Viral diarrhoea in children under two years of age in Dar es Salaam, Tanzania

Clinical and Molecular epidemiology

Sabrina Moyo

Dissertation for the degree philosophiae doctor (PhD)
at the University of Bergen

2014

Dissertation date: 08.12.2014
Scientific environment

This study was done in collaboration between University of Bergen, Norway and Muhimbili University of Health and Allied Sciences (MUHAS), Tanzania. The University of Bergen provided the funding for the work.
Acknowledgement

This work has been possible with the help, contributions, encouragements and support of a large number people. First and foremost I would like to thank my main supervisor Nina Langeland for believing in my abilities when I met her during the year of my specialisation in Microbiology (2006-2007), for that is when I started the journey to do research. She encouraged me not only to do research but also to pursue PhD. Words are not enough to express how thankful I am!. Thank you for believing in me and I highly appreciate your support and your guidance throughout the years of my training.

I also thank all my co-supervisors; Kurt Hanevik, thank you for being a good role model in research, for your contribution in my research studies and for always being there whenever needed help!;

Bjørn Blomberg for your guidance and support through different stages of this work I appreciate your support and contribution of this work;

Prof. Samuel.Y Maselle, thank you for being a mentor to me and advising me in my academic life. You have always encouraged me; I appreciate the guidance and support I received from you. Without the contribution from all my supervisors this work wouldn’t be possible.

I am grateful to all members of the team who worked on this project of diarrhoea diseases in children; Øyvind Kommedal at the Department of Microbiology, Haukeland University Hospital and Vainio Kirsti at Division of Infectious Disease Control, Norwegian Institute of Public Health, Oslo, Norway. You have contributed ideas in making sure that the project goes well. Special thanks to Svein Arne Nordbø of Department of Microbiology, University Hospital of Trondheim for his collaboration in this work.

I wish to thank Mary Ngange and Elisabeth Manda from the Section of Microbiology of Central Pathology Laboratory, (MNH) and Steinar Sornes from the department of clinical science (UiB) for their technical support.

I thank all my office mates for being so good to me, Marit Gjerde Tellevik, Astrid Spjelnaes and Christel Gill Haanshuss) during the last months of my work at the Center for Tropical Infectious Diseases at the Haukeland University Hospital, Bergen,
I thank the nurses who helped data collection at the hospitals where the study was conducted.

I thank all my friends here in Bergen for their good company, especially Dr. Hawa Nalwoga, thank you for being a good friend and a sister to me.

I am grateful to all mothers/guardians of children who consented to participate in this study, without whom this work would have been impossible.

I wish to acknowledge the former and the current head of department Prof M.I Matee and Prof Said Aboud and other members of department of Microbiology and Immunology Prof F.S Mhalu, and Prof E.F. Lyamuya, for allowing me to come and further my education here in Norway.

I thank my parents Angelina Sepeku and Bernard Sepeku and my siblings Josephine and Edmund Sepeku for their support during all the period of my study.

I give special thanks to my dear husband Joshua, thank you for encouraging me to take this opportunity to advance my carrier, for your support and for your patience during my training.

Most of all, I give glory and praise to my lord and saviour Jesus Christ for the grace and comfort through all the hardships that I have gone through during all the time of my studies.
Summary

Introduction
Diarrhoeal disease is one of the leading causes of illness and death in children below five years, especially in developing countries. It is estimated to cause 0.7 million deaths among children younger than 5 years of age, most of which occur in children below two years age. Viruses such as rotavirus, norovirus and adenovirus are known to be major causes of infectious diarrhoea.

Objectives
This study was done to investigate the molecular epidemiology of rotavirus, norovirus and adenovirus and their association with demographic and clinical parameters including HIV and in children under two years of age in Dar es Salaam, Tanzania.

Methods
This was a case control study conducted in Tanzania between August 2010 and July 2011. Cases were children admitted due to diarrhoea at three major hospitals of Dar es Salaam. Controls were children without history of diarrhoea for one month prior the study. These were either attending child health clinics for immunisation and growth monitoring or admitted due to diseases other than diarrhoea. Questionnaires were used to obtain demographic and clinical information. Stool specimens were collected from each study participant. We used ELISA methods for detection of rotavirus and adenovirus from stool specimens. Real time PCR was used for detection of norovirus. PCR followed by sequencing were used for typing all the three viruses.

Results
Viruses were detected in nearly half (48.8%) of the cases and in 17.5% of the controls. Dual viral infection was found in 7.4% and 8.1% in cases and controls respectively. Among cases, rotavirus was the most common detected virus (32.5%) followed by norovirus (18.3%), and adenovirus (3.5%). In controls the prevalence of rotavirus, norovirus and adenovirus was 7.7%, 9.2% and 2.4% respectively. The peak
prevalence of virus infection in cases was seen in the age group 7-12 months, 52.0%. Rotavirus G1P[8] predominated in cases and controls. Norovirus GII.4 predominated both in cases (87.9%) and controls (56.5%). Enteric adenoviruses contributed 50% in cases and 46% in controls. Rotavirus was more prevalent during cold than hot months during the study period. Rotavirus infection was significantly lower in HIV infected (15.4%) than HIV uninfected children (55.3%). There were no significant associations between HIV and either norovirus or adenovirus infection. There were no significant differences observed between viral infection and gender or parent level of education.

**Conclusion**

Viruses contribute to a high burden of diarrhoea in hospitalised children of Dar es Salaam, Tanzania. We found a high frequency of asymptomatic carriage of diarrhoea-associated viruses in healthy children; these children may be a source of infection in the community. We observed a large genetic diversity of the three viruses investigated in this study. Pre-rotavirus vaccination data showed predominance of rotavirus genotype G1. Our study suggests that rotavirus may not be an opportunistic pathogen in children infected with HIV.
List of Publications

Paper I


Paper II


Paper III

**Moyo SJ**, Kurt Hanevik, Bjørn Blomberg, Oyvind Kommedal, Svein Arne Nordbø, Samuel Maselle, Nina Langeland. Detection and molecular characterisation of human adenovirus in diarrhoeic children of Dar es Salaam, Tanzania; a case control study (Submitted)
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cAMP</td>
<td>Cyclic-AMP</td>
</tr>
<tr>
<td>CBC</td>
<td>Community based controls</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribo Nucleic Acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme Linked Immunosorbant Assay</td>
</tr>
<tr>
<td>EM</td>
<td>Electron Microscope</td>
</tr>
<tr>
<td>HAdV</td>
<td>Human adenovirus</td>
</tr>
<tr>
<td>HBC</td>
<td>Hospital based controls</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>HBGAs</td>
<td>Histo-blood group antigens</td>
</tr>
<tr>
<td>IgA</td>
<td>Immunoglobulin A</td>
</tr>
<tr>
<td>LA</td>
<td>Latex agglutination</td>
</tr>
<tr>
<td>MUHAS</td>
<td>Muhimbili University of Health and Allied Sciences</td>
</tr>
<tr>
<td>MNH</td>
<td>Muhimbili National Hospital</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>ORS</td>
<td>Oral Rehydration Solution</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse Transcriptase - Polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribo Nucleic Acid</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA–dependent RNA polymerase</td>
</tr>
<tr>
<td>RRV-1</td>
<td>Rhesus rotavirus vaccine candidate</td>
</tr>
<tr>
<td>VP</td>
<td>Viral proteins</td>
</tr>
<tr>
<td>VLPs</td>
<td>Virus-like particles</td>
</tr>
<tr>
<td>WHO</td>
<td>World health organization</td>
</tr>
</tbody>
</table>
Table of contents

Scientific environment ................................................................. 2
Acknowledgement ........................................................................ 3
Summary ...................................................................................... 5
List of publications ........................................................................ 7
Abbreviations ............................................................................... 8

1. Introduction .............................................................................. 11
   1.1 Burden of diarrhoeal disease ............................................. 11
   1.2 Risk factors for viral diarrhoea ......................................... 12
   1.3 Seasonality of viral diarrhoea .......................................... 12
   1.4 Viral causes of diarrhoea .................................................. 13
      1.4.1 Historical background ............................................... 13
      1.4.2 Rotavirus ................................................................ 14
      1.4.3 Norovirus .............................................................. 18
      1.4.4 Adenovirus ............................................................ 22
   1.5 Mechanisms of virus induced diarrhoea ........................... 24
   1.6 Management of viral diarrhoea ......................................... 25
      1.6.1 Treatment of viral diarrhoea ..................................... 25
      1.6.2 Preventive measures ............................................... 25

2. Aims of the thesis ................................................................. 26

3. Rationale of the study .......................................................... 27

4. Materials, methods and methodological considerations ......... 27
   4.1 Study design and study sites ............................................ 27
   4.2 Study population .......................................................... 27
   4.3 Specimen collection and storage ..................................... 28
   4.4 Study subjects ............................................................. 28
   4.5 Laboratory methods ....................................................... 30
      4.5.1 Extraction of total nucleic acid ......................... 30
      4.5.2 Detection of rotavirus (Paper I) ......................... 31
      4.5.3 Detection of human adenovirus (Paper III) .......... 32
      4.5.4 Detection of norovirus (Paper II) ....................... 32
      4.5.5 PCRs for typing rotavirus and adenovirus .......... 33
      4.5.6 PCRs for genotyping of norovirus ............... 34
      4.5.7 Sequencing and phylogenetic analysis of rotavirus, noro-
virus and adenovirus.

4.6 Statistical analysis

4.7 Ethical consideration

5. Summary of results

5.1 Characteristics of the study population

5.2 Overall prevalence of viruses

5.3 Prevalence of enteric viruses in different age groups

5.4 Paper I

5.5 Paper II

5.6 Paper III

6. Discussion

6.1 Discussion of the main findings

6.1.1 Prevalence of enteric viruses

6.1.2 Mixed viral infections

6.1.3 Distribution of enteric viruses by age groups

6.1.4 Molecular epidemiology of viral infections

6.1.5 Seasonality of viral infections

6.1.6 Association between enteric viruses and HIV

6.2 Methodological discussion

6.2.1 Study design, study population and study participants

6.2.2 Laboratory methods

7. Study limitations

8. Conclusions

9. Recommendations

10. Future perspectives

References

Papers 1-III

Appendices

i) Questionnaire for cases

ii) Questionnaire for controls
1. **Introduction**

1.1 **Burden due to diarrhoea diseases**

Diarrhoeal disease is one of the leading causes of illness and death in children under-five years in the world especially in developing countries. Studies conducted at different times have shown a substantial decrease in mortality due to diarrhoea in developing countries over time (Figure 1) [1-5]. In the year 1976 diarrhoea was estimated to cause five million deaths in children in developing countries [6], but recent studies have shown a decrease in mortality due to diarrhoea, with current estimates of 700,000 deaths annually among children younger than 5 years of age [5]. Of these deaths, 72% occur in children below two years of age. Despite the observed 13 each child experiencing an average of three episodes of diarrhoea per year [5, 7]. Apart from high morbidity and mortality associated with diarrhoeal diseases, diarrhoea also present an economic burden for the developing countries, with more than a third of the hospital beds for children being occupied by patients with diarrhoea, often involving the use of expensive intravenous fluids and drugs [8].
1.2 Risk factors for diarrhoea

Unlike in developed countries, children living in developing countries experience more diarrhoea episodes, severe episodes with dehydration, and a higher death rate. Several factors have contributed to this. Lack of accessibility to clean and safe drinking water, lack of appropriate sewage disposal, crowding and exposure to farm animals (free roaming chickens and pigs), poor hand hygiene and poor sanitary conditions, inaccessibility to medical care and low educational level [9, 10]. Poor handling of weaning food is also a factor which predisposes children to diarrhoea. If weaning food is prepared under unhygienic conditions, it is frequently heavily contaminated with pathogens and hence becomes a major factor in the cause of diarrhoeal diseases [11].

