

1 **Epidemiology and aetiology of moderate to severe diarrhoea in hospitalised HIV-infected patients**
2 **≥5 years old in South Africa, 2018-2021: a case-control analysis**

3 **Authors and affiliations**

4 Siobhan L. Johnstone^{1,2}, Linda Erasmus¹, Juno Thomas¹, Michelle J. Groome^{3,4}, Nicolette M. du
5 Plessis⁵, Theunis Avenant⁵, Maryke de Villiers⁶, Nicola A. Page^{1,7}

6 ¹Center for Enteric Diseases, National Institute for Communicable Diseases, a division of the National
7 Health Laboratory Service (NHLS), Johannesburg, South Africa; ²School of Public Health, Faculty of
8 Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; ³Division of Public
9 Health Surveillance and Response, National Institute for Communicable Diseases, a division of the
10 National Health Laboratory Service (NHLS), Johannesburg, South Africa; ⁴School of Pathology,
11 Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa;
12 ⁵Department of Paediatrics, Kalafong Provincial Tertiary Hospital, Faculty of Health Sciences,
13 University of Pretoria, Pretoria 0001, South Africa; ⁶Department of Internal Medicine, Kalafong
14 Provincial Tertiary Hospital, Faculty of Health Sciences, University of Pretoria, Pretoria 0001, South
15 Africa; ⁷Department of Medical Virology, Faculty of Health Sciences, University of Pretoria, Private
16 Bag X323, Arcadia, 0007;.

17 Corresponding author: Siobhan L. Johnstone; siobhanj@nicd.ac.za; siobhanjsa@gmail.com; (+27)
18 73 332 7905

19 Short title: Aetiology of diarrhoea in hospitalised HIV-infected patients

20 Key words: Diarrhoea, diagnostics, aetiology, HIV, South Africa

21 Authors' contributions

22 S.L.J, N.A.P, M.J.G, L.E and J.T conceptualised and designed the study. S.L.J, N.A.P, N.D., T.A. and
23 M.dV completed and managed fieldwork and data collection. S.L.J, N.A.P, M.J.G, L.E and J.T acquired,
24 analysed and interpreted the data. S.L.J drafted the work. All authors reviewed the manuscript.

25 Abstract

26 Diarrhoea is a recognised complication of HIV-infection, yet there are limited local aetiological data
27 in this high-risk group. These data are important for informing public health interventions and
28 updating diagnostic and treatment guidelines. This study aimed to determine the pathogenic causes
29 for diarrhoeal admissions in HIV-infected patients compared to hospital controls between July 2018
30 and November 2021.

31 Admitted diarrhoeal cases (n=243) and non-diarrhoeal hospital controls (n=101) ≥5 years of age
32 were enrolled at Kalafong, Mapulaneng and Matikwana hospitals. Stool specimens/rectal swabs
33 were collected and pathogen screening performed on multiple platforms. Differences in pathogen
34 detections between cases and controls, stratified by HIV status, were investigated.

35 The majority (n=164, 67.5%) of diarrhoeal cases with known HIV status were HIV-infected.

36 Pathogens could be detected in 66.3% (n=228) of specimens, with significantly higher detection in
37 cases compared to controls (72.8% versus 50.5%, $p<0.001$). Amongst HIV-infected participants,
38 prevalence of *Cystoisospora* spp. was significantly higher in cases than controls (17.7% versus 0.0%,
39 $p=0.028$), while *Schistosoma* was detected more often in controls than cases (17.4% versus 2.4%,
40 $p=0.009$). Amongst the HIV-uninfected participants, prevalence of *Shigella* spp., *Salmonella* spp. and
41 *Helicobacter pylori* was significantly higher in cases compared to controls (36.7% versus 12.0%,
42 $p=0.002$; 11.4% versus 0.0%, $p=0.012$; 10.1% versus 0.0%, $p=0.023$).

43 Diarrhoeal aetiology differed by HIV status, with *Shigella* spp. (36.7%) and *Salmonella* spp. (11.4%)
44 having the highest prevalence amongst HIV-uninfected cases and *Shigella* spp. (18.3%),
45 *Cystoisospora* (17.7%), and *Cryptosporidium* spp. (15.9%) having the highest prevalence amongst
46 HIV-infected cases. These differences should be considered for the development of diagnostic and
47 treatment guidelines.

48

49 Introduction

50 Diarrhoeal diseases pose a significant health burden across the globe, causing over 1.6 million deaths
51 annually amongst all ages¹. The majority of these deaths occur in low- and middle-income countries
52 (LMICs), specifically in south Asia and Sub-Saharan Africa (SSA)¹. Despite a decrease in global
53 diarrhoeal mortality rates over the past two decades, decreases have not been uniform across age
54 groups or settings². Unsafe water, poor sanitation and hygiene (WaSH)¹, as well as malnutrition and
55 compromised immunity³ are the main risk factors associated with diarrhoeal morbidity and
56 mortality. Other risk factors identified in LMICs include residence in a slum, use of communal toilets
57 and exposure to animals⁴. Children under the age of 5 years are at the highest risk, however there is
58 a significant burden in older age groups, specifically in LMICs^{1,5}. Despite global improvements in
59 WaSH⁶, much of the remaining burden of diarrhoeal disease may be attributed to foodborne
60 pathogens⁷. Food safety is of specific concern in LMICs where there is increased consumption of
61 unsafe foods, use of effluent in agriculture and changes in food distribution networks with bulk
62 production and increased distance between production and consumption as well as large, often
63 poorly regulated, informal sectors⁷.

