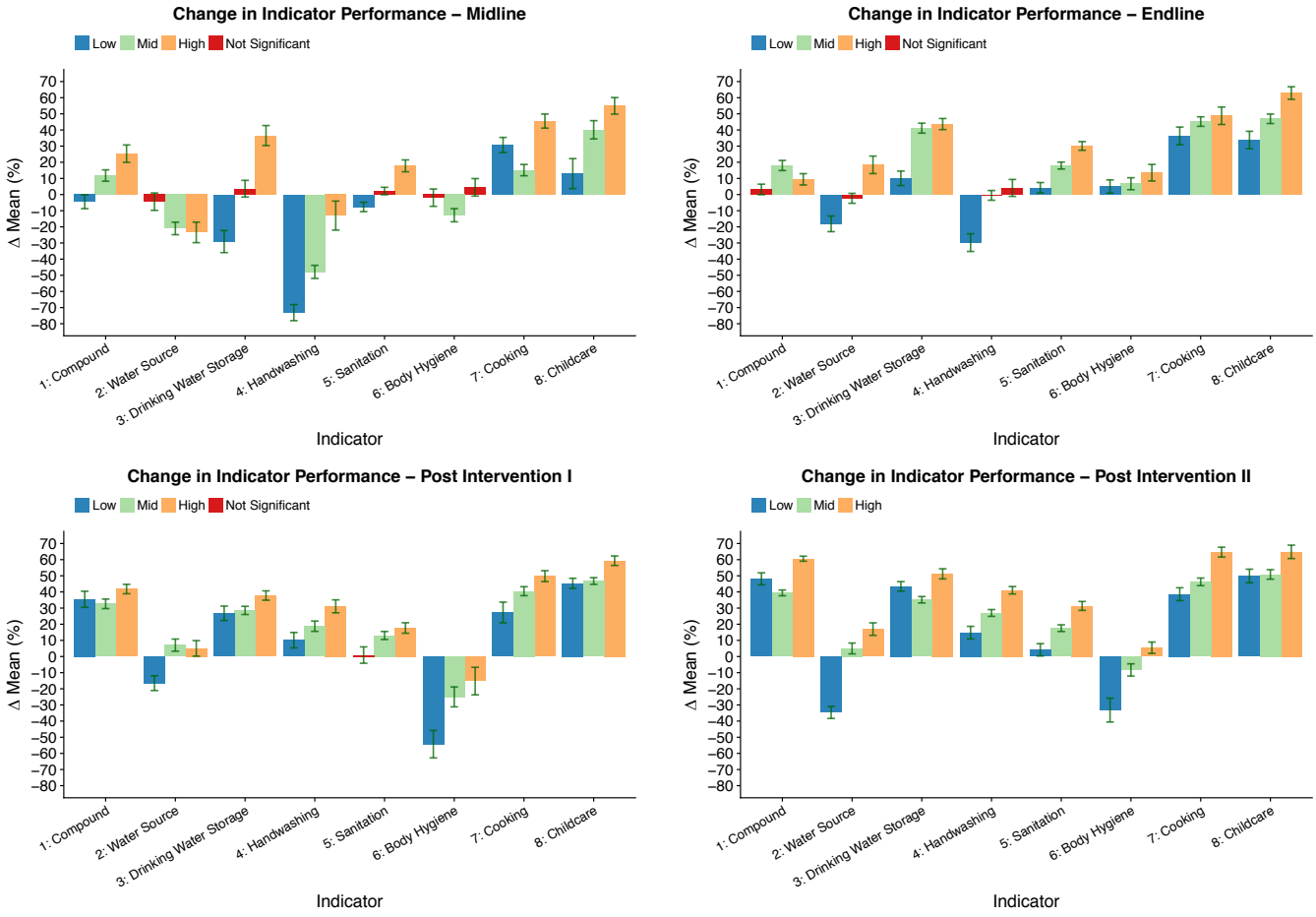

Figure 9 - Bar charts showing 95% confidence intervals of difference in mean performance per indicator per high- (orange), mid- (green) and low- (blue) performing CHCs for Midline (top left), Endline (top right), PI-I (bottom left) and PI-II (bottom right). Red bars indicate non-significant results with a p-value greater than 0.05.

Focusing only on the high-group results, it is also clear that the high-group shows more positive (and relatively less negative) change than statistically significant low-group results in all cases. The high-group, except for the case of indicator 1 (compound) in Endline and indicator 2 (water source) in PI-I, mostly shows either consistent improvement or at least constant positive behaviour change levels over time as one moves through the datasets sequentially. The low group also shows a similar trend and the same exceptions, however it shows worsening negative change for indicator 2 (water source). The low group also seems to struggle with indicator 6 (body hygiene), showing mostly negative change in PI-I and PI-II, and indicator 5 (sanitation), showing little change throughout the datasets. The low-group shows the biggest improvement for indicator 3 (drinking water storage) and 4 (handwashing). The biggest differences between high and low groups can be seen for the sanitation and handwashing indicators throughout the datasets.

In general, all groups seem to struggle in the areas of water source and body hygiene, showing either little or negative change throughout the datasets. A decline in performance relative to their Endline levels can also be seen for all groups for the indicators of sanitation (5), and body hygiene (6) in PI-I.

Supporting Results for Understanding of Behaviour Change Traits

Pearson's correlation tests were carried out between all of the eight household indicator scores for each of the five datasets. These produce correlation coefficient values, known as r values, which indicate the strength of correlation between two sets of independent variables. Typically, absolute values of r between 0 and 0.3 denote weak correlations, those above 0.5 denote strong correlations, and values in between denote moderate correlations. The test also outputs p-values indicating statistical significance between the variables.

These findings are displayed in Figure 10, in the form of heat maps, which have been assigned a colour scale corresponding to three levels derived from dividing the interval from 0 to the strongest calculated correlation coefficient of 0.572, into three equal portions. This is to allow relative visual comparisons of indicator correlations. The values shown in each cell correspond to the calculated r values. Negative values correspond to negative correlations, where an increase in one variable would result in a decrease in the other, while positive values correspond to correlations where both would change in the same direction. The colouration of cells in based on the absolute values of the associated r values. Cells that are left blank represent statistically non-significant results (p>0.05).