Lack of exclusive breast feeding was found to be associated with diarrhoeal diseases in young infants [12, 13]. Exclusively breastfed children get less diarrhoea episodes compared to no-breastfed infants for the first three months of life. This is because exclusively breast fed children are unlikely to take any other foods or fluids which are likely to be contaminated. Prolonged breast feeding also reduces the incidence or severity of certain types of diarrhoea diseases such as shigellosis and cholera [14-18]. Underlying conditions or host factors such as malnutrition, which is more common in developing countries, increase the risk of diarrhoea and severity [19, 20]. Diarrhoea is also a major source of morbidity and mortality in HIV infected children, particularly in developing countries [21-23].

1.3 Seasonality of viral diarrhoea

Seasonality is an important feature of rotavirus diarrhoea. Reports have shown that changes in the environmental conditions, such as humidity, temperature cycles, rain patterns and winds, are associated with seasonality of infectious diseases [24]. Globally, rotavirus infection is present all the year round, which suggests that low-level transmission could maintain the chain of transmission. Before rotavirus was discovered, the disease was called ‘winter diarrhoea and winter gastroenteritis [25-27]. In temperate climates seasonality of rotavirus infection has been well demonstrated [28, 29], where rotavirus is most frequently detected in the winter and rarely identified
in the summer months [25]. In contrast, in the tropics, it is detected all year round, with seasonal trends that are not well defined [25]. However, recently two studies demonstrated the inverse relationship between temperature and rotavirus incidence in the tropics, and the findings concur with those of temperate countries [26, 30].

Norovirus infections occur in epidemics as well as in sporadic cases. Seasonality of norovirus is poorly understood [24]. In temperate countries, norovirus epidemic characteristics, and timing, are consistent from year to year, with a peak incidence during the winter time. However, outbreaks of norovirus also occur during the summer [31], although at a reduced rate and with an absence of epidemic spread to geographically remote areas [24]. In contrast, in tropical countries, norovirus infection is observed all year round and does not show a clear seasonal pattern [32]. The observed temporal distribution of norovirus infections is most likely based on biological, environmental and behavioural factors that regulate the transmission, virulence and persistence of the virions in host populations [24].

Adenovirus of all types are circulating year round, but sometimes marked seasonal pattern has been reported, especially for non-enteric adenovirus types [33]. A recent study conducted in Bangladesh reported high prevalence of adenovirus infection during the rainy season [34]. Other studies have shown no consistent seasonal pattern of enteric adenovirus infection [35-37].

1.4 **Viral causes of diarrhoea**

Infectious diarrhoea can be caused by a wide range of viruses, bacteria, or parasites [38]. Studies have shown that diarrhoea can be caused by a single pathogen or a mixture of two or more pathogens [39, 40]. This thesis is dealing with diarrhoea caused by three common viruses, rotavirus, norovirus and adenoviruses.

1.4.1 **Historical background**

Viruses have been suspected of being important causes of gastroenteritis since the 1940s, but the aetiology remained unknown in most cases [41, 42]. In 1972 Norwalk virus became the first viral agent identified to cause diarrhoea, using immune electron microscopy (EM) after an outbreak of diarrhoea [43]. One year later rotaviruses were
identified as a cause of diarrhoea in children [44, 45]. In 1975 astroviruses were detected in the stools of children with acute diarrhoea [46, 47]. This was followed by the discovery of enteric adenoviruses [48], and human caliciviruses of two different genogroups now called Norwalk-like viruses (NLVs) and Sapporo-like viruses (SLVs) [49]. Other viruses which have been linked with gastroenteritis in humans are coronaviruses, Aichi viruses, picobirnaviruses, pestiviruses and toroviruses which produce diarrhoea in animals, and they are emerging as causes of viral gastroenteritis in humans, according to several studies [50-59]. The detailed structure and classification of the three viruses rotavirus, norovirus and adenovirus, which form this thesis, are described below.

1.4.2 Rotavirus
Rotavirus is the leading cause of severe diarrhoea in children worldwide [60]. It is estimated to cause 36% of diarrhoea hospitalizations among children aged less than 5 years, according to WHO rotavirus surveillance networks [61]. Rotavirus diarrhoea is estimated to cause 453,000 deaths in children younger than five years, more than half occur in developing countries of Africa and Asia [62].

Structure and classification
Rotaviruses are non-enveloped with icosahedral symmetry belonging to the genus Reoviridae [63]. Figure 2 shows the structure of rotavirus. Rotavirus possesses a genome consisting of 11 segments of double-stranded RNA, which are 18,555 nucleotides in total, enclosed in three protein layers. The genome segments encode six structural viral proteins which make viral particles (VPs) and five non-structural proteins (NSPs). Most segments encode a single polypeptide. Rotavirus NSPs have various functions in the virus replication cycle, in pathogenesis and evasion of the host immune response [64, 65].
The inner layer is composed of the core shell (VP2), the viral RNA-dependent RNA polymerase (VP1), and RNA capping enzyme (VP3), as well as the 11 double stranded RNA genome segments. The middle protein layer, also known as inner capsid, is composed of VP6 which surrounds the core shell. The VP6 is used for rotavirus classification and determines reactivity for the seven major groups (A through G) of rotavirus [64]. Rotavirus groups A, B and C are known to infect humans and various animals. Rotaviruses D, E, F and G so far have only been identified in animals, mostly birds. Epidemiologically, rotavirus group A is the most important for human infection and disease, responsible for more than 90% of all rotavirus infections. The outer protein layer, also called outer capsid, is composed of two structural proteins, VP4 and VP7. The VP7 makes up the outer capsid protein shell while VP4 forms spikes that protrude through the outer capsid shell [65]. The VP4 (a protease cleaved protein, P type antigen) and VP7 (a glycoprotein, G type antigen) induce neutralizing antibodies, correlate with protective immunity and serve as the basis of rotavirus typing system [66-68]. VP7 types are classified as serotypes by neutralization assays or as genotypes by sequencing. The 2 assays produce the same results, therefore rotaviruses are referred to by their G serotype alone e.g., G1, G2, G3, etc. [65, 69]. VP4 serotypes are also classified by neutralization and sequencing assays, but the results do not always agree, so there is a dual system for P typing. P serotypes are referred to by their serotype numbers (e.g., P1, P2) and P genotypes are denoted in brackets (e.g., P[8], P[4]). P genotyping is the most widely used method for classification because of difficulties in standardizing VP4 serotype assays [65, 69]. Currently there are 27 G genotypes and 35 P genotypes identified in human rotaviruses [70].
Molecular epidemiology of rotavirus

Globally, 80–90% of the childhood rotavirus disease burden are caused by genotypes G1P[8], G2P[4], G3P[8], G4P[8], and G9P[8] [71]. Recently rotavirus genotype G12 has been identified as an emerging genotype [72]. Three African countries; South Africa, Malawi and Nigeria, has reported the emergence of G12 occurring in combination with P[6] or P[8] [73-75]. Rotavirus genotype G8 has been reported worldwide, with high prevalence in some African countries [71, 76-83].

Vaccine status

Efforts to produce rotavirus vaccine started in early 1980s. The first rotavirus vaccine candidate was RIT4237, a live bovine rotavirus vaccine [84]. A single dose of this vaccine candidate was given orally before the winter epidemic season. The vaccine was given at birth and was able to induce 82-88% protection against severe disease in the first 2-3 years of life. Breastfeeding did not suppress the uptake of vaccine [84]. The vaccine was not pursued because of its failure in clinical trials conducted in the African region [85, 86]. A second vaccine RRV-1 Rhesus rotavirus vaccine candidate was an alternative to bovine rotavirus vaccine. This vaccine was later reasserted with human rotavirus to express human rotavirus antigen G1, 2, 3 and 4 and Rhesus G3. It was then known as tetravalent Rotavirus vaccine RRV-TV or Rotashield ™ [87]. In trials this oral vaccine was given in 3 doses and was able to induce 80-100% protection against severe disease. The vaccine was licenced in USA in 1998 and was discontinued one year later because it increased the risk of intussusception (at least 25-fold) within the first 10 days after administration [88].

From 2006 there are two licensed oral rotavirus vaccines which as in use, namely RotaTeq™ and Rotarix™. RotaTeq™ is a Bovine-Human re-assorted rotavirus vaccine produced by Merck. This pentavalent vaccine express human G1, 2, 3, 4 and human P8 on the Bovine W3 backbone [89]. The vaccine is given in three doses, trials in USA and developed countries have shown that the vaccine can induce 74% protection against any rotavirus diarrhoea, 87% protection against diarrhoea requiring a visit to physician and approximately 100% protection against severe rotavirus diarrhoea. The vaccine is not affected by breastfeeding and vaccine shedding has been
reported as infrequent and at a low level. Rotarix™ is a human rotavirus vaccine produced by GlaxoSmithKline. This monovalent vaccine is given in two doses; it gives 85% protection against severe diarrhoea and 100% protection against most severe diarrhoea [90, 91]. This vaccine is shed in greater amounts than RotaTeq, the bovine-derived vaccine. A decrease in morbidity and mortality attributable to rotavirus diarrhoea has been reported in high and middle income countries that have introduced rotavirus vaccines [92-99].

**Transmission and clinical symptoms of rotavirus infection**

Rotaviruses are mainly transmitted by the faecal-oral route [100]. Some studies speculate that transmission also occur through respiratory route [25, 100-102]. This theory is supported by the fact that there is rapid acquisition of rotavirus antibodies in the first few years of life in all settings regardless of hygiene standards [42]. There has also been reports of large outbreaks in which faecal-oral route could not be documented [100, 103]. In addition, there are reports of respiratory symptoms among patients with rotavirus diarrhoea [102].

The incubation period of rotavirus diarrhoea is estimated to be less than 48 hours [100]. Rotavirus produces spectrum of responses from subclinical infection to mild and severe rotavirus diarrhoea with dehydration [100]. Vomiting may occur and may precede diarrhoea symptoms. Diarrhoea symptoms can last between one and nine days [100].

Rotavirus can produce chronic symptoms in children who are immunocompromised causing prolonged shedding of the virus in stool [100, 104].