64 Studies from high-income countries (HICs) have shown the majority of diarrhoeal diseases in adults
65 seeking healthcare to be due to *Campylobacter* spp., *Salmonella* spp., norovirus and rotavirus^{8,9}.
66 Aetiology differs in SSA, with a meta-analysis amongst all ages identifying the main causative
67 pathogens as *Escherichia coli*, *Cryptosporidium*, *Cyclospora*, *Entamoeba* and *Shigella* spp.¹⁰ There are
68 several reasons for these differences in aetiology including different exposures and risk factors
69 associated with poorer living conditions and poor underlying health. One important consideration is
70 the high prevalence of human immunodeficiency virus (HIV) in SSA¹¹. HIV is a known risk factor for
71 diarrhoea, with episodes in HIV-infected individuals more likely to be severe, prolonged and result in
72 hospitalisation with higher mortality rates than in HIV-uninfected individuals^{12,13}. The aetiology of
73 diarrhoea in HIV-infected individuals is known to shift with HIV disease progression and treatment¹⁴.

74 During early disease stages, while CD4+ counts are high and viral loads low, diarrhoea is often
75 related to HIV seroconversion. As disease progresses, diarrhoea is frequently caused by
76 opportunistic pathogens related to decreased immune function¹⁴. These pathogens include bacterial
77 infections, such as *Mycobacteria*, parasitic infections, such as *Cystoisospora belli*, *Cyclospora*,
78 *Strongyloides*, *Cryptosporidia*, *Microsporidia* and viral infections, such as cytomegalovirus (CMV)^{14–16}.
79 Patients with low CD4+ cell counts are at increased risk for chronic diarrhoea⁴ and polyparasitic
80 infection¹⁷. The HIV infection itself may be responsible for a proportion of the pathogen-negative
81 cases, through HIV enteropathy and permanent damage to the gastrointestinal mucosa¹⁸, as well as
82 changes in gut microbial populations leading to dysbiosis¹⁹. Mucosal damage reduces small bowel
83 villous surface area, causing diarrhoea through malabsorption²⁰ as well as defects to cellular and
84 humoral defence mechanisms in the gastrointestinal tract²¹. Pathogen-negative cases associated
85 with advanced HIV disease may also be related to gastrointestinal tract involvement by
86 *Mycobacterium tuberculosis* and/or *avium*²² which is difficult to diagnose. A Dutch study found that
87 34% of diarrhoeal cases in untreated HIV-infected individuals could not be explained by a known
88 causative agent, and hypothesized that these unexplained cases were due to the HIV-infection
89 itself²³. Initiation of antiretroviral treatment (ART), and accompanying increase in CD4+ cell count, is
90 associated with a shift in the aetiology of diarrhoea from opportunistic infectious causes to non-
91 infectious causes²⁴. Data from both HICs and SSA show that although patients initiated on ART have
92 significantly reduced diarrhoeal incidence compared with those not on treatment, the overall
93 diarrhoeal incidence remains high^{18,25}. A proportion of the remaining episodes in patients with well-
94 controlled HIV may be drug-related¹⁸.

95 There are limited aetiological data on HIV-related diarrhoeal diseases in SSA, with the majority of
96 published studies focusing on children under the age of five years, with available data in older
97 children and adults being limited by diagnostic technology¹⁴. A review of laboratory records in
98 Botswana indicated that only 14% of stool specimens submitted for routine diagnostic testing for all
99 ages had a pathogen detected, of which 8% were bacteria and 6% parasites²⁶. *Shigella* spp. and