	Baseline								Post Intervention II							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
1: Compound																
2: Water Source	-0.052								0.253							
3: Drinking Water Storage	0.193								0.32	0.112						
4: Handwashing	0.236	0.142	0.241						0.37	0.306	0.264					
5: Sanitation	0.2		0.233	0.17					0.486	0.27	0.335	0.402				
6: Body Hygiene	-0.026	-0.057	0.144	0.185	0.255				0.449	0.304	0.266	0.434	0.572			
7: Cooking	0.16	0.04	0.086	0.187	0.129	0.043			0.462	0.259	0.35	0.342	0.486	0.453		
8: Child Care	0.109		0.174	0.117	0.063	0.036	0.04		0.325	0.121	0.303	0.28	0.201	0.229	0.387	

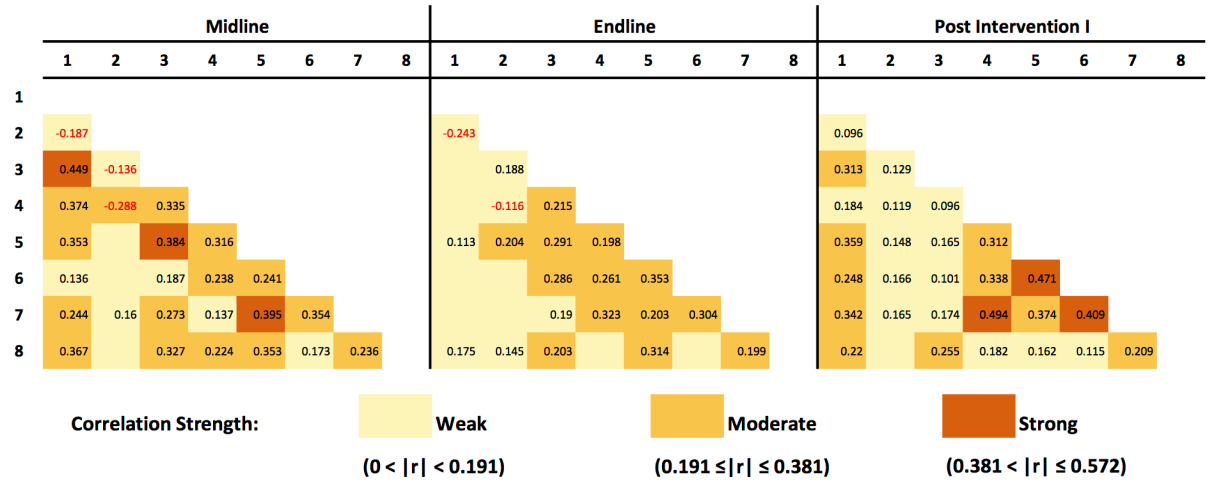

Figure 10 - Pearson's correlations between CHC indicators per dataset

The results show that a clear increase in the number and strength of correlations has taken place between Baseline and PI-II, with a slight slowing in the trend in Endline. PI-I and PI-II show similar patterns, with nearly all of the strongest correlations overlapping. At least one strong correlation in Midline also overlaps with that of PI-II. Those that don't overlap in Midline and Endline typically show relatively high moderate values in PI-II as well.

Focusing on the PI-II results, the strongest correlation can be found between indicators 5 (sanitation) and 6 (body hygiene), with an r-value of 0.572. Going from left to right, it is clear that indicator 1 (compound) is well correlated with indicators 5,6, and 7 (cooking). Next, indicator 4 (handwashing) also shows strong correlations with indicators 5 and 6, while indicator 5 is also strongly correlated with indicator 7 (in addition to indicator 6, mentioned above). Finally, indicator 6 shows strong correlation with 7, and indicator 7 shows strong correlation with indicator 8 (cooking). The moderately correlated pairs that were strongly correlated in PI-I and Midline are indicator 3 (drinking water storage) with indicator 5, and indicator 4 with indicator 7 as well indicator 1 (compound) with indicator 3 (drinking water storage). These results indicate a fairly consistent pattern, showing that five particular indicators repeatedly correlate with one another or with other individual indicators. These consist of indicators 1 (compound), 4 (handwashing), 5 (sanitation), 6 (body hygiene), and 7 (cooking).

The next set of results aim to shed light on whether differences can be seen between the average levels of attendance per performance group for each dataset. Summary statistics for the mean number of CHC sessions attended and the standard deviations about their means

are presented in Table 9. It is important to note that the CHCs within each group are different between each of the datasets.

	Attendance of CHC sessions					
	High Group		Middle Group		Low Group	
Dataset	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Midline	1.40	0.49	1.15	0.431	2.21	1.98
Endline	19.6	1.52	19.3	2.39	18.7	3.48
Post Intervention I	18.8	3.70	16.1	8.33	14.6	8.23
Post Intervention II	20.7	3.96	19.2	4.58	16.8	7.36

Table 9 - Summary statistics on attendance of CHC sessions by different relative CHC performance groups.

Focusing on the mean values for each of the groups, the results show that surprisingly few sessions have been attended on average by all groups by Midline. By Endline, the high performing group has attended approximately one more session on average than the low performing one. This difference is more pronounced for PI-I and PI-II, where the high performing groups have attended approximately four more sessions than the low performing ones. Focusing now on the standard deviation values, it is clear that high performing groups seem to more consistently attend CHC sessions, as illustrated by the fact that their standard deviations are mostly less than half of those for the low performing groups.

The final set of results relate to a qualitative analysis that was conducted in order to track whether particular CHCs consistently appeared in the same performance groups throughout the Midline to PI-II datasets. Only CHCs that appeared at least three times throughout the datasets were included in the analysis. For the purposes of clarity, the results presented in Table 10 only correspond to CHCs that appeared in all four datasets. The other results can be found in the Annex (Table 11). The table is vertically split by performance group with the high group at the top and low at the bottom. A ranking percentage is given for each CHC, indicating its position within each overall dataset ranking, with 100% indicating the top of the ranking. The group ranking column indicates the position within each of the group rankings.

Midline			Endline			Post Intervention I			Post Intervention II		
CHC	Ranking (%)	Group Ranking	CHC	Ranking (%)	Group Ranking	CHC	Ranking (%)	Group Ranking	CHC	Ranking (%)	Group Ranking
Karambo N	94%	High	Karambo N	89%	Middle						
Gaseke	82%	Low				Karambo N	81%	Low	Gaseke	85%	Low
Gakenke	71%	High	Gakenke	67%	High						
Ruhwa	53%	Middle	Gaseke	61%	High	Gaseke	36%	Low			
			Ruhwa	22%	High				Ruhwa	25%	High
						Gakenke	17%	Middle	Karambo N	15%	Middle
						Ruhwa	3%	Low	Gakenke	5%	Low

Table 10 - CHC performance tracking chart with four data points per CHC.