Rotavirus infection has also been associated with intussusception, which may occur when a child is infected with wild type rotavirus [105, 106]. The theory is supported by the report of intussusception following oral administration of quadrivalent rhesus rotavirus based vaccine [88]. A study conducted in Japan has also found that infants and young children with intussusception had evidence of rotavirus infection, which was also detected in stool [105, 106]. The shedding of rotavirus in stool can last between four to 29 days when measured by ELISA or four to 57 days if PCR is used to measure the virus in stool [107].
**Immune response to rotavirus infection**

The immunological mechanisms responsible for protection against infection by rotavirus are still not well known. Intestinal IgA is probably the most important mechanism for long-term protection against rotavirus [108]. Studies showed that the first infections are generally the most severe, with severity decreasing as the number of infections increases [109-113]. Furthermore, protection conferred by infection was greatest against moderate to severe disease, less against mild illness, and least against asymptomatic infection. Asymptomatic infection provides protection to a degree comparable to that achieved by symptomatic infection [109, 110, 114, 115]. A trend was observed where repeated infections with the same G serotype is less likely to occur, suggesting the first infection with rotavirus elicits a homotypic neutralizing antibody response [116], with heterotypic responses in subsequent infections [109, 110, 116]. The presence of neutralizing antibodies directed towards the proteins VP4 or VP7 does not correlate with protection against this disease [117]. Another study has also shown that protein NSP4 produces a cellular immunity-mediated response [118].

**1.4.3 Norovirus**

Noroviruses are among the leading cause of sporadic cases and outbreaks of acute gastroenteritis across all age groups [119-121]. It is estimated to cause 18% of all cases of sporadic acute gastroenteritis in children younger than 5 years and older children and adults were associated with norovirus [122]. In developing countries, noroviruses are estimated to cause more than 1.1 million hospitalisations and 200,000 deaths in children <5 years of age each year [123]. When norovirus infection occurs in outbreaks, they are usually difficult to control because of the prolonged shedding of the virus, environmental stability and resistance to disinfection, low infectious dose, and contribution of virus-induced vomiting to the spread of the infectious agent [124].

**Structure and classification**

Noroviruses belong to the genus Norovirus in the family *Caliciviridae*. The virions are composed of a single structural capsid protein, with icosahedric symmetry [125]. This
capsid protein has a continuous shell with protrusions in the shape of an arch (Figure 3). A key characteristic is the existence of cup-shaped depressions, situated on the axes of the icosahedron, from whose Latin designation, calyx, the virus derives its name [126]. The genome of the norovirus consists of positive-sense, single-stranded RNA of approximately 7.5 kbp. The genome is organized into three open reading frames (ORFs). ORF1 at the 5’end is the largest and encodes non-structural polyproteins such as RNA-dependent RNA polymerase (RdRp), proteinase and NTPase (Figure 3B). It is followed by ORF2 encoding a single viral capsid protein [127] and ORF3 at the 3’ terminus encoding a basic protein which regulate the expression and stability of the viral capsid protein [128]. Based on RNA polymerase or capsid gene sequence analysis noroviruses in humans have been divided into three genogroups (G), GI, GII and GIV, whereas GIII and GV Noroviruses have been found in animals only. Within GI and GII, several genotypes are recognized [129, 130]. To date 14 GI and 29 GII polymerase genotypes and 8 GI and 23 GII capsid genotypes have been described [131].

**Molecular epidemiology of norovirus**

Most human norovirus infections are caused by genogroup GII, which account for an average of 96.0% [131, 132]. Norovirus genotype (GII.4) predominates and is
responsible for approximately 55–85% of the norovirus gastroenteritis cases worldwide [131, 132]. Norovirus GII.4 has also been reported to cause severe outcomes of hospitalizations and deaths [133]. Within norovirus GII.4 seven different variants have been reported since 1990s to early 2013, these variants were associated with global epidemics of gastroenteritis [132]. A review by Hoa Tran and co-workers has shown that norovirus GI occurs less common, at an average of 3.6% [131]. Norovirus GI is also reported to be associated with water-borne outbreaks [132]. Mixed norovirus infections of GI and GII were also reported with an average of 0.4% [131]. Some countries have reported high prevalence of norovirus GI in approximately one third of norovirus strains; in Madagascar, Egypt, Yemen and Thailand [134-136].

**Vaccine status**

Currently there is no licensed norovirus vaccine. Promising vaccine candidates are under development. Studies have shown that both orally and nasally administered virus-like particles (VLPs) can be well tolerated and are immunogenic in phase 1 trials [137, 138]. A nasally administered investigational GI.1 norovirus VLP vaccine was also shown to confer 47% protection against norovirus gastroenteritis in healthy adults in a human experimental challenge model [139]. Recently, a bivalent intramuscular norovirus virus-like particle (VLP) vaccine was reported to be well tolerated and immunogenic in adults [124].

**Transmission and clinical symptoms of norovirus infection**

Noroviruses can infect humans via multiple routes such as the oral route, through contact with fecal matter or aerosolized vomitus from infected people, as well as contaminated surfaces, food, or water [140-142]. Transmission of norovirus is very efficient because of certain characteristics of the virus that enhance their ability to spread. The virus is highly infectious, and a very small inoculum, as few as 10–100 virions, is sufficient to cause infection [141, 143]. In addition, the virus is relatively stable in the environment. It survives freezing, heating to 60° C, and exposure to chlorine [141]. Norovirus can cause food-borne related outbreaks in healthcare settings such as nursing homes and hospitals where the virus is predominantly spread
from person to person [142, 144]. Norovirus is also among major causes of sporadic gastroenteritis in children and adults [142].

After an incubation period of 12-48h, the illness may start with one or all of the following symptoms: projectile vomiting, non-bloody diarrhea, nausea, abdominal cramps, and low-grade fever [142, 145]. In healthy individuals the duration of symptoms is usually 12-72 hours [141] and in most patients the disease is self-limiting [142]. However, young children are at increased risk for more severe and prolonged illness leading to hospitalization [142, 146], while for immuno-compromised patients the disease is increasingly recognized as an important cause of chronic gastroenteritis [146]. The shedding of the virus in stool can continue for up to three weeks [144].

**Immune response to norovirus infection**

Protective immunity to norovirus is complex and also not well understood. It is hypothesised that innate host factors and acquired immunity contributes to the susceptibility of norovirus infection [147]. Some studies have shown that infected volunteers were susceptible to reinfection with the same or different strains [147, 148]. Furthermore, it was demonstrated that people with pre-existing antibodies are not protected from infection unless repeated exposure to the same strain occurs within a short period of time. In these studies it was also noted that homologous antibody protection is short lived lasting from 8 weeks to 6 months [147, 148]. Proposed receptors for norovirus are Histo-blood group antigens (HBGAs), including H type, ABO blood group, and Lewis antigens. Strain-specific susceptibility to norovirus infection has been associated with the expression of HBGAs [144, 149-154]. In addition, resistance to norovirus infections has been associated with mutations in the 1, 2-fucosyltransferase (FUT2) gene which lead to lack of expression of HBGAs on the surface of intestinal cells [149-151, 155]. People with normal FUT2 gene and express HBGAs antigens are termed as "secretors". People with mutations in the FUT2 gene who will not express HBGA are termed as "non-secretor", these are people who are less susceptible to norovirus infection. However, secretor status alone does not entirely explain the observed differences among infected and uninfected people, for all strains of norovirus. Therefore it is most likely that other mechanisms
of immunity are involved and this needs further research. There are also evidence suggesting that new GII.4 variants evolve to escape the build-up of both acquired and innate immunity resistance in the human population [156, 157].

1.4.4 Adenovirus

Human adenovirus causes acute diarrhoea sporadically, as well as in outbreaks. Besides acute diarrhoea, adenoviruses cause other diseases such as respiratory diseases, conjunctivitis and haemorrhagic cystitis [158]. Adenoviruses have also been associated with persistent infections in both immuno-competent and immuno-compromised individuals [158]. In young children enteric adenoviruses are reported with variable prevalence of infection. In developed countries, the prevalence varies from 1% to 8% [36, 159-162], whereas in developing countries, prevalences from 2–31% has been reported [163-167].

Structure and classification

Adenoviruses represent the largest non-enveloped viruses measuring 90–100 nm, icosahedral in shape. The nucleocapsid contains a double stranded DNA genome. They belong to the family Adenoviridae and genus Mastadenovirus. Their name derives from their initial isolation from human adenoids in 1953 [158]. Figure 4 shows the structure of adenovirus. The protein capsid is composed of 252 capsomers with hexones/ pentones and structures called fibers that protrude to the outside. The hexones contain proteins II, VI, VIII, and IX, which participate in the stability and assembly of the viral particle. The pentone proteins (III and IIIa) have the function of cellular penetration. The protruding fibers are hemagglutinins and are responsible for binding the virus to receptors [168-170]. The core is made of at least eight proteins which maintain the integrity of the genome, and participate in enzymatic activity [158]. The adenovirus genome is linear, non-segmented double-stranded (ds) DNA that is between 26 and 48 Kbp [158, 171].
Adenoviruses have a broad range of vertebrate hosts; in humans, there are 52 types of adenoviruses identified, grouped into seven species A to G on the basis of their resistance to neutralisation by antisera to other known human adenoviruses [171, 172] or genome analysis [172].

**Vaccine status**
Currently there is no vaccine for enteric adenovirus in children [173]. However adenoviruses are commonly used as vectors for various other vaccines [173, 174].

**Transmission and clinical symptoms of adenovirus infection**
Enteric adenoviruses are transmitted directly through person-to-person contact via the fecal- oral route, as well as indirect transfer through contact with contaminated surfaces or shared utensils [166]. Human adenovirus causes acute diarrhoea sporadically, especially in children, as well as in outbreaks [175]. The incubation period is approximately 7-8 days [175, 176]. Symptoms include non-bloody watery diarrhoea with or without vomiting. Low grade fever may occur, dehydration and respiratory tract infection may be present [175]. The infection last for 5-12 days, which is longer than other types of viral gastroenteritis, but usually the disease is milder compared to other viruses [175, 177].

**Molecular epidemiology of adenovirus**
The disease pattern of adenoviruses varies according to species. Adenovirus species F, types 40 and 41, has been found to be regularly associated with gastroenteritis and they are referred to as enteric adenoviruses and are responsible for most cases of acute gastroenteritis. Other species such as A (types 12, 18 and 31) [169], C (types 1, 2 and
5) and D (types 28, 29, 30, 32, 37, 43 to 46) [158, 178, 179] have also been associated with diarrhoea.

**Immune response to adenovirus diarrhoea**

After adenovirus infection, type specific neutralizing antibodies develop in most cases. These can provide protection both in the current illness and against reinfections by the same serotype, although patients may continue to eliminate the virus in their faeces for months after an effective humoral response [158].

### 1.5 Mechanisms of virus induced diarrhoea

Enteric viruses, namely rotavirus, norovirus and enteric adenoviruses, are stable at low pH and resistant to digestive enzymes. The viruses infect the intestinal villus and spare the crypt cells [180]. Once the virus is inside the cell, virus replication occurs, which later causes cell lysis and release of viral progeny which will infect neighbouring cells [180]. This cytolytic infection leads to villus atrophy, which will cause decrease in digestive and absorptive capacity of the intestine, and is responsible for the malabsorption component of diarrhoea. Both cellular and viral proteins will induce immune cells to release variety of chemical mediators. The interaction between the immune cells and enteroendocrine systems lead to increased intracellular Ca$^{2+}$ and cAMP, these in turn will result in one or more of the following:

i) Activation of Cl$^{-}$ channels causing increased Cl$^{-}$ secretion into intestinal lumen

ii) Act on cytoskeleton and tight junctions resulting in weaker adherence between enterocytes leading to paracellular leakage and to increased permeability;

iii) Alterations in intestinal motility. This can also be a results of villus ischaemia

iv) Decreased enzyme expression and decreased nutrient and water absorption this is virus-induced down-regulation of the expression of absorptive enzymes [180].