100 *Salmonella* spp. were the most commonly detected bacteria while *Cystoisospora* spp. and
101 *Cryptosporidium* spp. were the most commonly detected parasites. Viral testing was not done.
102 Although HIV status was not available for the participants in this analysis, Botswana is a high HIV
103 prevalence setting with an estimated 23.9% of the population between 15-49 years old being HIV-
104 infected during the study time period²⁶. Contrary to these findings, a longitudinal cohort study of
105 Zambian adults with a 31% HIV seroprevalence (higher than the estimated population HIV
106 seroprevalence of 22%) found that pathogens could be detected in 99% of stool samples collected
107 (diarrhoeal and non-diarrhoeal patients combined), with *Cryptosporidium* spp., *Cystoisospora* spp.
108 and *Citrobacter* spp. being significantly more common in HIV-infected than HIV-uninfected
109 individuals⁵. They found that HIV-infected adults in the pre-ART era were 2.4 times as likely to suffer
110 from diarrhoea than HIV-uninfected adults and that this increased risk lasted throughout the
111 infection rather than being limited to those with low CD4+ cell counts⁵. Meta-analysis data indicates
112 a high burden of *Cryptosporidium*, microsporidia and *Cystoisospora* in SSA specifically¹⁶. Another
113 recent meta-analysis identified South Africa as having the highest global *Cryptosporidium* prevalence
114 (57.0%, CI 95%: 24.4-84.5%), although they recognised that these estimates were based on limited
115 data²⁷. A study in rural South Africa found that 60% of patients suffering from chronic diarrhoea
116 were HIV-infected and that the majority of these infections were due to *Campylobacter* spp. (20%),
117 *Plesiomonas shigelloides* (17%), *Aeromonas* spp. (13%), *Shigella* spp. (10%), *Salmonella* spp. (10%)
118 and *E. coli* spp. (10%)²⁸. Testing did not include parasites or viruses and results were not stratified by
119 ART status. Other studies in African children have identified HIV-infection as a significant risk factor
120 for rotavirus infection^{13,29}. Very few studies in SSA in adults included screening for viruses, despite
121 data from HICs showing that viruses (specifically norovirus and rotavirus) were responsible for as
122 much as 44% of diarrhoea in adults presenting to emergency departments⁹. Patients with advanced
123 HIV are at increased risk for diarrhoea due to CMV infection³⁰. As CMV is mainly diagnosed on
124 endoscopic biopsy³¹, prevalence is likely underestimated.

125 South Africa currently has the highest HIV burden in the world, representing 19% of the global HIV-
126 infected population³². It is estimated that 86% of HIV-infected adults in South Africa have been
127 diagnosed, with 57% of these on ART (50.8% - 72.7% by province)³³. Women, elderly patients, those
128 with inadequate access to WaSH facilities and those not yet initiated on ART, have the highest risk
129 for developing diarrhoea amongst HIV-infected South Africans³⁴. The majority of data on HIV-related
130 diarrhoea comes from HICs with low HIV prevalence¹⁴. Since diarrhoeal aetiology is likely to vary
131 between settings³⁵, there is a need for more targeted epidemiological studies¹⁸, specifically in LMICs
132 with high HIV prevalence. There are few studies in HIV-infected populations which investigate a
133 comprehensive range of pathogens or infections with multiple pathogens¹⁹. These data are
134 important for informing public health interventions and updating diagnostic and treatment
135 guidelines, especially since treatment is often based on guidelines rather than individual patient
136 diagnostics³⁶. This study aimed to determine the pathogenic causes for diarrhoeal admissions in HIV-
137 infected patients compared to controls without diarrhoea at three South African hospitals between
138 July 2018 and November 2021.

139

140 Methods

141 *Case and control enrolment*

142 Patients of all ages hospitalised with diarrhoea were enrolled at Kalafong, Mapulaneng and
143 Matikwana hospitals through the African Network for improved Diagnostics, Epidemiology and
144 Management of common Infectious Agents (ANDEMIA)³⁷. Kalafong Provincial Tertiary Hospital is
145 based on the western outskirts of Pretoria (Gauteng Province) and serves the generally low-income
146 communities residing in urban townships to the west of Pretoria³⁸. The rural sites included
147 Mapulaneng, a 180 bed district hospital, and Matikwana, a 250 bed regional hospital³⁹, both located
148 in Bushbuckridge district in Mpumalanga Province.

149 Cases were defined as patients admitted with diarrhoea (three or more loose or liquid stools over a
150 24 hour period) for any duration. Unmatched controls were defined as individuals presenting to the
151 hospital or clinic for reasons other than diarrhoea (vaccination clinic, orthopaedic or surgical wards),
152 without gastrointestinal symptoms (vomiting or diarrhoea) in the past 3 weeks. The study was
153 explained to patients fulfilling the study definition and an information leaflet provided. Written
154 informed consent was signed by patients ≥ 18 years of age and by parents/guardians for patients < 18
155 years. An assent form was signed for patients between 7-17 years. Surveillance Officers completed
156 investigation forms by interviewing the patient, parent/guardian and reviewing clinical notes and
157 laboratory results where available. HIV status was determined from the clinical notes and available
158 laboratory results. Stool specimens or rectal swabs were collected from both cases and controls.
159 Patients were enrolled within 48 hours of admission to exclude nosocomial infections. Cases were
160 enrolled from July 2018 to June 2021 and controls enrolled from October 2019 to November 2021.
161 Cases and controls unable to provide specimens and patients < 5 years old were excluded from this
162 analysis.

163 *Laboratory testing*

164 Pathogen screening was performed on Fast-track Diagnostics (viral and bacterial gastroenteritis and
165 stool parasite kits), Seegene Allplex (GI-Parasite and GI-Helminth(I) assays) as well as TaqMan Array
166 Cards (TAC) (included pathogens indicated in Table S1). Monoplex PCR was used to determine final
167 outcome where there were discrepant results between testing platforms for a specific target. Due to
168 limited available TAC cards, cases with unknown HIV status were excluded from testing. Controls
169 with unknown HIV status were included in the overall case-control analysis (since there were limited
170 control specimens available) but were not included in the HIV specific case-control analysis.