The overall results from the analysis were generally inconclusive, particularly those relating to CHCs that appeared only three times across the datasets. The results above however do show a general declining trend in the ranking of CHCs that started off in the high and mid performing groups in Midline, with the exception of one CHC (Gaseke) that oscillated between the mid and high groups. The number of CHCs is unfortunately too low however in order to draw any general conclusions.

8. Discussion

The discussion of the key findings of this study is guided by the primary research questions and objectives. Where possible, the discussion is supported by external sources of information.

8.1. Research Question I

Did significant hygiene behaviour change take place within Rusizi district CHCs? If so, to what extent, and with how much consistency did these changes take place across observed CHCs?

The findings of this study indicate that statistically significant positive behaviour change took place for the majority of sampled CHC households over time, both during and after the implementation of CBEHPP training within the 50 Classic Rusizi district CHCs. This was also true at the CHC level, however some CHCs showed negative change within the Midline, Endline, and PI-I datasets. A wide range of behaviour change levels was also found across all of the CHCs in the post-Baseline datasets, indicating that a lack of consistent change might have taken place across CHCs.

Midline

Beginning with the Midline dataset level results, it is clear that overall change relative to baseline was only moderately significant, as shown by the t-test results. The boxplot and kernel density results showed that while the Midline results had a slightly higher median value and rightwards shifted distribution peak respectively, the distribution was also relatively broader than that of Baseline. Without interpreting the CHC level results yet, this implies that some households in Midline had begun to already show change, while others had been left behind, thereby stretching the distribution in the positive direction.

This can be explained by the fact that not all CHCs began their training programmes at exactly the same time and that those that started early would only have been part way through the training syllabus. This explains the apparent 'inertia' that the behaviour change progression seems to have to overcome before the large jump in performance seen by Endline.

The image given by the dataset level statistics however changes when the corresponding CHC level results are inspected. These results show that many CHCs had in fact experienced negative overall behaviour change relative to their Baseline levels and had not just been left behind as previously thought.

There are two possible explanations for this result, both relating to the period of data collection that took place between April and May 2014. The first explanation is that CHC attendance and the completion of homework assignments took a low priority for most CHC members due to the fact that April is one of the wettest months of the year in Rusizi district (World Climate Guide, 2012). This would have put additional pressure on agricultural families to spend more time in the fields rather than attending CHC sessions or making home improvements. The second explanation, is that April also constitutes the national annual memorial month for the Rwandan Genocide, resulting in the slowdown of many aspects of everyday life including the availability of goods and services. This hypothesis is supported by the consistently low attendance levels shown in Table 9 for Midline. Both of these factors would have compounded the progress of behaviour change efforts surveyed in Midline.

The low levels of CHC session attendance would have also affected the level of 'common unity' (Waterkeyn and Waterkeyn, 2013) being achieved by CHCs and could explain why some households were achieving positive behaviour change while others were left behind.

Endline

The Endline dataset results show the biggest jump in behaviour change performance out of all of the datasets. The narrow overall performance distributions shown by the corresponding dataset level boxplot and kernel density plot suggest that a uniform shift in improved performance had taken place for the majority of sampled CHC members. The narrow, tall peak of the density plot also implies a much higher level of consistency for the achieved level of overall performance. The corresponding CHC level results also show the lowest level of behaviour change variation (~45% between top and bottom CHCs) among all of the datasets.

These results suggest that by the end of the training period, some sort of major widespread change had taken place throughout the surveyed CHC households. The relatively 'tight' distribution of overall performance and hygiene behaviour change also suggests that 'common unity' had been achieved by CHCs, resulting in most members progressing together through the programme and emerging at the end with similar levels of community wide detectable changes in behaviour. The fact that the training period between the Midline and Endline datasets had taken place during the dry season, could have also contributed to high levels of behaviour change. This is supported by the high attendance levels across all performance groups shown in Table 9. The benefits of dry season CHC training are supported by (Waterkeyn, 2006).

Post Intervention I

Dataset level results for PI-I indicate that no statistically significant change took place when compared to Endline, as shown by the t-test results. This means that behaviour change levels were maintained across observed CHCs for more than one year after the end of the training programme. The rightwards shifted peak and broader distribution of the associated kernel density plot and boxplot respectively indicate both a wider range of behaviour change relative to Endline, and continued progress by high achievers. This broadening of the distribution could be attributed to the natural loss of excitement or interest that might have taken place after the end of the training programme for CHCs with lower levels of common unity. The maintenance of the average dataset level of overall performance however indicates that CHCs are capable of long term sustainable behaviour change.

Due to the low sample sizes, the CHC level results concerning CHCs that might have experienced negative change are difficult to interpret due to the lack of statistical

significance of many of these results. Low sample sizes could also explain the distributional trends mentioned above.

One possible influencing factor on the overall performance of the dataset could have been the IPA-Endline survey which took place shortly the AA PI-I survey. This could have resulted in a temporary Hawthorne effect, which could have artificially improved the performance of CHC households. Another factor that might have resulted in the broadening of the distribution could have been the effects of the rainy season, and the memorial month for the Rwandan Genocide, due to the fact that the PI-I data collection took place between April and May 2016.

Post Intervention II

The dataset level results for PI-II show the highest levels of overall performance out of all of the datasets. The highly skewed nature of its performance distribution suggests that high performers had continued to push themselves to adopt more positive hygiene practices over time. The raised position of the lower whisker and bottom of the PI-II IQR seen in the associated boxplot seems to indicate that the effects of common unity had continued to filter down to lower performers.

The CHC level results also show that no CHCs showed statistically significant negative behaviour change, and that all CHCs had shown improvement relative to their Baseline states. A wide range still existed however between the top and bottom CHCs.

One possible explanation for the high overall performance scores seen in PI-II was that district level competitions had been held between CHCs prior to PI-II data collection. This could have motivated CHC members to put exceptional effort into showing that they were practising the recommended behaviour practices.