Usually viruses cause diarrhoea by one or more of the above four mechanisms [180].
1.6 Management of diarrhoea

1.6.1 Treatment of viral diarrhoea

Rehydration is the mainstay of treatment for viral gastroenteritis and most cases resolve spontaneously without the need for additional interventions [181-183]. Rehydration can be done by administering oral rehydration solution (ORS) to prevent or correct dehydration which is a consequence of diarrhoea [181]. For children who are unable to tolerate ORS via the oral route (with persistent vomiting), nasogastric feeding can be used to administer ORS. Intravenous fluids may be given in cases with severe dehydration [181]. Currently there is a more effective, lower-osmolality ORS (with reduced concentrations of sodium and glucose, associated with less vomiting, less stool output, and a reduced need for intravenous infusions in comparison with standard ORS) has been developed for global use [184]. Table 1 shows the constituents of ORS solution.

Table 1. Constituents of Oral rehydration solution (ORS)

<table>
<thead>
<tr>
<th>Constituent</th>
<th>mmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sodium</td>
<td>75</td>
</tr>
<tr>
<td>Chloride</td>
<td>65</td>
</tr>
<tr>
<td>Glucose, anhydrous</td>
<td>75</td>
</tr>
<tr>
<td>Potassium</td>
<td>20</td>
</tr>
<tr>
<td>Citrate</td>
<td>10</td>
</tr>
<tr>
<td>Total osmolality</td>
<td>245</td>
</tr>
</tbody>
</table>

1.6.2 Prevention of viral diarrhoea

The following are preventive measures for diarrhoea diseases:

i) Interruption of transmission of the causative agents include measures such as good infection control practices, these are important in hospitals and centres which care for small children in order to interrupt transmission and eventually prevent spread of infection [182, 185]. Improving practices related to the preparation and storage of weaning foods (to minimize microbial contamination and growth). Other measures are availability of safe water supply, washing hands (before and after preparing food or eating), safe disposing of faeces, including infant faeces.

ii) Exclusive breast feeding for 6 months and improved weaning practices are important measures to reduce the incidence of diarrhoea [182].
iii) Measures to improve host defences and thus diminish the risk of diarrhoea include: continuing to breast-feed for at least the first year of life [186], for rotavirus, two licenced vaccines in use have shown to be able to reduce morbidity and mortality [92, 93, 96, 98].
iv) Zinc deficiency is common among children in developing countries and it has been associated with infectious disease including diarrhoea [187-189]. A number of trials have supported zinc supplementation as an effective agent in treating and preventing diarrheal disease [190-193].

2.0. Aims of the thesis

2.1 General aim
Detection and molecular characterisation of viruses in hospitalised diarrhoeic children and asymptomatic controls in Dar es Salaam, Tanzania.

2.2 Specific aims
2.2.1. To determine the prevalence of rotavirus, norovirus and adenovirus in diarrhoeic children (Paper I, II and III).
2.2.2 To determine the prevalence of asymptomatic carriage of rotavirus, norovirus and adenovirus in non-diarrhoeic children (Paper I, II and III).
2.2.4 To explore potential associations between HIV, seasonality, clinical and demographic characteristics with viruses (rotavirus, norovirus and adenovirus) in hospitalised diarrhoeic children compared with asymptomatic controls (Paper I, II, III)
3.0 Rationale of the study
The information generated from this study will provide the prevalence of enteric viruses namely rotaviruses, noroviruses and adenoviruses in childhood diarrhoea in Tanzania. The study will also provide a comprehensive picture of genetic diversity of these viruses which is relevant information required for the development of efficacious vaccines. Furthermore, from early 2013, Tanzania started to implement a live attenuated monovalent vaccine (Rotarix) in the national childhood vaccination schedule. This study describes the molecular epidemiology of rotavirus prior to implementation of vaccination. The study also highlights seasonal variation, impact of HIV and clinical characteristics of viruses causing diarrhoea in children of Dar es Salaam, Tanzania.

4.0 Materials, Methods and methodological considerations

4.1 Study design and study sites
This study was a cross sectional case control study. The study was conducted between August 2010 and July 2011 in Dar es Salaam Tanzania. Dar es Salaam is a city with a population of about five million. The study was conducted at three major hospitals; Muhimbili National Hospital (MNH), Amana and Temeke Municipal Hospitals. MNH, with a bed capacity of 1200, is the largest hospital in the country and serve as a tertiary and national referral hospital. Amana and Temeke are Municipal district hospitals of Dar es Salaam.

4.2 Study population
The study included children less than two years of age. Cases were children admitted due to diarrhoea in the three hospitals studied. Controls were children without history of diarrhoea for one month prior to the study. They were either children attending child health clinics for routine immunisation and growth monitoring, hereafter referred to as community based controls (CBC) or children admitted to the same hospitals as the cases due to diseases other than diarrhoea, hereafter referred to as hospital based controls (HBC). Details of recruitment of study participants, collected
demographic and clinical information, case definition and exclusion criteria are described in paper I and II. HIV testing was done by HIV-DNA PCR as previously described [194]. HIV testing results were obtained from patient files. Weight and length measurements were performed as previous described [195]. Nutritional status of the study participants was assessed using Z-scores according to WHO criteria [2].

4.3 Specimen collection and storage

For diagnosis of viral pathogens causing diarrhoea, specimen can be whole stool or rectal swabs [196]. Whole stool specimens are preferred over rectal swabs because of the higher quantity of virus present in whole stool [197]. Furthermore specimens should be obtained during the acute phase of illness (within 48-72 hours after onset) while the stools are still liquid or semisolid and viral excretion is at its peak [196]. It is also recommended that collected stool specimens should be kept refrigerated at 4°C if testing occurs within 2-3 weeks or if testing is expected to occur over 3 weeks after collection or for archiving purposes, stool samples should be frozen at -20°C or -70°C [196]. At this temperature, specimens can be stored without compromising diagnostic yield. Norovirus RNA can be detected by using molecular techniques after at least 5 years when samples are stored under appropriate conditions[196, 198]. In this study we collected a single whole stool specimen from each child on the day the child was admitted due to diarrhoea or the day of enrolment for controls. A wide mouthed sterile plastic container was used for stool collection. The intention was to obtain at least 10g of stool; however, to be considered acceptable, the sample as an approximation resembles the size of 3 peas. The specimen was immediately placed into a Styrofoam container containing a cold pack and transported to the Central Pathology Laboratory (CPL). Upon arrival in the laboratory, two aliquots of the portion of stool specimen were frozen at -70°C the same day as it was collected. At the end of data collection stool specimens were shipped on dry ice to Bergen, Norway for further analysis.

4.4 Study subjects for paper I, II and III

Sample collection was performed during two seasons, starting in August 2010 and in March 2011, aiming for minimum 300 cases and 300 controls in each period. The
target for cases was reached in January 2011 and in June 2011, while enrolment of controls continued in February 2011 and July 2011, thus enrolment continued for one complete year.

The study included a total of 1303 stool samples from cases and controls. Of these, 1266 samples had sufficient material to extract total nucleic acid and all these samples (705 cases and 561 controls) were used for analysis of Paper II.

A total of 1235 samples were adequate for detection of rotavirus and adenovirus using ELISA (690 cases and 545 controls) and therefore were used for analysis of paper I and III.

For paper I and III the number in each group of controls i.e. CBC and HBC were 310 and 235, respectively. For paper II the number of CBC and HBC were 312 and 249 respectively.

**Figure 5: The flow chart of the number of samples used for each study**
4.5 Laboratory methods

4.5.1 Extraction of total nucleic acid (RNA/DNA)

This study used PCR for detection and typing of norovirus (paper II) and typing of rotavirus and adenovirus (Paper I and III). The reliability of PCR depends on the efficient isolation of nucleic acids together with sufficient removal of inhibitory substances from the sample material [126, 145, 199]. Therefore the method used for nucleic acid extraction has to be well considered. Several commercial assays for the extraction of nucleic acids are available and used in routine diagnostics. The number of samples to be tested affects the choice of the method to be chosen. When dealing with large number of samples, automated method of nucleic acid extraction will be the right choice since it will decreases the hands-on time per sample and improves assay performance including precision [199]. A study by Verheyen et al. evaluated the performance of five automated methods of nucleic acid extraction from stool, these methods were 1; easyMAG, bioMerieux, 2; m2000sp, Abbott, 3; MagNA Pure LC 2.0, Roche, 4; QiaSymphony, Qiagen and 5: sample preparation module of the VERSANT kPCR Molecular System, Siemens. The five automated extraction methods yielded comparable results [199].

In this study we used MagNA Pure LC 2.0, Roche for extraction of total nucleic acid. In order to extract both RNA and DNA from the same sample material we used Magna Pure LC High Performance Total Nucleic Acid Isolation Kit (Roche Applied Science, Mannheim, Germany). Details of the method can be found in paper I. We first tested the method ability to extract both RNA and DNA by testing known norovirus genogroup I and II and giardia positive samples. The testing results could detect all the known positive and negative samples. This method was preferred because of large number of samples if this study and also the instrument was made available for use in this study by Microbiology laboratory, Haukeland University hospital.
4.5.2 Detection of rotavirus (Paper I)

Initially detection of rotavirus was mainly done using electron microscope (EM) [100, 200-202]. This method has an advantage of high specificity because of the distinctive appearance of rotavirus [100]. Although EM is still useful it is limited to reference laboratories [203]. Cell culture is not considered to be useful for diagnostic purposes, as it is technically cumbersome and slow [203]. The latex agglutination (LA) technique although gives rapid results and does not require expensive instruments, but it is reported to have lower sensitivity than the enzyme-linked immunosorbent assay (ELISA) [204, 205]. In addition LA may not be appropriate choice when dealing with large number of samples. ELISA is now well established as a sensitive and specific technique for the detection of rotavirus antigen in stools and was adopted by WHO as the standard diagnostic technique for use in epidemiological studies in developing countries [206]. There are several commercial available ELISA kits for detection of rotavirus which has been evaluated [207]. Some of the ELISA tests are monoclonal antibody-based, these are more costly but have the advantage of highest sensitivity compared to polyclonal antibody-based tests which are the least expensive but may be less sensitive [208]. However, a recent evaluation study found no significant difference in sensitivity of two monoclonal vs. one polyclonal ELISA kits [207]. These ELISA kits are RIDASCREEN® Rotavirus (R-Biopharm AG, Darmstadt, Germany) and the Premier™ Rotaclone® kit both uses monoclonal antibodies raised against rotavirus structural protein VP6. The third ELISA kit is ProSpecT™ Rotavirus, this kit uses polyclonal antibodies directed against rotavirus structural proteins. ProSpecT™ Rotavirus kit was reported by manufacture (Oxoid, Ltd) to have 100% and 99.2% sensitivity and specificity respectively when compared to EM and a commercial EIA kit as gold standard. We therefore chose to use the ProSpecT™ Rotavirus kit for detection of rotavirus in paper I. This ELISA kit was also reported to have 75% and 100% sensitivity and specificity, respectively, when compared to RT-PCR as a gold standard [207]. The ProSpecT™ Rotavirus kit we used is a replacement kit for the widely used rotavirus IDEIA™ Rotavirus EIA (Dako Diagnostics Ltd., Ely, UK), which was discontinued in March 2009. IDEIA™ Rotavirus kit was used in our
previous study [195]. We preferred this kit not only because of good reported sensitivity but also to get comparable results from our previous study.