171 *Data management and statistical analysis*

172 Results for CD4+ cell count were extracted from the National Health Laboratory Service (NHLS)
173 database for HIV-infected cases and were categorised into three classes (< 200 cells/ μ l; 200-500

174 cells/ μ l; >500 cells/ μ l). Only CD4+ cell count results within 12 months of the enrolment were
175 included and where multiple results were available for a single patient, the results closest to the
176 time of enrolment were used. Demographic and socioeconomic characteristics for cases and controls
177 were compared. Clinical presentation for HIV-infected versus HIV-uninfected cases were compared.
178 Pathogen detection amongst case and control specimens were compared, stratified by HIV status.
179 Clinical presentation and pathogen detection for HIV-infected cases was compared between CD4+
180 cell count categories and for ART and cotrimoxazole prophylaxis treatment. X^2 -test and Fisher's exact
181 test were used for categorical variables and t-test for continuous variables. *P-values* <0.05 were
182 considered to be statistically significant. Stata software (version 14) was used for all analyses.

183 *Ethical considerations*

184 The ANDEMIA study was approved by the Human Research Ethics Committee (Medical) of the
185 University of the Witwatersrand (approval numbers M170403) and the University of Pretoria Faculty
186 of Health Sciences Research Ethics Committee (approval number 101/2017). The HIV sub-analysis
187 was approved by the Human Research Ethics Committee (Medical) of the University of the
188 Witwatersrand (approval numbers M190663).

189 Results

190 *Enrolment and patient characteristics*

191 A total of 378 specimens for patients ≥ 5 years were collected during the study period, including 277
192 cases and 101 controls (Figure 1). Amongst cases, 164 (59.2%) were HIV-infected, 79 (28.5%) were
193 HIV-uninfected and 34 (12.3%) had unknown HIV status. Amongst controls, 23 (22.8%) were HIV-
194 infected, 50 (49.5%) were HIV-uninfected and 28 (27.7%) had unknown HIV status. Cases with
195 unknown HIV status ($n=34$) were excluded (due to limited testing capacity), hence a total of 344
196 specimens were included in the analysis.

197

198 Figure 1: Specimens included in the analysis

199 The median age of included patients was 34 years (IQR of 23-47), with cases being slightly older
 200 (median 36; IQR 26-51) than controls (median 31; IQR 17-38) (Table 1). The majority of patients were
 201 female (59.4%) and from the rural sites (67.4%), with proportionally more controls coming from the
 202 rural site than cases (82.2% versus 61.3%, $p < 0.001$). Cases and controls were similar with respect to
 203 household crowding and dietary habits. Both cases and controls had good access to electric or gas
 204 cookers (86.1%), refrigerators (88.4%) and improved WaSH (94.8% access to improved water source
 205 and 96.4% access to private latrines/flush toilets). CD4+ cell counts were available for 99 (60.4%) of
 206 the 164 HIV-infected cases. Eight (8.1%) had counts > 500 cells/ μ l, 21 (21.2%) were between 200-500
 207 cells/ μ l and the majority ($n=70$, 70.7%) had < 200 cells/ μ l. Information regarding ART was available
 208 for 156 (95.1%) of the HIV-infected cases, of which 130 (83.3%) were on treatment. Mean CD4+ cells
 209 counts were similar between those on ART (229.23 cells/ μ l) and those not on ART (169.7 cells/ μ l),
 210 ($p=0.418$). Information regarding cotrimoxazole treatment was available for 151 HIV-infected
 211 patients, of which 63 (41.7%) were on treatment.

212

213 Table 1: Comparison of patient characteristics for cases and controls

	Total (n=344)	Cases (n=243)	Controls (n=101)	p-value
Age (years) – median (IQR)	34 (23-47)	36 (26-51)	31 (17-38)	<0.001^f
Male	129 (40.6%)	88 (36.2%)	41 (40.6%)	0.445
Rural	232 (67.4%)	149 (61.3%)	83 (82.2%)	<0.001
HIV-infected^a	187/316 (59.2%)	164/243 (67.5%)	23/73 (31.5%)	<0.001
Education^b				0.002^g
None	13/337 (3.9%)	10/237 (4.2%)	3/100 (3.0%)	
<= 6 years	53/337 (15.7%)	47/237 (19.8%)	6/100 (6.0%)	
7-10 years	107/337 (31.8%)	78/237 (32.9%)	29/100 (29.0%)	
>10 years	164/337 (48.7%)	102/237 (43.0%)	62/100 (62.0%)	
Household crowding/people per room – mean (std. dev)	2.0 (1.2)	2.0 (1.9)	2.0 (0.9)	0.733 ^f
Electric or gas cooker	296 (86.1%)	210 (86.4%)	86 (85.1%)	0.757
Refrigerator	289/327 (88.4%)	200/231 (86.6%)	89/96 (92.7%)	0.115