Baseline

Dataset level overall performance results for the Baseline dataset indicate that the majority of CHC households were already practising over half of the recommended practices prior to CBEHPP intervention. With the training sessions starting shortly after the Baseline data collection round that took place between October to November 2013, the Rusizi intervention would have started at the height of the rainy season. Both of these factors – a high baseline level of positive hygiene practices and starting a CHC programme in the middle of the rainy season, have been known to affect the ultimate consistency of achieved

behaviour change. The effects of the rainy season are mentioned by Waterkeyn et al. (2006), while those of starting with a high baseline are mentioned by Waterkeyn et al. (2005). Important factors mentioned by the latter include the importance of identifying previous interventions planned CHC areas, and the fact that the most impressive and consistent CHC level results are achieved by communities who have received little subsidised health related support in the past.

Robustness and Limits of the Data for Drawing General Conclusions

Steps were taken throughout this study to establish the level of uncertainty associated with all of the generated results. By computing 95% confidence intervals, as well as the population standard deviation and standard error of overall performance per dataset, it was possible to estimate how well the sampled household data would reflect all CHC households. These tests showed that at the dataset level, a low level of uncertainty could be achieved using the available data, and that general conclusions could be drawn concerning the likely performance of all Rusizi district CHCs. This also applied to the evaluation of dataset level hygiene behaviour change estimates based on t-test results.

At the CHC level, 95% confidence intervals were computed as part of the t-tests used to evaluate hygiene behaviour change. These allowed for the production of error bars which could be visually evaluated for the purpose of understanding the likely range of values where the true difference in mean performance could be found. In this way, uncertainty was also tracked at the CHC level. Out of the post-Baseline datasets, the PI-II dataset has lowest level of uncertainty owing to its balanced sample size statistics and seen by its relatively small error bars. The PI-I dataset has relatively high uncertainty for individual CHC behaviour change values due to small sample sizes, but benefits from the fact that nearly all CHCs were sampled. The Midline dataset is the most problematic, with large confidence intervals and many statistically non-significant results.

Similar measures were taken for the indicator level bar charts. These benefited from large sample sizes however due to the relatively large performance groups and corresponding clusters. This resulted in relatively low levels of uncertainty for results that showed reasonably non-zero behaviour change.

8.2. Research Question II

Do common traits exist among high-performing and low-performing CHCs with respect to hygiene behaviour change? If so, can these be linked to elements of the CBEHPP training and M&E programme implementations?

The findings of this study show that when CHCs are categorised into high-, mid- and low-performing groups, common trends can be identified among them. The most striking result in this regard was the positive change found in all groups for the child care and cooking indicators, where a steady increase in positive behaviour change was observed over time. This result was further validated through the Pearson's correlation test for PI-II, which detected a strong positive correlation between the same indicators. This implies that the two indicators are possibly intrinsically linked. Even though no gender analysis was carried out in this study, this result could be an indication of the positive effect the Clubs have on women. Women in general are the primary family caregivers in Rwanda and their responsibilities include child care as well as cooking. Hence, this could also suggest that these indicators were a priority for improvement by CHC members, as the majority of the Clubs tend to attract high proportions of women (Waterkeyn and Cairncross, 2005).

The indicators that all groups struggled with, were body hygiene, water source and handwashing. In particular, water source seemed to show a consistent negative behaviour change trend for low performing CHCs over time. The nature of the associated sub-indicators for the main water source indicator, found in Table 5, could possibly explain the challenges faced by the low performing groups in achieving improvement. It can be argued that the sub-indicators are both highly location dependent, as well as being dependent on the socio-economic status of a household. Examples include indicator 2b – 'Is the walking distance to the water source 30 minutes or less' (location dependent), and indicator 2d – 'Are there 15 litres or more of water available per household member each day' (socio-economic status dependent). This possibly suggests that high performing groups tend to be composed of households that are located close to improved water sources or those that have relatively higher socio-economic statuses.

The other two indicators that all groups struggled with, body hygiene and handwashing, both required the construction and maintenance of a bath shelter and a handwashing facility. One of the conclusions of the Whaley and Webster (2011) paper was that CHC

members, particularly those from low socio-economic backgrounds, are less likely to repair broken handwashing facilities. This statement could also be extended to bath shelters. Sanitation also corresponded to an indicator where large disparities existed between low and high performing households in terms of their ability to practice positive behaviours. Considering the overall performance of all groups over time, this indicator did not show large amounts of change between datasets, indicating some sort of barrier that was impacting all of the groups' abilities to achieving higher levels of change. One possible barrier is a household's ability to procure the relevant materials and resources to construct handwashing facilities, bathing shelters, and latrines. This was also proposed by Whaley et al. (2011), who suggested that the relative wealth of a household was found to significantly affect the likelihood of latrine construction over other aspects of sanitation focused behaviour change programmes.

Interestingly, the Pearson's correlation results for PI-I, showed a strong relationship between the indicators of body hygiene and sanitation, suggesting that if socio economic barriers are overcome, and sufficient resources can be found for constructing latrines, then it is most likely that resources would also be available for constructing bathing shelters. Referring to the bar chart in Figure 9, which shows a greater positive change for sanitation than for body hygiene, it can be assumed that latrine construction is most likely to happen prior to bath shelter construction.

The compound indicator also showed a gradual positive behaviour change trend for all performance groups over time. This is possibly because it is an easy indicator to achieve, requiring little to no resources in order to show improvement. Additionally, when looking at the Pearson's correlation results for PI-II, the compound indicator showed strong correlations with the indicators of cooking, body hygiene and sanitation. This can be explained by the traditional layout of rural homes in Rwanda, where most of the cooking is done outside, at the back of their compound, and where bath shelters and latrines are typically separate from the main house. This result suggests that if compounds are well looked after, it is likely that positive behaviours are also being practised in the indicators of cooking, body hygiene, and sanitation. This might mean that the compound indicator could be used as a 'super-proxy' for predicting other behaviours and could thus be used for quick inspections of CHC households.

A trait that stands out between high and low groups is that in general, high groups show more consistent (as well as higher levels of) positive behaviour change than low groups across all indicators and datasets. This implies that high groups achieve early successes during the training programme and maintain these successes over time. Low groups often seem to lag behind, however often show substantial improvement given enough time.