4.5.3 Detection of human adenoviruses (Paper III)

Enteric adenovirus (type 40 and 41) are fastidious; they cannot be cultivated in cell lines which have traditionally been used for isolation of adenovirus. However, they can be isolated in Graham 293, a human embryonic kidney cell line transformed by adenovirus type 5-sheared DNA [158]. Therefore, cell culture is not routinely used for detection of adenovirus from stool samples. Antigen detection tests such as latex agglutination and ELISA are now widely used for detection of adenoviruses from stool samples. Furthermore, nested PCR which detects all described adenovirus types have recently been used in epidemiological studies. Studies which have used PCR for detection of adenovirus have shown high sensitivity by reporting high prevalence of adenovirus from stool samples [165, 166, 209] compared to studies which have used either LA or ELISA techniques [164, 195, 210-213]. However PCR is relatively expensive compared to ELISA especially when the expected prevalence is low. Among ELISA test kits available, some detect group or species specific adenovirus F40 and F41 and others are designed to detect all the types of adenovirus. Due to the fact that other types of adenovirus than F40 and F41 have been implicated in diarrhoea cases [166, 179], the use of tests which detect other types than enteric adenovirus is important in order not to underestimate the prevalence of adenoviruses. In paper II we used ELISA test kit ProSpecT™ Adenovirus kit. According to the manufacturer (Oxoid, Hants, UK), this kit has sensitivity and specificity of 90.1% and 99.0% respectively. The test utilizes a genus-specific monoclonal antibody designed to detect epitopes common to 51 human adenovirus serotypes (Oxoid, Hants, UK). The testing was performed according to the manufacturer’s instructions.

4.5.4 Detection of norovirus (Paper II)

Isolation of norovirus in cell cultures has not yet been achieved [126]. Although in some laboratories EM is used for detection of noroviruses, the sensitivity of EM is
low, requiring at least $10^6$ viral particles per ml of stool [214]. Therefore direct examination of norovirus using EM is not useful, firstly because of the shedding of low concentration of the virus in stool and secondly it is difficult to differentiate norovirus from other caliciviruses [145]. Immunoassays such as ELISA kits are available and in use for detection of norovirus in stool, however, these test kits are less sensitive compared to RT-PCR for detection of norovirus [215, 216]. Furthermore the sensitivity of ELISA kits is low, less than 70%, therefore results need to be interpreted with caution [142]. Therefore, currently RT-PCR is widely used for detection of noroviruses in stool specimens. There is an increase in the use of real time RT-PCR because it allows rapid detection of the virus [214]. When choosing the RT-PCR for norovirus, there is a need to consider the choice of appropriate primers to be used because of existence of genetic diversity among circulating strains [145]. Several primer pairs from two different regions of the RNA i.e. ORF1, Polymerase region and ORF2, capsid region, have been described [217-225]. In paper II we used single-plex, one step, real time PCR for the detection of norovirus genogroups GI and GII as described in paper II. Primers used in this PCR were previously described [214, 226]. These primers are from the junction of ORF1-ORF2 which is regarded as the most conserved region in the norovirus genome [214]. This method was also chosen because it has already been established and used in the routine microbiology laboratory where the work was performed. Taking into consideration the large number of samples in this study one step real time RT-PCR was the appropriate choice. The details of the PCR parameters can be found in paper II. The advantage of the method chosen used was its ability to detect and differentiate norovirus genogroup I and II at the same time.

4.5.5 **PCRs for typing of rotavirus and adenovirus**

Several methods can be used for typing of rotavirus. ELISA using serotype specific monoclonal antibodies or polyclonal antibodies has previously used to detect specific G and or P types from stool samples [227-230]. Although useful for typing rotavirus, ELISA may be cumbersome when dealing with a large number of specimens and continuous designing of new monoclonal/polyclonal antibodies against new strains is
required. Since the 1990s multiplexed semi-nested RT-PCR have been developed and used in most countries for typing of rotavirus [231-233]. The European rotavirus surveillance network has developed a multiplex semi-nested RT-PCR for rotavirus and the methods are available online. Primers designed in this method have been widely used in developing as well as developed countries with good results [195, 234-236]. For paper I, we therefore adopted and used the same method which we previously used, multiplex semi-nested RT-PCR for G and P typing [237].

Adenovirus serotyping can be done by neutralization or hemagglutination inhibition or by restriction endonuclease analysis of full-length adenovirus DNA. Both methods involve cell culture which is time consuming and also does not favour fastidious adenoviruses [238, 239]. PCR is useful for fastidious adenoviruses i.e. F40 and F41 to be typed. The choice of method depends on the need to type some or all types of adenovirus. In paper II, real time multiplex PCR was used to type 51 types of adenoviruses with primers and TaqMan probes from the conserved region of the hexon-coding gene of adenoviruses [240]. The details of the PCR parameters have been described in paper II.

4.5.6 PCRs for genotyping norovirus

Dual nomenclature systems exist for norovirus genotypes, RNA polymerase region in ORF1 and VP1 sequences [241]. PCR followed by nucleotide sequencing of small regions of ORF1 or ORF2 (CAP) of the norovirus genome are used to genotype norovirus strains [142]. For paper II genotyping of norovirus was done by conventional RT-PCR followed by sequencing of the polymerase region to genotype all norovirus positive samples. This method was chosen because it was previously established, published and used by our collaborators [242]. Randomly selected, 33 samples also underwent real time PCR followed by sequencing of the capsid region using the primers described by Kojima et al. [243]. This was done to confirm and compare the results from PCR of the polymerase region.
4.5.7 **Sequencing and phylogenetic analysis of rotavirus, norovirus and adenovirus**

Sanger sequencing was used for rotavirus, norovirus and adenovirus. This was followed by phylogenetic analysis using the RipSeq interpretation software (Isentio AS., Bergen, Norway) or nucleotide BLAST service (NCBI) for paper I and III. For paper II, genotyping of norovirus and the characterization of norovirus GII.4 variants were done by the norovirus automated genotyping tool [244].

4.6 **Statistical analysis**

Analysis of weight-for age, length for age and weight for length Z-scores was done using EPI Info (USD, Inc., Stone Mountain, GA, USA). The Statistical Package for the Social Sciences (SPSS for IBM-PC, release 20.0; SPSS Inc., Chicago, IL, USA) was used for data analysis, and a p-value of < 0.05 was considered significant. Pearson’s Chi-squared test or Fishers exact test were used to compare differences in prevalence of rotavirus/ norovirus and adenovirus infection in cases and control. For paper I, II and III, the age distribution of children in cases and control was compared using Mann-Whitney U test.

Univariate analysis using cross tabulations was performed to assess relationship of norovirus (paper II) and adenovirus positivity (paper III) with season, demographic and clinical characteristic.

In paper I, we assessed the association between rotavirus infection and clinical/demographic characteristics using odds ratio (OR) in a logistic regression model. Because of small number of HIV tested participants, we performed a second regression analysis which included HIV as a factor; this was a manual, backwards, stepwise regression analysis by elimination of non-significant factors with p value of > 0.05.

4.7 **Ethical considerations**

This study was conducted in accordance with existing ethical guidelines of Muhimbili University of Health and Allied Sciences in Dar es Salaam, Tanzania and the Regional Committee for Medical and Health Research Ethics (REK) in Bergen, Norway. Ethical clearance was obtained from both locations. The permission to conduct the
study was also obtained from the respective Hospital authorities where recruitment of study participants took place. Parent/guardian of the child signed written informed consent form.

Table 2: Demographic and clinical characteristics of the study population

<table>
<thead>
<tr>
<th>Demographic/clinical characteristic</th>
<th>Paper I and III</th>
<th></th>
<th>Paper II</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td></td>
<td>N=690</td>
<td>N=545</td>
<td>N=705</td>
<td>N=561</td>
</tr>
<tr>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>422 (61.2)</td>
<td>296 (54.3)</td>
<td>436 (61.8)</td>
<td>306 (54.5)</td>
</tr>
<tr>
<td>Female</td>
<td>268 (38.8)</td>
<td>249 (45.7)</td>
<td>269 (38.2)</td>
<td>255 (45.5)</td>
</tr>
<tr>
<td><strong>Age groups in months</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-6</td>
<td>191 (27.7)</td>
<td>120 (22.0)</td>
<td>192 (27.2)</td>
<td>129 (23.0)</td>
</tr>
<tr>
<td>7-12</td>
<td>343 (49.7)</td>
<td>211 (38.7)</td>
<td>348 (49.4)</td>
<td>215 (38.3)</td>
</tr>
<tr>
<td>13-18</td>
<td>92 (13.3)</td>
<td>159 (29.2)</td>
<td>90 (12.8)</td>
<td>161 (28.7)</td>
</tr>
<tr>
<td>19-24</td>
<td>64 (9.3)</td>
<td>55 (10.1)</td>
<td>75 (10.6)</td>
<td>56 (10.0)</td>
</tr>
<tr>
<td><strong>Place of residence (District)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinondoni</td>
<td>202 (29.3)</td>
<td>158 (29.0)</td>
<td>211 (29.9)</td>
<td>162 (28.9)</td>
</tr>
<tr>
<td>Ilala</td>
<td>338 (49.0)</td>
<td>296 (54.3)</td>
<td>341 (48.4)</td>
<td>305 (54.4)</td>
</tr>
<tr>
<td>Temeke</td>
<td>150 (21.7)</td>
<td>91 (16.7)</td>
<td>153 (21.7)</td>
<td>94 (16.8)</td>
</tr>
<tr>
<td><strong>Parent level of education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary education</td>
<td>531 (77.0)</td>
<td>413 (75.8)</td>
<td>543 (77.0)</td>
<td>424 (75.6)</td>
</tr>
<tr>
<td>Secondary education</td>
<td>137 (19.9)</td>
<td>127 (23.3)</td>
<td>139 (19.7)</td>
<td>130 (23.2)</td>
</tr>
<tr>
<td>Higher education</td>
<td>22 (3.2)</td>
<td>5 (0.9)</td>
<td>23 (3.3)</td>
<td>7 (1.2)</td>
</tr>
<tr>
<td><strong>Type of diarrhoea</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute diarrhoea</td>
<td>611 (88.8)</td>
<td>NA</td>
<td>627 (88.9)</td>
<td>NA</td>
</tr>
<tr>
<td>Persistent diarrhoea</td>
<td>79 (11.4)</td>
<td>NA</td>
<td>78 (11.1)</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Hydration status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dehydration</td>
<td>521 (75.5)</td>
<td>NA</td>
<td>533 (75.6)</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Nutrition status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>i) Underweight</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malnourished</td>
<td>390 (56.5)</td>
<td>217 (39.8)</td>
<td>402 (57.0)</td>
<td>222 (39.6)</td>
</tr>
<tr>
<td>*ii) Stunting (HAZ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malnourished</td>
<td>469 (68.0)</td>
<td>301 (55.2)</td>
<td>480 (68.1)</td>
<td>306 (54.5)</td>
</tr>
<tr>
<td>*iii) Wasting (WLZ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malnourished</td>
<td>200 (28.9)</td>
<td>111 (20.4)</td>
<td>212 (30.1)</td>
<td>117 (20.9)</td>
</tr>
</tbody>
</table>

NA: not applicable
5 Summary of results

5.1 Characteristics of the study population
Table 2 shows demographic and clinical characteristics of study population for paper I, II and III. HIV testing results were available on 423 children, of whom 33 and 390 children tested positive and negative, respectively.