Improved water source^c	326 (94.8%)	229 (94.2%)	97 (96.0%)	0.603 [§]
Treat drinking water^d	45 (13.1%)	31 (12.8%)	14 (13.9%)	0.782
Toilet type				
Private flush	86/337 (25.5%)	62/241 (25.7%)	24/96 (25.0%)	0.604 [§]
Private latrine	239/337 (70.9%)	170/241 (70.5%)	69/96 (71.9%)	
Communal flush	8/337 (2.3%)	7/241 (2.9%)	1/96 (1.0%)	
Communal latrine	2/337 (0.6%)	1/241 (0.4%)	1/96 (1.0%)	
Other	2/337 (0.6%)	1/241 (0.4%)	1/96 (1.0%)	
Dairy^e	156 (45.4%)	107 (44.0%)	49 (48.5%)	0.447
Deli meat^e	103 (29.9%)	79 (32.5%)	24 (23.8%)	0.107
Eggs^e	183 (53.2%)	127 (52.3%)	56 (55.5%)	0.590

214 ^a For participants with known HIV status only; ^b highest level of education that the patient (or parent/caregiver
215 if patient is <15 years) has obtained; ^c piped water to an inside or outside tap (as opposed to river, rain or well
216 water or water from a truck); ^d by boiling or chemical treatment; ^e In the preceding 4 weeks; ^f t-test; [§] fisher's
217 exact test; denominators given when different to column headings due to missing data.

218 *Clinical presentation of cases*

219 HIV-infected cases were likely to experience symptoms for a longer duration before admission
220 (median of 5 days versus 2 days, $p<0.001$) and were more likely to have chronic or persistent
221 diarrhoea (8.5% versus 0.0%, $p=0.006$) than HIV-uninfected cases (Table 2). The most common
222 symptoms experienced by all cases were fatigue (182, 74.9%), weight loss (176, 72.4%), vomiting
223 (172, 70.8%), fever (168, 69.1%), abdominal pain (166, 68.3%) and nausea (165, 67.9%). HIV-infected
224 cases were more likely to suffer from weight loss (81.1% versus 54.4%, $p<0.001$) and nausea (72.0%
225 versus 59.5%, $p=0.051$), while HIV-uninfected cases were more likely to suffer from dysentery (15.2%
226 versus 4.3%, $p=0.003$).

227 Table 2: Clinical presentation of diarrhoeal cases

	Total (n=243)	HIV-infected (n=164)	HIV-uninfected (n=79)	p-value
Duration of symptoms before admission (days) – median (IQR)	4 (2-8)	5 (3-8)	2 (1-4)	<0.001^d
Chronic/persistent diarrhoea^a	13 (5.4%)	13 (7.9%)	0 (0.0%)	0.011^e
Fatigue	182 (74.9%)	124 (75.6%)	58 (73.4%)	0.712
Weight loss	176 (72.4%)	133 (81.1%)	43 (54.4%)	<0.001
Vomiting	172 (70.8%)	118 (72.0%)	54 (68.4%)	0.564
Fever^b	168 (69.1%)	113 (68.9%)	55 (69.6%)	0.910

Abdominal pain	166 (68.3%)	106 (64.6%)	60 (76.0%)	0.076
Nausea	165 (67.9%)	118 (72.0%)	47 (59.5%)	0.051
Respiratory symptoms	105 (43.2%)	77 (47.0%)	28 (35.4%)	0.090
Headache	100 (41.2%)	68 (41.5%)	32 (40.5%)	0.887
Chills	63 (25.9%)	48 (29.3%)	15 (19.0%)	0.087
Arthralgia	51 (21.0%)	38 (23.2%)	13 (16.5%)	0.229
Dermatological symptoms	31 (12.8%)	25 (15.2%)	6 (7.6%)	0.094
Myalgia	24 (10.0%)	16 (9.8%)	8 (10.1%)	0.928
Dysentery^c	19 (7.8%)	7 (4.3%)	12 (15.2%)	0.003
Neurological symptoms	10 (4.1%)	6 (3.7%)	4 (5.1%)	0.732 ^e
Painful lymphadenopathy	4 (1.7%)	3 (1.9%)	1 (1.3%)	>0.99 ^e

228 ^a diarrhoea for 28 days or longer; ^b current temperature of >38°C or history of fever in the past 10 days; ^c
229 Dysentery defined as self-reported blood in the stool; ^d t-test; ^e fisher's exact test