Another observation between the performance groups is that in general, households in the low performing groups attended fewer CHC sessions than those high performing groups, with the exception of midline. High performing groups also attended sessions with much more consistency than low performing ones. This possibly explains the higher levels of consistent behaviour change shown by the high groups. This is supported by the study carried out by Waterkeyn at al. (2013) in Zimbabwe, who stated that 'complete health knowledge' had a more positive effect on the uptake of positive hygiene practices than 'incomplete knowledge'. This relates to the fact that less consistent CHC session attendance means that a given CHC member has less chance to absorb the required level of knowledge in a participatory way (also missing out on the benefits of common unity), so as to translate it into well embedded hygiene behaviours.

An additional observation of the average attendance levels over time, show that the number of sessions attended by high group households generally increases steadily between datasets. This not the case for the low group, where the highest level of attendance takes place in Endline, followed by a sharp decline in PI-I, then a slight increase by PI-II. This suggests that high group CHC members continue to practise common unity and attend CHC sessions well after the end of the training period, whereas low group members show a reduction in their commitment to attend sessions.

8.3. Influencing Factors and Limitations of the Study

The largest limitation of this study was the fact that out of a maximum of 55 possible sub-indicators, only 29 could be used, thereby presenting an uneven picture of hygiene behaviour change in Rusizi district CHCs.

Furthermore, the lack consistency between main indicator categories with respect to the number of sub-indicators assigned to them, could have introduced errors through the quantitative methods used for this study. Steps were taken to make the most of this uneven data structure however and to minimise errors.

The subjective nature of the dataset and indicator alignment process might have been another source of error that could have made comparisons between datasets from different survey versions unreliable. Comparisons within survey versions should not have been affected however, thereby not invalidating the most significant behaviour change improvement seen between the Midline and Endline datasets.

The last factor that could have affected the reliability of results was the large variations between datasets in the sample sizes per CHC. These were taken into account through the various uncertainty calculation efforts that were carried out during the study.

9. Conclusions

Main Conclusions

The findings of this study conclude that statistically significant hygiene behaviour change took place for the majority of sampled CHC households from the 50 Classic villages of the Rusizi district CBEHPP intervention. These changes took place over a course of approximately three years and generally increased over time. All were statistically significant when compared to Baseline levels. Moderate positive change was found between the Baseline and Midline datasets. The greatest positive change then took place from Midline to Endline. This was followed by no statistically significant change between Endline and the Post Intervention I dataset. This meant that long term positive behaviour change had been sustained for more than one year after the end of the CBEHPP training programme. A final positive increase was detected nearly one year afterwards in the PI-II dataset.

Low levels of uncertainty were found for the dataset level results based on sample statistics. It was found that predicted levels of behaviour change for the dataset level samples would most likely reflect those of the population (all CHC households) fairly well, due to low levels of uncertainty.

Signs of the effects of common unity were found in the data by evaluating the dataset level distributions of overall performance. The Endline and PI-II datasets showed relatively narrow distributions indicating consistent performance for the majority of sampled households. The opposite was found for the Midline and PI-I datasets, where distributions had widened compared to previous datasets indicating increased disunity within CHCs. Datasets with wider distributions and less impressive performance among households were identified as

being collected during both the annual rainy season in Rusizi district and the national memorial month for the Rwandan Genocide. Analysis of attendance levels confirmed that attendance dropped for these datasets. This helped to confirm the relationship between household performance distributions, attendance levels, common unity, and overall dataset performance.

CHC level studies confirmed that a wide range of behaviour change took place within each dataset. This implied that there might have been issues concerning the consistency of positive behaviour change across CHCs. Statistically significant negative changes were detected in all post-Baseline datasets, except for PI-II. The level of statistical significance was estimated for each CHC and was found to be heavily dependent on the quality and consistency of CHC level sample sizes.

The CHC level results were partitioned into high, mid, and low performing groups and indicator level behaviour change was evaluated per group per dataset. Common trends were identified among high and low performing groups. All groups showed consistent positive change for the indicators for compound, cooking, and childcare over time. Low performing groups consistently struggled with the indicators of body hygiene, water source, sanitation, and hand washing. High performing groups performed more consistently over time and achieved higher scores than the low groups in every dataset. High groups were also found to attend CHC sessions across datasets more consistently than low groups.

Correlation tests were conducted between all indicators for each dataset and showed that the indicators for sanitation and body hygiene were strongly correlated in two datasets. Strong correlations were also found between the indicators for compound, sanitation, body hygiene, and cooking within the PI-II dataset. Similar trends were seen in some of the other datasets.

Sample size analyses were conducted and it was found that the PI-II dataset corresponded to the best quality data, while the Midline dataset corresponded to the lowest quality data. An optimum sample size of 30 households per CHC is recommended in order to maximise accuracy.

Recommendations for Future Research

Based on the possible impact of geographic location and socio-economic household status on indicators such as water source and sanitation, found in this study, it is recommended

that the effect of these factors should be investigated further. This could be achieved by including GPS data in future studies, along with appropriate socio-economic indicators.

Investigation at the sub-indicator level using methods such as frequency counts could also shed more light on the main indicator results found in this study.

Recommendations for Improvements to the CBEHPP Training Programme

Based on the findings of this study, it is recommended that CBEHPP training programmes should start in the dry season, and that efforts are made to establish common unity as soon as possible within CHCs.

Based on the performance group analysis, the indicators of water source, body hygiene, hand washing, and sanitation should be paid special attention due to the fact that all groups struggled in these areas. Low performing groups should be identified as soon as possible in order to provide additional support. This could be identified by conducting intermediate surveys and looking out for the 'problem indicators' that low group members struggle with in particular, namely water source and body hygiene.

Recommendations for Improvements to the M&E Programme

Based on the correlation findings above, it is recommended to consider using the compound indicator as a 'super-proxy' indicator for the purpose of getting quick estimates of the levels of the other indicators with which it was strongly correlated. This could be used to quickly identify low-performing households early on in an intervention.

Another recommendation for the M&E programme would be to track CHCs over time, by ensuring that the same CHCs appear in surveys during at least three separate time periods.

Ideal sample sizes per CHC should be calculated based on recorded CHC member numbers in order to ensure consistent confidence intervals.

A similar numerical scoring system and analytical methodology, as used in this study, should be followed in order to quantify behaviour change levels.

10. Bibliography

Africa AHEAD (2015a) *Annual Report*. Kigali. Available at: https://www.africaahead.org/2015/08/rwanda-africa-ahead-2015-annual-report/.

Africa AHEAD (2015b) *Annual Report*. Cape Town. Available at: https://www.africaahead.org/about-us/annual-reports/.