5.2 Overall prevalence of viruses

At least one of the three viruses was detected in 352 cases, which is 48.8% of the case samples tested. Out of these 352 cases, 326 were infected with only one virus (92.9%) whereas 26 (7.4%) were infected with more than one virus. Among controls at least one virus was detected in 99 (17.6%) of the samples tested, of these 90 (90.9%) had single virus infection and 8 (8.1%) had mixed viral infection.

5.3 Prevalence of enteric viruses in different age groups

Among cases there was significant difference in prevalence of viruses in different age groups. Figure 6 shows the prevalences of viruses in different age groups. The peak prevalence of virus infection was seen in the age group 7-12 months (n=183, 52.0%, P= 0.004, OR=2.15, 95% CI: 1.275-3.63). The prevalence of at least one virus was also significantly higher in the age group of 0-6 months compared to older children.
aged 19 months and above (P=0.016, OR=1.98, 95% CI: 1.14-3.45). Overall prevalence of viruses did not differ significantly among controls.

5.4 Paper I

Genetic diversity of circulating rotavirus strains in Tanzania prior to the introduction of vaccination.


In this study, we included 690 children admitted to hospital due to diarrhoea (cases) and 545 children without diarrhoea (controls). The prevalence of rotavirus was significantly higher in cases (32.5%) than in controls (7.7%, P<0.001). The prevalence of rotavirus did not differ significantly between community based controls (7.7%, 24/310) and hospital based controls (7.7%, 18/235, P>0.05).

**G-genotypes:** Sequencing results detected rotavirus G8 in 33 cases, which were misallocated to G12 when G genotype was assessed by PCR alone. Rotavirus G-typing based on sequencing results of 211 (190 cases and 21 controls) samples which were successfully sequenced, showed that genotype G1 (68.6%), was most common followed by G8 (17.3%), G12 (11.5%) and G4 (2.1%) in cases. Rotavirus G1 also predominated in controls and accounted for 66.7%, followed by G12 (19.0%) and G8 (14.3%). Genotype G2, G3 and G10 were not detected. The Tanzanian G1 variants displayed 94% similarity with the Rotarix vaccine G1 variant.

**P-genotypes:** Rotavirus P genotypes were detected in 236 samples (211 cases and 25 controls) using RT-PCR. The commonest P genotypes were P[8] 66.8%, P[4] 19.0%, and P[6] 14.2%. We did not detect rotavirus P[9], P[10]and P[11]. Rotavirus G/P combinations were found in 211 samples that were successfully typed for both G and P genotypes (among the 236 P-typed samples, 25 could not be G-typed). The commonest G/P combination in cases was G1P[8] accounting for 123 samples, followed by G8P[4] (n=27) and G12P[6] (n=21).
Rotavirus prevalence was significantly higher in cool (23.9%) than hot months (17.1%) of the year (P=0.012). Rotavirus was most frequently found in the age group of four to six months. The prevalence of rotavirus in cases was lower in stunted children (28.9%) than in non-stunted children (40.1%, P=0.003) and significantly lower in HIV-infected (15.4%, 4/26) than in HIV-uninfected children (55.3%, 42/76, P<0.001).

5.5 Paper II

Genetic diversity of norovirus in hospitalised diarrhoeic children and asymptomatic controls in Dar es Salaam, Tanzania


This study included 705 cases and 561 controls. Norovirus was detected in 14.3%, 181/1266 children. The prevalence of norovirus was significantly higher in cases (18.3%, 129/705) than in controls, (9.2%, 52/561, P < 0.02, OR = 2.2, 95% CI: 1.56–3.09). The prevalence of norovirus did not differ among the two control groups; community based controls (9.3%, 29/312) and hospital based controls (9.2%, 23/249).

Genogroups: Except for one child who had double infection with GI and GII all 129 cases had norovirus GII. Among norovirus positive controls, 23.1% had GI and 76.9% had GII. The proportion of norovirus GI was significantly higher in CBC 34.5%, (10/29) than in HBC 8.7%, (2/23), whereas GII was significantly higher in HBC 91.3% (21/23) than in CBC (65.5%, 19/29, P = 0.028, OR = 5.5, 95% CI: 1.1–28.5).

Genotypes: For cases, 108 GII samples were genotyped and for controls 23 and four samples were GII and GI genotyped respectively. Norovirus GII.4 predominated, both in cases and controls, but it was significantly more prevalent in cases than in controls (87.9% vs. 56.5%, P = 0.001).

Other genotypes detected among cases were GII.16 (4.6%), GII.21 (3.7%), GII.g (2.8%) and GII.13 (0.9%). Apart from GII.4, other genotypes detected in controls were GII.21 (26.1%), GII.g (8.7%) and one sample of each of GII.16 and GII.e
The prevalence of norovirus was significantly higher during the first than the second year of life (109/540, 20.2% vs. 20/165, 12.1% \( P = 0.02 \), OR 1.81, 95% CI: 1.08–3.02). In controls, there was no significant difference in the prevalence of norovirus in children below and above one year.

The highest numbers of norovirus were detected in the month of April 2011 which is the month of long rains and the beginning of cold season. However, we did not find significant differences in number of norovirus detected between dry and rainy season in cases (109/580, 18.8%, vs. 20/125, 16.0%, \( P = 0.46 \)) nor in controls (7.9%, 19/239 vs. 10.2%, 33/323, \( P = 0.36 \)).

The prevalence of norovirus in HIV-positive and negative children was (21.2%, 7/33) and (10.3%, 40/390, \( P = 0.05 \)) respectively, regardless of diarrhoea symptoms. No significant difference in gender, parent’s level of education or nutritional status with norovirus infection was observed within cases or controls.

### 5.6 Paper III

**Detection and molecular characterisation of human adenovirus in diarrhoeic children of Dar es Salaam, Tanzania; a case control study**

Moyo S, Kurt Hanevik, Bjørn Blomberg, Oyvind Kommedal, Svein Arne Nordbø, Samuel Maselle, Nina Langeland. (submitted)

This study included 690 children admitted to hospital due to diarrhoea (cases) and 545 children without diarrhoea (controls). Human adenovirus (HAdV) was detected in 37 children, corresponding to a prevalence of 3.5% (24/690) in diarrhoeic and 2.4% (13/545) in non-diarrhoeic children (\( P > 0.05 \)). There was no significant difference in the prevalence’s of HAdV in the two control groups CBC (3.2%, 10/310) and HBC (1.3%, 3/235, \( P = 0.14 \)).
**HAdV types**: The proportions of enteric adenoviruses (type 40 and 41) were not significantly different in diarrhoeic and non-diarrhoeic children (50%, 12/24 vs. 46%, 6/13, P=0.82).

Other HAdV types detected were; 1, 2, 7, 18, 19 and 31.

Among HAdV-infected children, the median age was significantly lower in diarrhoeic than in non-diarrhoeic children (10 vs. 14 months, p<0.001). More than half of HAdV infected (54.2%) were dehydrated as compared to diarrhoeic children without HAdV (45.8%, P=0.01).

The prevalence of adenovirus was not significantly different between rainy and dry seasons. HAdV was not detected in the 33 known HIV positive children. There was no significant association between HAdV infection and gender, nutritional status of the child and parent educational level.

### 6 Discussion

#### 6.1 Discussions of the main findings

##### 6.1.1 Prevalence of enteric viruses

Diarrhoea is one of the most common diseases and a significant cause of morbidity and mortality worldwide. Viruses are recognized as important causes of diarrhoea especially in children during their first years of life. A number of viral agents are associated with diarrheal disease in children with three of these being most common, namely; rotavirus, norovirus and human adenovirus. This study was conducted in order to better understand the clinical and molecular epidemiology of enteric viruses causing diarrhoea in young children. The present study used real time PCR and ELISA to detect the viruses causing diarrhoea in children. One or more viruses were detected in nearly half of the samples tested, 48.8% among children admitted to hospital due to diarrhoea. This is higher compared to our previous study which found at least one virus in 32% of children. Prevalences reported in other studies also differ with regard to presence of viruses, ranging from 30% to 67% in Tunisia, Venezuela, Turkey, southeast China, and China [161, 213, 245-247]. Several factors may explain
the observed discrepancies in reported prevalences of enteric viruses. One is the
detection methods used, some methods are more sensitive than others, such as PCR vs
ELISA for norovirus [248-250]. The second reason is different seasons when the
studies were conducted. Some viruses such as rotavirus shows peak prevalence in
certain months of the year, therefore studies conducted during the peak season report
higher prevalences [247] compared to studies conducted during non-peak season
[195]. The other reasons are the number of enteric viruses detected. If the study has
included the most common viruses it is likely to find higher prevalence of viruses than
studies which have omitted some very common viruses. For example a study
conducted in Tunisia [213] did not search for norovirus, whereas norovirus is known
to be the second common to rotavirus in causing diarrhoea[251]. Different age groups
or different inclusion criteria of the study population included and geographical
variations may also be reasons for the observed variations. Of the three viruses
detected in the present studies, rotavirus was the most frequent followed by norovirus
and adenovirus. Children admitted with diarrhoea were six times more likely to be
infected with rotavirus than those without diarrhoea. The study confirms findings from
other studies twenty years ago in the same location and elsewhere [252, 253]
indicating that rotavirus is still a major pathogen causing diarrhoea in children in
Tanzania. The presence of rotavirus among controls may represent reservoirs for
transmission in the community. Norovirus was the second common virus detected,
which is similar to our previous study [195]. Despite using the same method for
detection of norovirus i.e. RT-PCR, there is much variation in the reported prevalence
in African countries ranging from 6.5% to 25.5% in Malawi, Tunisia, South Africa,
Ghana, Ethiopia, Libya, Botswana and Nigeria [75, 254-259]. This could be due to
different age groups included in these studies or geographical variation. The
prevalence of norovirus in children without diarrhoea, both community controls and
hospital controls, was significantly lower compared to children with diarrhoea as
reported in most other studies [260-264]. The findings from these studies, and results
from the present study, confirm the role of norovirus in causation of diarrhoeal
disease, although in some countries the prevalence of norovirus infection has been
found to be similar in children with and without diarrhoea [258, 265]. Human
adenovirus was the least frequent virus detected in our study. When we compared adenovirus prevalence in this study with reports from other African countries, we observe large variations. Studies detecting human adenovirus by PCR, which is known to be more sensitive, have reported higher prevalences [165, 166, 209] compared to studies employing EIA [212, 213]. The finding of an almost similar prevalence of human adenovirus in diarrhoeic and non-diarrhoeic children, but at higher median age of infection in non-diarrhoeic children, suggest prolonged shedding of adenovirus in stool after previous infection more than one month prior to the study. Alternatively, it could be asymptomatic adenovirus infections in children who may have acquired immunity from previous infections.