230 *Pathogen detection in specimens of cases and controls*

231 Pathogens were detected in 66.3% (228) of specimens tested, with significantly higher detection in
232 cases compared with controls (72.8% versus 50.5%, $p < 0.001$; Table 3). Bacteria were more prevalent
233 in cases compared with controls (50.2% versus 28.7%, $p < 0.001$), specifically *Shigella* spp., *Salmonella*
234 spp. and *Helicobacter pylori* (24.3% versus 10.9%, $p = 0.005$; 8.6% versus 1.0%, $p = 0.007$ and 8.6%
235 versus 0.0%, $p = 0.001$, respectively). Detection of viruses was significantly higher in cases compared
236 with controls (21.8% versus 10.9%, $p = 0.018$), specifically for adenovirus (11.0% versus 4.0%,
237 $p = 0.038$). Parasites were detected in 37.8% (130) of specimens tested, however, detection was not
238 significantly higher in cases than controls (40.7% versus 30.7%, $p = 0.080$) with the available sample
239 size. Prevalence of *Cystoisospora*, *Cryptosporidium* spp. and *Enterocytozoon* spp. was significantly
240 higher in cases than controls (11.9% versus 0.0%, $p < 0.00$; 11.1% versus 3.0%, $p = 0.012$; 4.5% versus
241 0.0%, $p = 0.038$, respectively). *Blastocystis* and *Schistosoma* were more prevalent amongst controls
242 than amongst cases (23.8% versus 10.3%, $p = 0.001$; 6.9% versus 2.1%, $p = 0.025$, respectively). Cases
243 were more likely than controls to have multiple pathogens detected (37.9% versus 22.8%, $p = 0.007$).
244 Amongst HIV-infected participants, *Cystoisospora* spp. prevalence was significantly higher in cases
245 than in controls (17.7% versus 0.0%, $p = 0.028$), while *Schistosoma* was detected more often in
246 controls compared with cases (17.4% versus 2.4%, $p = 0.009$). Amongst the HIV-uninfected

247 participants, *Shigella* spp., *Salmonella* spp. and *Helicobacter pylori* were more prevalent amongst
248 cases than controls (36.7% versus 12.0%, $p=0.002$; 11.4% versus 0.0%, $p=0.012$; 10.1% versus 0.0%,
249 $p=0.023$). Norovirus GII, *C. difficile*, *Campylobacter*, *Cystoisospora*, *Cryptosporidium* spp. and
250 *Enterocytozoon* spp. were more prevalent amongst HIV-infected than HIV-uninfected participants
251 (6.4% versus 1.6%, $p=0.032$; 7.0% versus 0.8%, $p=0.010$; 6.4% versus 0.8%, $p=0.018$; 15.5% versus
252 0.01%, $p<0.001$; 15.0% versus 1.6%, $p<0.001$; 5.4% versus 0.8%, $p=0.031$). *Shigella* spp. and
253 *Blastocystis* were more prevalent amongst HIV-uninfected than HIV-infected participants (27.1%
254 versus 17.1%, $p=0.032$; 25.6% versus 8.0%, $p<0.001$). *Shigella* spp. (18.3%) and *Cystoisospora* (17.7%)
255 were the most prevalent pathogens detected in HIV-infected diarrhoeal cases, while *Shigella* spp.
256 (36.7%) and *Salmonella* spp. (11.4%) were the most prevalent pathogens in HIV-uninfected
257 diarrhoeal cases. Although *Blastocystis* was prevalent in HIV-uninfected cases, it was not associated
258 with diarrhoea as prevalence was significantly higher in HIV-uninfected controls (40.0% versus
259 16.5%, $p<0.001$).