Ajzen, I. and Fishbein, M. (1980) *Understanding Attitudes and Predicting Social Behaviour*. New Jersey: Prentice Hall.

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. New Jersey: Prentice Hall.

Bandura, A. (2017) *Social Learning Theory Bandura*, *Learning-Theories.com*. Available at: https://www.learning-theories.com/social-learning-theory-bandura.html (Accessed: 26 February 2018).

Cairncross, S., Waterkeyn, J. A., Waterkeyn, A. J., Uwingabire, F., Ntakarutimana, A., Pantoglou, J., Katabarwa, J. T., Bigirimana, Z., Cumming, O. and Carter, R. C. (no date) 'Journal of Water , Sanitation and Hygiene for Development Measuring hygiene behaviour change in Community Health Clubs : methodological questions arising from a cluster-randomised controlled trial in Rwanda .', *Journal of Water Sanitation and Hygiene for Development*.

Condo, J., Mugeni, C., Naughton, B., Hall, K., Tuazon, M. A., Omwega, A., Nwaigwe, F., Drobac, P., Hyder, Z., Ngabo, F. and Binagwaho, A. (2014) 'Rwanda's evolving community health worker system: a qualitative assessment of client and provider perspectives', *Human Resources for Health*, 12(71), pp. 1–7. doi: 10.1186/1478-4491-12-71.

Curtis, V. C. S. Y. R. (2000) 'Domestic hygiene and diarrhoea', *Tropical Medicine and International Health; 2000*, 5(1), pp. 22–32. doi: 10.1046/j.1365-3156.2000.00512.x.

Janz, N. K. and Becker, M. H. (1984) 'The Health Belief Model: A Decade Later', *Health Education Quarterly*, 11(1), pp. 1–47. doi: 10.1177/109019818401100101.

Kawachi, I. (1999) 'Social capital and community effects on population and individual health.', *Annals of the New York Academy of Sciences*, 896, pp. 120–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10681893 (Accessed: 26 February 2018).

Kawachi, I. and Berkman, L. F. (2003) *Neighbourhoods and Health*. New York: Oxford University Press Inc, USA.

Lumley, T., Diehr, P., Emerson, S. and Chen, L. (2002) 'The Importance of the Normality Assumption in Large Public Health Data Sets', *Annual Reviews of Public Health*, 23, pp. 151–169. doi: 10.1146/annurev.publheath.23.100901.140546.

Ministry of Health (2004) *Health Sector Strategic Plan 2005 - 2009*. Kigali: Government of Rwanda.

Ministry of Health (2009) *Health Sector Strategic Plan July 2009 - June 2012*. Kigali: Government of Rwanda.

Ministry of Health (2010) *Roadmap for CBEHPP*. Kigali: Environmental Health Desk,

Government of Rwanda.

Ministry of Health (2015) *Rwanda: Community Health Insurance Scheme Credited for Reducing Child Mortality (All Africa)*. Available at: http://www.moh.gov.rw/index.php?id=34&tx_ttnews%5Btt_news%5D=400&cHash=cd6e6e8fab558bfc946ec4fd2ae56b6e (Accessed: 26 February 2018).

Neighbors, C., Foster, D. W. and Fossos, N. (2013) 'Peer Influences on Addiction', in *Principles of Addiction*. Elsevier, pp. 323–331. doi: 10.1016/B978-0-12-398336-7.00033-4.

Peal, A., Evans, B. and Van der Voorden, C. (2010) *Hygiene and Sanitation Software: An Overview of Approaches*, *Water Supply & Sanitation Collaborative Council*. Geneva: Rosseels S. A. - Switzerland. Available at: www.wsscc.org.

Republic of Rwanda Ministry of Finance and Economic Planning (2000) *Rwanda Vision 2020*. Kigali.

Sinharoy, S. S., Schmidt, W. P., Wendt, R., Mfura, L., Crossett, E., Grépin, K. A., Jack, W., Rwabufigiri, B. N., Habyarimana, J. and Clasen, T. (2017) 'Effect of community health clubs on child diarrhoea in western Rwanda: cluster-randomised controlled trial', *The Lancet Global Health*. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license, 5(7), pp. e699–e709. doi: 10.1016/S2214-109X(17)30217-6.

United Nations (2003) *Indicators for Monitoring the Millenium Development Goals*, *ST/ESA/STAT/SER.F/95*. New York: United Nations Publications.

University of Twente (2017) *Health Communication | Health Belief Model*. Available at: https://www.utwente.nl/en/bms/communication-theories/sorted-by-cluster/Health Communication/Health_Belief_Model/ (Accessed: 26 February 2018).

Waterkeyn, J. (2006) 'Community Health Clubs - District Health Promotion using the Consensus Approach', (July), pp. 1–25. Available at: http://www.africaahead.org/wp-content/uploads/2011/01/2006_WELL1.pdf1.pdf.

Waterkeyn, J. (2011a) *Community Based Environmental Health Promotion Program (CBEHPP): Manual for CHC Facilitators and Community Health Workers*. WB-WSP/UNI. Kigali: Environmental Health Desk, Ministry of Health Rwanda. Available at: https://www.africaahead.org/documentation/training-tools/.

Waterkeyn, J. (2011b) *Start-up of Community Hygiene Clubs in Rwanda: Draft Manual for Trainers of Environmental Health Officers*. Kigali: Environmental Health Desk, Government of Rwanda. Available at: https://www.africaahead.org/documentation/training-tools/.

Waterkeyn, J. A. and Waterkeyn, A. J. (2013) 'Creating a culture of health: hygiene behaviour change in community health clubs through knowledge and positive peer pressure', *Journal of Water Sanitation and Hygiene for Development*, 3(2), pp. 144–155. doi: 10.2166/washdev.2013.109.

Waterkeyn, J. and Cairncross, S. (2005) 'Creating demand for sanitation and hygiene through Community Health Clubs: A cost-effective intervention in two districts in Zimbabwe', *Social Science & Medicine*, 61(9), pp. 1958–1970. doi: 10.1016/j.socscimed.2005.04.012.

Waterkeyn, J. and Waterkeyn, A. (2005) 'Taking Phast the Extra Mile Through Community

Health Clubs : The A.H.E.A.D Methodology (Applied Health Education and Development)', *Water and Sanitation Programme*. Available at: https://africaahead.org/wp-content/uploads/2008/01/taking-phast-the-extra-mile.pdf.