6.1.2 Mixed viral infections
In this study we observed a relatively high percentage (7.4%) of dual viral infections among positive samples in cases, most were a combination of rotavirus and norovirus (18/26), and followed by norovirus and adenovirus (6/26) and two were rotavirus and adenovirus. The prevalence of dual infection was not statistically different from what we reported in the previous study (8.0%) in the same study setting [195]. Other studies have reported almost similar rate of dual viral infections in children [213, 247, 266]. Whether a single virus is the main reason for the diarrheal illness or whether they potentiate each other still remains unclear. Mixed infections with viruses, bacteria and/or parasites may also occur, but was not investigated in this study.

6.1.3 Distribution of enteric viruses by age groups
We observed the peak age of viral infection in infants aged 7-12 months in cases (52.0%). This could be due to decreased protection from maternal antibodies which starts to decline at this age. The relatively high prevalence of enteric viruses in infants less than 6 months in cases may suggest lack of complete protection from maternal antibodies. There was a decline of viral infection in children aged above 13 months. The low prevalence of viruses in these older children could be partly due to immunity acquired through previous exposures. The mean age of rotavirus infection in this study compares with the mean age of rotavirus diarrhoea in studies from other developing
countries, which ranges from 6–9 months [267]. The median age of symptomatic norovirus infection was 9 months, which is in agreement with results from Libya [255]. The high prevalence of norovirus infection in children below 1 year has also been documented in many other developing countries, suggesting that children tend to become infected at a young age [255, 258, 268, 269]. For adenovirus, the median age was significantly higher in adenovirus-infected children than adenovirus non-infected children in our study (12 vs. 10 months).

6.1.4 Molecular epidemiology of viral infections
This is the first report to present comprehensive molecular diversity for three important viruses causing diarrhoea in children in Tanzania. PCR and sequencing were used to characterise the viruses in this study. We found wide genetic diversity of all the three viruses examined.

**Rotavirus**
Among rotavirus, we detected five G and three P genotypes. The most common G/P combination was G1[P8]. This genotype combination is responsible for 50–65% of rotavirus infections in children worldwide [270]. In this study we found more than 60% of the study participants were affected by G1[P8] which is the genotype combination included in the current vaccine introduced in Tanzania (Rotarix). We therefore assume that the vaccine will be protective, given that the circulating genotype is stable. The second common genotype detected was G8. We report for the first time the presence of rotavirus genotype G8 and G12 in Tanzania. G12 genotype was previously documented only in three African countries, namely Malawi, South Africa and Nigeria where it was detected by using sequencing [73-75]. Rotavirus genotype G9 was the commonest genotype in our previous study which was conducted in Dar es Salaam eight years earlier [195]. However, in the present study only one strain of G9 in combination with P[8] was found. This emphasizes the need for continuous rotavirus strain surveillance.

Through phylogenetic analysis, the results of this study showed that the rotavirus G1P[8] strains from the studied population are distantly related to G1P[8] of the
vaccine strains in the Rotarix and RotaTeq vaccines. This suggests that circulating G1P[8] strains may have changed over time through accumulated mutations making them different from original vaccine strains which were isolated over twenty-six years ago [271, 272]. In this study we observed an increase in the prevalence of P[6] and P[4] rotavirus positive samples compared to the previous study in the same study setting [195]. Our results showed that genotype P[6] in this study is closely related to P[6] from other African countries such as Malawi and South Africa. Lower vaccine efficacy has been reported in developing countries [273]. This may partly be due high prevalence of circulating genotypes which are not included in the vaccine such as G8, G12, P[4] and P[6]. However, other possible contributing factors such as maternal antibodies, chronic enteropathy, and malnutrition need to be investigated [274, 275].

**Norovirus**

Among detected noroviruses, seven different norovirus genotypes were found. The presence of norovirus diversity has also been observed in other African countries such as South Africa, Tunisia and Malawi which shows that several different genotypes can co-exist at the same time [256, 258, 259]. The phylogenetic analysis of the norovirus, based on both RdRp and capsid region in the current study, showed predominance of norovirus GII.4 genotype, this is in agreement with previous studies conducted in the African region [134, 255, 256, 258, 259, 276-278]. This genotype has been circulating worldwide from the mid-1990s to recent years [279-286], confirming that it plays a major role in causing diarrhoea in children. Throughout the study period all detected norovirus GII.4 clustered with strain New Orleans 2009. We are reporting presence of this strain in Tanzania 2 years after it was first detected in winter 2009 in the USA [287] where it caused outbreaks. Since its emergence, the New Orleans strain has spread widely and has been reported in other countries such as Finland, Brazil and Vietnam etc.[280, 288, 289]. This is the first report of norovirus genotype GII.21 in an African region. The genotype was detected in children with and without diarrhoea. Norovirus GII.21 was first detected in humans in 2002 in Japan from patients with gastroenteritis and was reported as norovirus GII.16 [226, 290]. This genotype has also been reported recently in Brazil [291].
Adenovirus
Sequence analysis showed that a wide variety of human adenovirus species (five) and types (nine) circulate among diarrhoeic and non-diarrhoeic children in Dar es Salaam. Half of the adenovirus types were enteric adenovirus 40 and 41 in both diarrhoeic and non-diarrhoeic children. In the current cross sectional study, adenovirus type 40 and 41 occurred at equal frequency in diarrhoeic children. This observation is consistent with what has been reported in other studies [292, 293]. In non-diarrhoeic children, adenovirus type 40 was more prevalent than type 41. Some studies have reported antigenic drift of adenovirus type 41, whereby an increase of adenovirus type 41 at the expense of adenovirus type 40 has been observed. In order to detect serotype drift in the study setting, future studies are needed over a prolonged period of time, as reported elsewhere [36, 294, 295]. Human adenovirus types 1, 2, 3, and 7 which are associated with respiratory infections [173, 296], are also shown to be associated with diarrhoeal disease [179]. In the present study these adenovirus types were detected from both diarrhoeic and non-diarrhoeic children. When these HAdV types infect respiratory sites, they can be shed in the faeces of an infected person for months [173, 296]. Hence findings of this study partly support the theory of prolonged shedding of these human adenovirus species in faeces. However, we cannot rule out their role in diarrhoea etiology because these types have also been reported to cause diarrhoea [179].

6.1.2 Seasonality of viral infections

Seasonality is an important feature of viral diarrhoea; most studies have demonstrated that viral diarrhoea peak during winter season. In paper I, we observed a high prevalence of rotavirus infection during the cooler months of the year. This is in agreement with studies conducted during cool months of the year [253, 297] which found high prevalences of rotavirus infection compared to studies conducted during hot months of the year [195, 298] in Dar es Salaam, Tanzania. In paper II, we observed a high number of norovirus infections during the month of April, which is regarded as a month of long rains and beginning of the cool months of the year in our
study setting. Therefore this finding partly supports winter month seasonality of norovirus associated diarrhoea, as documented by several studies in the past [247, 299, 300]. However, there were no significant differences in number of norovirus positive children during the rainy season compared to the dry season in this study. In paper III, human adenoviruses were found to be more prevalent during the short rain months of the study period, but not for the two rainy seasons combined. However, other studies have reported no specific seasonal pattern of adenovirus infection [36, 37]. Proper understanding of seasonality of viral infections may provide not only some insights into the possible modes of transmission of viruses but also is useful when considering the appropriate timing of immunization booster programs. This will be applicable in regions which have reported low efficacy of rotavirus vaccine and have demonstrated strong seasonality. Rotavirus vaccine is administered at the age of 6, 10 and 14 weeks in developing countries. However, if booster vaccination programs were to be considered for older children lacking immunity, vaccination during the pre-rotavirus season has been recommended [30].

### 6.1.3 Association between enteric viruses and HIV

Infection with HIV is common among children of sub-Saharan Africa countries, and diarrhoeal disease is a leading cause of morbidity and mortality in HIV-infected children in these areas [22, 23]. In paper I, we found significantly lower incidence of rotavirus in HIV infected children compared to HIV negative among cases. This is in agreement with findings of previous studies in Tanzania and other developing countries where rotavirus is not found at all or found at low prevalence among HIV-positive children with diarrhea [22, 234, 301-304]. In paper II, we found a higher prevalence of norovirus infection in HIV infected children than HIV negative children, but the association was not statistically significant. Our finding is different from a previous study in the same study setting [301] which reported a significant association of HIV infection and norovirus. The lack of significant association in the current study can partly be explained by the inclusion of acute/persistent diarrhoea, while the previous study included only children with chronic diarrhoea which is
reported to be common in HIV infected individuals [146]. In paper III, adenovirus was not detected in HIV infected children, which is in agreement with studies conducted in the neighbouring country Kenya and elsewhere [166, 302]. These studies suggest that adenovirus is not an important cause of diarrhoea in young HIV positive children. Understanding the causative agents of diarrhoea in HIV infected children needs further investigation.
6.2 Methodological discussion

6.2.1 Study design, study population and study participants

Although this was a case control study, it was not matched by age and sex, and hence may not an ideal study design. However, the large samples size of our study enabled us to do stratification of data by age and sex during analysis to take care of confounding, and multivariate analysis was also used for paper I. The source population used for cases was from secondary and tertiary hospitals and as a result the study used a specialized study population. This could be a limiting factor because we dealt with only severe cases admitted in these hospitals. However, this study was designed for analysis of the viruses in hospitalized diarrhoeic children, these are severe cases of gastroenteritis and hence make the source population appropriate for this purpose.

The two control groups used in this study were from child health clinics or children admitted due to diseases other than diarrhoea. Although the prevalence of viruses was significantly lower in these controls compared to cases, the control groups may not be appropriate, because they may not represent the same source population from which cases arose. An alternative way was to obtain controls from a well-defined source population where cases reside and make the study a population based study.

6.2.2 Laboratory methods

Extraction of nucleic acids is one of the crucial stages to get optimal PCR results [145]. Extracted RNA or DNA should be free from inhibitors of the PCR reaction. To avoid getting a false-negative result because of interference from inhibitors, an internal control must be incorporated in every molecular assay [199, 305]. In the present study MagNA Pure LC 2.0 (Roche) with total nucleic acid extraction kit was used. The method may not be able to remove all inhibitory substances present in stool. This insufficient removal of stool inhibitors is also expected in other extraction methods available [199]. We did not include internal controls in our PCR methods;
hence our results could have been affected by false negatives from stool inhibitors. This was especially seen when typing rotavirus, not all rotavirus ELISA positive samples were PCR positive. However all adenovirus ELISA positive samples were positive in PCR analysis during typing. If indeed stool inhibitors were present in the extracted nucleic acids, this may have caused underestimation of norovirus prevalence which was detected using direct RT-PCR. The study could thus have benefited from inclusion of internal controls in each PCR run.

ELISA was used for detection of rotavirus and adenovirus. Studies have indicated that ELISA is cheaper, easier to perform and relatively cheap compared to PCR [207, 306, 307]. At present, ELISA is most widely used for detection of these viruses, and we found prevalences of these viruses comparable to other studies in developing countries [212, 213, 253, 298, 308, 309]. Although PCR is reported to be more sensitive compared to ELISA for detection of these viruses, it may not be relevant for epidemiological studies, as well as routine diagnostics in developing countries, because of shedding of large amounts of viral particles in stool.