260 Table 3: Pathogen detection in specimens of cases and controls, stratified by HIV-infection status

	Overall – n ^a (%)				HIV-infected – n ^a (%)				HIV-uninfected – n ^a (%)				p-value ^c
	Total (n=344)	Cases (n=243)	Controls (n=101)	p-value ^b	Total (n=187)	Cases (n=164)	Controls (n=23)	p-value ^b	Total (n=129)	Cases (n=79)	Controls (n=50)	p-value ^b	
Any pathogen	228 (66.3)	177 (72.8)	51 (50.5)	<0.001	138 (73.8)	124 (75.6)	14 (60.9)	0.132	84 (65.1)	53 (67.1)	31 (62.0)	0.555	0.097
Virus	64 (18.6)	53 (21.8)	11 (10.9)	0.018	43 (22.1)	40 (24.4)	3 (13.0)	0.296 ^e	20 (15.5)	13 (16.5)	7 (14.0)	0.707	0.101
Adenovirus	31 (9.0)	27 (11.1)	4 (4.0)	0.038^e	20 (10.7)	19 (11.6)	1 (4.4)	0.476 ^e	10 (7.8)	8 (10.1)	2 (4.0)	0.314 ^e	0.380
Adenovirus 40/41	3 (0.9)	3 (1.2)	0 (0.0)	0.558 ^e	2 (1.1)	2 (1.2)	0 (0.0)	>0.99 ^e	1 (0.8)	1 (1.3)	0 (0.0)	>0.99 ^e	0.637 ^e
Norovirus	18 (5.2)	14 (5.8)	4 (4.0)	0.603 ^e	13 (7.0)	13 (7.9)	0 (0.0)	0.374 ^e	5 (3.9)	1 (1.3)	4 (8.0)	0.074 ^e	0.246
Norovirus GI	4 (1.2)	2 (0.8)	2 (2.0)	0.584 ^e	1 (0.5)	1 (0.6)	0 (0.0)	>0.99 ^e	3 (2.3)	1 (1.3)	2 (4.0)	0.559 ^e	0.308 ^e
Norovirus GII	14 (4.1)	12 (4.9)	2 (2.0)	0.248 ^e	12 (6.4)	12 (7.3)	0 (0.0)	0.367 ^e	2 (1.6)	0 (0.0)	2 (4.0)	0.148 ^e	0.032 ^e
Enterovirus	16 (4.7)	12 (4.9)	4 (4.0)	0.787 ^e	12 (6.4)	10 (6.1)	2 (8.7)	0.645 ^e	4 (3.1)	2 (2.5)	2 (4.0)	0.641 ^e	0.296 ^e
CMV	3 (0.9)	3 (1.2)	0 (0.0)	0.558 ^e	3 (1.6)	3 (1.8)	0 (0.0)	>0.99 ^e	0 (0.0)	0 (0.0)	0 (0.0)	-	0.273 ^e
Astrovirus	5 (1.5)	4 (1.7)	1 (1.0)	>0.99 ^e	3 (1.6)	3 (1.8)	0 (0.0)	>0.99 ^e	2 (1.6)	1 (1.3)	1 (2.0)	>0.99 ^e	>0.99 ^e
Rotavirus	2 (0.6)	2 (0.8)	0 (0.0)	>0.99 ^e	1 (0.5)	1 (0.6)	0 (0.0)	>0.99 ^e	1 (0.8)	1 (1.3)	0 (0.0)	1.000 ^e	>0.99 ^e
Sapovirus	1 (0.3)	1 (0.4)	0 (0.0)	>0.99 ^e	1 (0.5)	1 (0.6)	0 (0.0)	>0.99 ^e	0 (0.0)	0 (0.0)	0 (0.0)	-	>0.99 ^e
>1 virus detected	9 (2.6)	7 (2.9)	2 (2.0)	>0.99 ^e	7 (3.7)	7 (4.3)	0 (0.0)	0.600 ^e	2 (1.6)	0 (0.0)	2 (4.0)	0.148 ^e	0.318 ^e
Bacteria	151 (48.9)	122 (50.2)	29 (28.7)	<0.001	87 (46.5)	76 (46.3)	11 (47.8)	0.894	61 (47.3)	46 (58.2)	15 (30.0)	0.002	0.894
Shigella spp.	70 (20.4)	59 (24.3)	11 (10.9)	0.005	32 (17.1)	30 (18.3)	2 (8.7)	0.378 ^e	35 (27.1)	29 (36.7)	6 (12.0)	0.002	0.032
Salmonella spp.	22 (6.4)	21 (8.6)	1 (1.0)	0.007^e	13 (7.0)	12 (7.3)	1 (4.4)	>0.99 ^e	9 (7.0)	9 (11.4)	0 (0.0)	0.012^e	0.993
C. difficile	14 (4.1)	13 (5.4)	1 (1.0)	0.074 ^e	13 (7.0)	12 (7.3)	1 (4.4)	>0.99 ^e	1 (0.8)	1 (1.3)	0 (0.0)	>0.99 ^e	0.010^e
Campylobacter	13 (3.8)	12 (4.9)	1 (1.0)	0.119 ^e	12 (6.4)	12 (7.3)	0 (0.0)	0.367 ^e	1 (0.8)	0 (0.0)	1 (2.0)	0.388 ^e	0.018^e
STEC	2 (0.6)	0 (0.0)	2 (2.0)	0.086 ^e	0 (0.0)	0 (0.0)	0 (0.0)	-	2 (1.6)	0 (0.0)	2 (4.0)	0.148 ^e	0.166 ^e
ETEC	11 (3.2)	9 (3.7)	2 (2.0)	0.519 ^e	7 (3.7)	5 (3.1)	2 (8.7)	0.207 ^e	4 (3.1)	4 (5.1)	0 (0.0)	0.157 ^e	>0.99 ^e
EPEC	21 (6.1)	14 (5.8)	7 (6.9)	0.631	15 (8.0)	11 (6.7)	4 (17.4)	0.094 ^e	6 (4.7)	3 (3.8)	3 (6.0)	0.676 ^e	0.237
EAEC	41 (11.9)	28 (11.5)	13 (12.9)	0.725	25 (13.4)	20 (12.2)	5 (21.7)	0.208	14 (10.9)	8 (10.1)	6 (12.0)	0.739	0.504
O157	4 (1.2)	4 (1.7)	0 (0.0)	0.325 ^e	3 (1.6)	3 (1.8)	0 (0.0)	>0.99 ^e	1 (0.8)	1 (1.3)	0 (0.0)	>0.99 ^e	0.648 ^e
Plesiomonas	1 (0.3)	1 (0.4)	0 (0.0)	>0.99 ^e	1 (0.5)	1 (0.6)	0 (0.0)	>0.99 ^e	0 (0.0)	0 (0.0)	0 (0.0)	-	>0.99 ^e
Helicobacter pylori	21 (6.1)	21 (8.6)	0 (0.0)	0.001^e	13 (7.0)	13 (7.9)	0 (0.0)	0.374 ^e	8 (6.2)	8 (10.1)	0 (0.0)	0.023^e	>0.99