Whaley, L. and Webster, J. (2011) 'The effectiveness and sustainability of two demand-driven sanitation and hygiene approaches in Zimbabwe', *Journal of Water, Sanitation and Hygiene for Development*, 1(1), p. 20. doi: 10.2166/washdev.2011.015.

Wood, S., Sawyer, R. and Simpson-Herbert, M. (1998) *PHAST step-by-step guide: a participatory approach for the control of diarrhoeal disease*. WHO/EOS/98. Geneva: World Health Organization.

World Climate Guide (2012) *Climate in Rwanda: temperature, precipitation, when to go, what to pack*. Available at: https://www.climatestotravel.com/climate/rwanda (Accessed: 1 March 2018).

11. Annexes

Midline			Endline			Post Intervention I			Post Intervention II		
CHC	Ranking (%)	Group Ranking	CHC	Ranking (%)	Group Ranking	CHC	Ranking (%)	Group Ranking	CHC	Ranking (%)	Group Ranking
			Isangano I	94%	High	Murinzi	97%	High	Rukuraza	100%	High
Rukuraza	88%	Middle				Kibare	83%	Low			
Busekanka	76%	Low	Rubona	83%	Low	Ruhondo	78%	Low			
						Rukaraza	75%	High	Kibare	75%	High
Isangano I	65%	High							Isangano I	70%	High
Njambwe	59%	High							Murinzi	65%	High
						Gasharu	53%	Middle	Gasharu	55%	Middle
Rugugna	47%	Middle	Ruhondo	44%	Middle				Ruhondo	40%	Low
Gasharu	41%	Low							Mukenke	35%	Low
Kibare	35%	Low	Busekanka	33%	Low				Njambwe	30%	Low
			Murinzi	28%	High	Njambwe	19%	High	Nyambeho	20%	High
Nyambeho	18%	Middle	Nyambeho	17%	Middle	Busekanka	14%	Middle			
Mukenke	12%	Low	Mukenke	11%	Low	Rugunga	11%	Middle	Ruguna	10%	Low

Table 11 – CHC performance tracking chart with three data points per CHC.

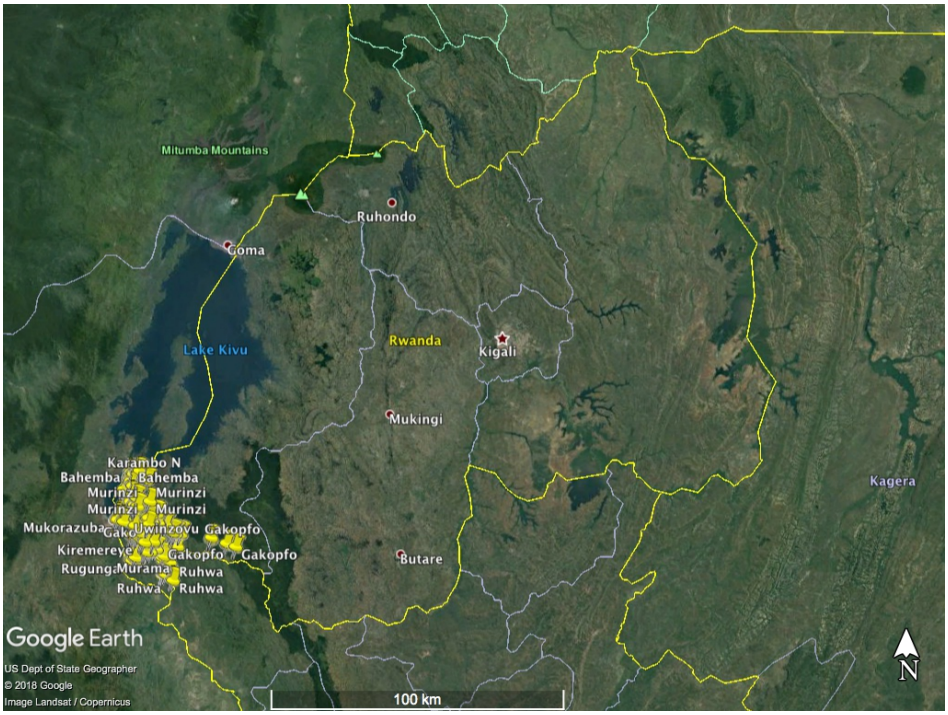

Figure 11 – Map of Rwanda showing the locations of CHC households in Rusizi. GPS data was recorded from surveyed households.

Figure 12 - Map of Rusizi district showing the locations of CHC households.

12. Declaration of Originality of Work

Declaration

This thesis is the result of independent investigation. Where my work is based on the work of others, I have made appropriate acknowledgements.

I declare that this study has not already been accepted for any other degree nor is it currently being submitted in candidature for any other degree.

Date and candidate's signature

13. Acknowledgements

I would like to express by deepest gratitude to a number of people without whom this thesis would not have been possible to do. First, I would like to thank Dr Juliet Waterkeyn and Anthony Waterkeyn for taking a chance on me and allowing me to use their work as part of this thesis. I would like to particularly thank Dr Juliet Waterkeyn for her support and encouragement of my work as well as her useful insights.

My greatest appreciation goes to Dr Richard Carter, my supervisor, for his patience and guidance throughout the development and writing of this thesis. As well as for his encouraging support during the challenging periods.

I would also like to thank the Africa AHEAD team in Rwanda, who were a crucial part of the Rusizi project and thus the development of this thesis. In particular Joseph Katabarwa, Etienne Havumiragira, Mercy Mbirira, Andrew Ndahiro, Jeanne Gasengayire, Fausca Uwingabire, Zachary Bigirimana and Amans Ntakarutimana for his assistance in obtaining ethical approval. As well as Tristan Waterkeyn, for his hard work on the monitoring and evaluation tools.

I would very much like to thank Dr Mathias Borchert, the former course coordinator for his valuable guidance throughout the Masters course and for the development of the thesis.

Additionally, I would like to thank my family for their loving support and encouragement as well as Sahir Khan for dedicating countless hours to proof reading my work.

Lastly, I would like to thank the Institute of Tropical Medicine and International Health at Charite Medical University, particularly the Committee and Admissions Department (CAD) and administrative department for their patience and support.