Semi nested RT-PCR is widely used for typing of rotaviruses both in developing and developed countries. Mistyping of rotavirus using PCR can result, due to constant accumulation of point mutations through genetic drift; emergence of novel genotypes; and possibly zoonotic transmission and subsequent assortment. In paper I we observed mistyping of rotavirus using RT-PCR, this was also reported by other studies [75, 310-312]. The strength of this paper was that sequencing was also used to confirm PCR

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Binding site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer_G8-F</td>
<td>5'-TTTRCGCACATTGGTGAAT-3'</td>
</tr>
<tr>
<td>Our G8 variant A</td>
<td>5'-TGTCACACCATTGGTAACT-3'</td>
</tr>
<tr>
<td>Our G8 variant B</td>
<td>5'-TGTCATACCATTGGTAACT-3'</td>
</tr>
<tr>
<td>Primer_G12-F</td>
<td>5'-GGTTATGTAATCCGATGGACG-3'</td>
</tr>
<tr>
<td>Our G8 variant A+B</td>
<td>5'-GGTTATCGAATCCGATGGACG-3'</td>
</tr>
</tbody>
</table>

**Figure 7: Rotavirus G8 primers used and mismatching on during sequencing**

Red underscored letters indicate positions with a mismatch against the above primer. Without performing sequencing, we could not only have reported wrong molecular epidemiology of rotavirus but also underestimated the presence of rare emerging genotypes such as G8 in Tanzania.
typing results. In paper I, we first used semi nested multiplex RT-PCR for G and P typing of all ELISA positive rotavirus samples. But sequencing of 43 randomly selected samples revealed that eight rotavirus G8 samples were mistyped as rotavirus G12. This necessitated re-analysing all G types by using sequencing. We then found that 33 G8 samples were mistyped as G12 using RT-PCR. Figure 7 below shows mismatch of primes of G8 used for RT-PCR. In the present study we observed two variants of rotavirus G8 A and B through sequencing, variant A was the most common G8 sequence variant. The Tanzanian rotavirus G8 variants displayed 3-4 mismatches against the G8-primer whereof 3 mismatches are located in the 3’-end portion. The 3’-terminal mismatch will be removed by the proof-reading polymerase leaving only 2 effective mismatches in the mid-part of the rotavirus G12-primer. This unintended binding site lies close to the expected binding-site on true rotavirus G12 isolates producing amplicons indistinguishable from true G12 amplicons. Without performing sequencing, we could not only have reported wrong molecular epidemiology of rotavirus but also underestimated the presence of rare emerging genotypes such as G8 in Tanzania. Figure 8 shows molecular epidemiology of rotavirus in this study using PCR and sequencing results. A different and a wrong molecular epidemiology could have been reported in this study if sequencing was not used.
7 Study limitations

This study dealt with only three viruses known to be common in the setting. Other viruses such as sapoviruses, astroviruses and coronaviruses were not investigated. This may have underestimated the burden of viral diarrhoea in children in the study setting. In order to fully understand the aetiology of viral diarrhoea, a study which will search for more viruses is needed.

Acute or chronic diarrhoea can be caused by a combination of viruses, bacteria or parasites. As this study dealt with viruses only, it is difficult to ascertain if diarrhoea was caused by viruses or not. This is especially important when trying to associate a causative agent with clinical features such as dehydration and the type of diarrhoea, acute vs. persistent diarrhoea.

There were missing data in some months during the study period. This could have affected the assessment of seasonality of viral infections. This was especially observed in paper III were a clear seasonality was not found. Despite this limitation of missing data in some months during the study, in paper I we were able to show a clear pattern of seasonality for rotavirus.

This study was conducted in one region of Tanzania only, Dar es Salaam. In order to better understand the strain distribution of rotavirus which has been reported to show geographical variations even within the same country [313], a comprehensive study involving other regions of Tanzania may be required. This is especially important for viruses for which vaccines are available such as rotavirus.

Data on breast feeding is missing, this could have given additional information on how breast feeding protect against viral infections.

HIV testing results was not available for all study participants. A relatively small number of children with diarrhoea were tested for HIV. The small sample size for tested HIV participants lowered the power of the study.

Unsuccessful sequencing of a number of positive samples for both rotavirus and norovirus may have underestimated presence of other genotypes which could have been detected if sequencing was successful.
8 Conclusions

Based on the objectives of this study, the following conclusions were reached:

1. Viruses contribute a high burden of diarrhoea in hospitalised children of Dar es Salaam Tanzania. Children aged 7-12 months were especially affected.
2. There is high asymptomatic carriage of diarrhoea-associated viruses in children (paper I, II and III), which may be explained by prolonged shedding of the virus from previous infection, or new viral infection without symptoms because of low infective dose. These children may serve as a source of infection in the community.
3. Pre-rotavirus vaccination data in Tanzania (Paper I) showed predominance of rotavirus genotype G1, which is phylogenetically distantly related to the vaccine strains. We also observed the emergence of rotavirus genotype G8 and G12.
4. Genetic analysis of norovirus causing infection in Tanzania (Paper II) showed predominance of norovirus GII.4 both in diarrhoeic cases and among controls.
5. We showed high genetic diversity of human adenovirus types that circulate in the study setting, with equal proportion of enteric types and non-enteric types in cases and controls.
6. Rotavirus seasonality, as seen in Paper I, provides insights important for vaccination strategies, including potential shifts in seasonal peaks and duration of outbreaks. No clear pattern of seasonality was seen in norovirus infection (Paper II) and adenovirus infection (Paper III).
7. Our data suggests that rotavirus may not be an opportunistic pathogen in children infected with HIV (Paper I). We did not see an association between HIV and norovirus infection (Paper II) or adenovirus infection (Paper III).

9 Recommendations

Due to our findings of mistyping of rotavirus when using PCR to identify strains, due to higher mutation-rates in viral genomes, we recommend to interpret rotavirus typing results with caution when G typing is based on multiplex PCR. Furthermore, our
results emphasize the higher robustness for typing of rotaviruses obtained by sequencing.

This study and previous ones have shown that norovirus is second to rotavirus in causing diarrhoea in children hospitalised for diarrhoeal disease. There is no licenced vaccine for norovirus, the efforts to reduce disease burden have focused on effective disease surveillance and limiting disease transmission. There is no routine diagnosis of norovirus in our study setting, without which diagnosis and detection of norovirus in diarrhoeic children is difficult. This will hinder detection of an outbreak due to norovirus, in which the virus has been reported to have severe outcomes and death [122, 314]. We therefore recommend routine testing of norovirus to be established in the study setting for proper prevention and control measures.

10 Future perspectives

Diarrhoea continues to be a major health problem in children worldwide. Detection of enteric viruses using proper diagnostic methods will add important knowledge on the contribution of each virus in diarrhoea and hence proper preventive measures to be put in place. Due to the possibility of change in pattern of circulating serotypes/genotypes of viruses causing diarrhoea, it will be important to continue rotavirus surveillance in the study setting. This phenomenon was observed in paper I when compared to our previous study [195]. Rotavirus vaccination was started in Tanzania in 2013, and it will be important to conduct studies post rotavirus vaccination period, to evaluate the efficacy of the vaccine or if there is reduced incidence of hospitalization due to rotavirus diarrhoea or if this is overtaken by other viruses such as norovirus. Further studies need to be conducted for longer periods of time, more than a year with equal number of samples every month, in order to better understand seasonality of these viruses, as well as to detect any shift of serotypes in virus infection such as adenovirus. Secondly, it will help to confirm consistency of the seasonal trends observed in this study.

Although this study used control groups, these were not necessarily children living in the same community as cases, future studies need to be done to include controls from
the community population in order to differentiate new infection or reinfection of the virus from prolonged carriage of the virus. Adding analysis of bacterial, as well as parasitic pathogens will be useful to give a clearer picture of the role mixed infections in diarrhoea. Further studies which will include a higher number of HIV infected participants is required to understand the burden of norovirus and adenovirus in HIV infected children.
References


APPENDIX I

QUESTIONNAIRE FOR CASES

MUHIMBILI UNIVERSITY OF HEALTH AND ALLIED SCIENCES (MUHAS)

Title: Rotavirus infection in hospitalized children with diarrhoea and asymptomatic children below five years at Muhimbili National Hospital, Dar es Salaam Tanzania

1. Child ID Number
2. Name and initials of the child ...........................................
3. Informant (a) Parent (mother / father) □ (b) Care taker □
4. Residential area/ place of stay ...........................................
5. Which district are you coming from?
   a) Ilala □
   b) Temeke □
   c) Kinondoni □
6. What is the parent/guardian’s highest level of education?
   a) Informal □
   b) Primary □
   c) Secondary □
   d) Post-secondary education □

Particulars of the child

7. Sex (a) male □  (b) Female □
8. Age ..............................
9. Date of birth (from MCH card) .................................
10. Weight (in kilograms) ....................... ...........................
11. Height (in centimetres) .....................
12. Mid Upper Arm Circumference (MUAC) (in centimetres) ...........

Symptoms and signs

13. What is the duration of Diarrhoea?
   (a) ≤ 13 days □
   (b) ≥ 14 days □
14. Is the stool blood stained?
   (a) Yes ☐ (b) No ☐

15. Has the child been treated with antibiotics the last 2 weeks?
   (a) Yes ☐ (b) No ☐

16. If so when/why/ name of drug?
   .................................................................

17. Hydration status of the child
   a) no dehydration
   b) some dehydration
   c) severe dehydration

   **Breast feeding practice**

18. Is the child on exclusive breast feeding (if below six months)
   (a) Yes ☐ (b) no ☐

19. Is the child still breast feeding (if the child is above (six- twelve months)
   (a) Yes ☐ (b) no ☐
APPENDIX II

QUESTIONNAIRE FOR CONTROLS

MUHIMBILI UNIVERSITY OF HEALTH AND ALLIED SCIENCES (MUHAS)

Title: Rotavirus infection in hospitalized children with diarrhoea and asymptomatic children below five years at Muhimbili National Hospital, Dar es Salaam Tanzania

1. Child ID Number □□□□□□
2. Name and initials of the child ...........................................
3. Informant (a) Parent (mother / father) ..................................
   (b) Care taker □
4. Residential area/ place of stay ........................................
5. Which district are you coming from?
   a) Ilala □
   b) Temeke □
   c) Kinondoni □
6. What is the parent/guardian’s highest level of education?
   a) Informal □
   b) Primary □
   c) Secondary □
   d) Post secondary □

Particulars of the child
7. Sex (a) male □ (b) Female □
8. Age ............................
9. Date of birth (from MCH card) .........................
10. Weight (in kilograms) .........................
11. Height (in centimetres) .........................
12. Mid Upper Arm Circumference (MUAC) (in centimetres) ............
13. Has the child been treated with antibiotics the last 2 weeks?
   (a) Yes □ (b) No □
14. If so when/why/ name of drug?
   ..........................................................
Breast feeding practice

15. Is the child on exclusive breast feeding (if below six months)
   (a) Yes ☐  (b) no ☐

16. Is the child still breast feeding (if the child is above six- twelve months)
   (a) Yes ☐  (b) no ☐