14. Curriculum Vitae

Name:	Julia Pantoglou	**Date of Birth:**	12th March 1987
Email:	jpantoglou@gmail.com	**Telephone No.:**	+49 1577 346 1996

Education

Sep 2013 – Present **Masters in International Health, Charité – Medical University Berlin, Germany**
Master of Science (MSc) - Completion pending thesis
Core Modules: Epidemiology, Qualitative and Quantitative Analysis, Public Health Planning, Health Economics, Quality Management, Tropical Medicine & Infectious Diseases, Medical Anthropology

Elective Modules:

Jan 2014 – May 2014 **University College London, United Kingdom**
- Advanced Statistical and Epidemiological Methods: Quantitative Analysis
- Health Management: Project Planning and Programme Design
- Climate Change and Health

Feb 2014 – Mar 2014 **London School of Hygiene and Tropical Medicine, United Kingdom**
- Tropical Environmental Health: Water Sanitation and Hygiene (WASH)

Additional University Experience: Course Representative for class 2013

Sep 2008 – Jun 2012 **Bachelors in Biomedical Science, Royal Holloway University of London, United Kingdom**
Bachelor of Science (BSc) with Upper Second-Class Honours
Relevant Courses: Biology of Parasitic Diseases, Molecular Basis of Diseases, Human Physiology in Health and Disease and Clinical Diagnosis of Disease

Relevant University Projects & Experience

Individual Research Project: *Gene Repair through Gene Therapy: Determining the Specificity of Zinc-Finger Nuclease Targeting Mouse Prkdc*

The project was based on determining the specificity of zinc-finger nucleases by analysing 10 different zinc-finger treated off-target sites in mouse DNA.

- The success of my project led to it being taken up by Sangamo BioSciences Inc. for further research.
- Exhibited a poster regarding my project at the British Society for Gene and Cell Therapy 2012 annual conference in London.
- Gained independent lab experience, the ability to analyse and evaluate results, while learning how to make effective modifications in order to obtain reliable data.

Dissertation: *To Investigate Different Approaches for the Eradication of Malaria through Vaccine Development*

The study involved a literature review of the current research that was being undertaken in the field of antimalarial vaccine development. Mainly focusing on the RTS,S/A01 pre-erythrocytic vaccine as well as others targeting different stages in the parasitic cycle. The RTS,S/A01 vaccine was, at the time the only one that had progressed to Phase III trials.

Rare Disease Day – Volunteer
I volunteered for the Rare Disease Day event hosted by Royal Holloway University of London for two years in a row (2011 and 2012), to raise awareness for the need of further research and funding.

Sep 2005 – Feb 2008 **Medical Qualification, Semmelweis Medical University, Budapest, Hungary**
Studies continued in the field of Biomedical Science at Royal Holloway University of London
Key Subjects studied: Medical Biology, Statistics, Histology, Anatomy, Embryology, Medical Chemistry, Medical Physics and Latin

Aug 2003 – May 2005 **International Baccalaureate (IB) Diploma, Vienna International School, Austria**

Work Experience

Sep 2015 – Aug 2016 **Africa AHEAD, Kigali, Rwanda**
Monitoring and Evaluation Intern/Assistant to the CEO
www.africaahead.org

An international NGO focusing on sustainable hygiene behaviour change in rural communities through the implementation of Community Health Clubs.

Responsibilities and Achievements included:

- Managing of website data on monitoring Community Health Clubs in Rusizi District.
- Assisting with the design of the questionnaire for monitoring Community Health Club participants on behaviour change using a household survey.
- Contributing to the preparation of proposals for USAID, Rwanda Development Board and report writing on field activities.
- Assigned as project manager, for giving feedback to IT company, Overtone Limited, on the design of the data collection tool, using the open source software ODK Collect.
- Supervised with the training of the environmental health officers on using the data collection tool.
- Monitoring and evaluating environmental health officers' performance in the field and providing feedback.

May 2015 – Aug 2015 **Body & Soul – Leading UK HIV Charity, London, United Kingdom**
Case Worker
http://www.bodyandsoulcharity.org

A London based charity supporting individuals living with or affected by HIV, offering various free workshops, counselling sessions, legal advice once a week.

Responsibilities and Achievements included:

- Carrying out advocacy work for individual charity attendees.
- Actively researching different services in the UK available to HIV positive individuals are eligible for e.g. applying for healthy food support programmes, small grants for HIV positive individuals as well as housing benefits.
- Advocating on behalf of members with various governmental bodies in the UK e.g. social services, immigration office.

Oct 2014 – Dec 2014 **Irise International in collaboration with Kampala International University, Ishaka and Jinja, Uganda**
Research Assistant
http://www.irise.org.uk

A UK based charity carrying out a research study on the impact of menstrual hygiene management on the school attendance of girls in Bushenyi District, Uganda.

Responsibilities and Achievements included:

- Gaining experience with quantitative data collection and data analysis of school surveys.
- Taking charge of logistical planning for my research team in a resource poor environment.
- Training, assessing and managing local personnel for carrying out field research.
- Voluntarily assisting local university students with scientific research. This included:
    - Helping a Biomedical student to prepare a publication on local remedies for treating Diabetes related symptoms.

Summer 2007 **National Health Centre Kassandria, Greece**
Internship – Medical Assistant

Summer 2005 **Dr Michael Theodorakis Private Medical Clinic, Vienna, Austria**
Internship – Medical Assistant

Publications

Ajagun-Ogunleye, M. , Tirwomwe, M. , Mitaki, R. , Ejekwumadu, J. , Kasozi, K. , Pantoglou, J. , Mbiydzenyuy, E. and Mitaki, N. (2015) Hypoglycemic and High Dosage Effects of Bidens pilosa in Type-1 Diabetes Mellitus. Journal of Diabetes Mellitus, 5, 146-154. doi: 10.4236/jdm.2015.53018.

Pantoglou, J., Abdul-Razak, HH., Rocca, C., et al. (2012) Specificity of a zinc-finger nuclease targeting mouse Prkdc: preparation of samples for analysis of off-target cutting. Human Gene Therapy. [Online] 23 (5), A18-A18.

Skills

Languages

English (fluent oral and written skills)
German (fluent oral and written skills)
Greek (fluent oral and written skills)
French (basic oral and written skills)

Computer Skills

Stata 13.0 – Experienced user
SPSS 24.0 – Experienced user
Microsoft Office 2017 – Experienced user
R – Basic user
ODK Collect – Basic user
EpiData Software – Basic user