>1 bacteria detected	45 (13.1)	37 (15.2)	8 (7.9)	0.067	23 (12.3)	20 (12.2)	3 (13.0)	>0.99 ^e	11 (8.5)	8 (10.1)	3 (6.0)	0.528 ^e	0.287
Parasite^d	130 (37.8)	99 (40.7)	31 (30.7)	0.080	85 (45.5)	78 (47.6)	7 (30.4)	0.122	43 (33.3)	21 (26.6)	22 (44.0)	0.042	0.031
<i>Cystoisospora</i>	29 (8.4)	29 (11.9)	0 (0.0)	<0.001^e	29 (15.5)	20 (17.7)	0 (0.0)	0.028^e	0 (0.0)	0 (0.0)	0 (0.0)	-	<0.001^e
<i>Cryptosporidium</i> spp.	30 (8.7)	27 (11.1)	3 (3.0)	0.012^e	28 (15.0)	26 (15.9)	2 (8.7)	0.537 ^e	2 (1.6)	1 (1.3)	1 (2.0)	>0.99 ^e	<0.001^e
<i>Blastocystis</i>	49 (14.2)	25 (10.3)	24 (23.8)	0.001	15 (8.0)	12 (7.3)	3 (13.0)	0.403 ^e	33 (25.6)	13 (16.5)	20 (40.0)	0.003	<0.001
<i>Giardia</i> spp.	12 (3.5)	9 (3.7)	3 (3.0)	>0.99 ^e	7 (3.7)	7 (4.3)	0 (0.0)	0.600 ^e	4 (3.1)	2 (2.5)	2 (4.0)	0.641 ^e	>0.99 ^e
<i>Enterocytozoon</i> spp.	11 (3.2)	11 (4.5)	0 (0.0)	0.038^e	10 (5.4)	10 (6.1)	0 (0.0)	0.614 ^e	1 (0.8)	1 (1.3)	0 (0.0)	>0.99 ^e	0.031^e
<i>Schistosoma</i>	12 (3.5)	5 (2.1)	7 (6.9)	0.025	8 (4.3)	4 (2.4)	4 (17.4)	0.009^e	4 (3.1)	1 (1.3)	3 (6.0)	0.298 ^e	0.767 ^e
<i>Dientamoeba</i>	2 (0.6)	1 (0.4)	1 (1.0)	0.502 ^e	0 (0.0)	0 (0.0)	0 (0.0)	-	2 (1.6)	1 (1.3)	1 (2.0)	>0.99 ^e	0.166 ^e
>1 parasite detected	19 (5.5)	13 (5.4)	6 (5.9)	0.827	24 (12.8)	22 (13.4)	2 (8.7)	0.744 ^e	10 (7.8)	6 (7.6)	4 (8.0)	>0.99 ^e	0.152
Mixed infections	40 (11.6)	34 (14.0)	6 (5.9)	0.034		23 (14.0)	2 (8.7)	0.744 ^e	14 (10.9)	10 (12.7)	4 (8.0)	0.564 ^e	0.504
Virus-bacteria					25 (13.4)								
Virus-parasite	35 (10.2)	29 (11.9)	6 (5.9)	0.094	25 (13.4)	23 (14.0)	2 (8.7)	0.744 ^e	11 (8.5)	7 (8.9)	4 (8.0)	>0.99 ^e	0.183
Bacteria-parasite	72 (20.9)	60 (24.7)	12 (11.9)	0.008	46 (24.6)	41 (25.0)	5 (21.7)	0.734	23 (17.8)	16 (20.3)	7 (14.0)	0.366	0.152
>1 pathogen detected	115 (33.4)	92 (37.9)	23 (22.8)	0.007	73 (39.0)	66 (40.2)	7 (30.4)	0.366	40 (31.0)	26 (32.9)	14 (28.0)	0.557	0.143

261 ^aTotal is 344 for overall analysis and 316 for the HIV analysis due to 28 controls with unknown HIV-status being excluded from the HIV analysis; ^b*p*-value comparing
262 pathogen detection in cases versus controls (for HIV-infected and HIV-uninfected groups respectively); ^c*p*-value comparing pathogen detection in HIV-infected versus HIV-
263 uninfected persons (cases and controls combined); ^dThe following parasites were screened for but not detected: *Vibrio cholerae*, *Yersinia enterocolitica*, *E. histolytica*,
264 *Strongyloides* spp., *Cyclospora*, *Hymenolepsis*, *Ascaris*, *Taenia*, *Trichuris*, *Ancylostoma*, *Enterobius*, *Necator*: ^eFisher's exact test

265 *Effect of CD4+ cell count and treatment on clinical presentation and pathogen detection amongst*
266 *HIV-infected diarrhoeal cases*

267 For HIV-infected cases, presentation did not differ by CD4+ cell count for duration of symptoms,
268 fever, vomiting, fatigue, nausea, respiratory symptoms, abdominal pain, dermatological symptoms,
269 chills, headache, dysentery, arthralgia, myalgia, neurological symptoms or painful lymphadenopathy
270 (Table S2). There was an increased proportion of cases reporting weight loss with lower CD4+ cell
271 counts (91.4%, 81.0% and 62.5% for CD4+ cell counts of <200, 200-500, >500 cells/mm³ respectively;
272 $p=0.032$). Pathogens were detected more frequently in cases with low CD4+ cell counts (75.7%,
273 42.9% and 37.5% for CD4+ cell counts of <200, 200-500, >500 cells/mm³ respectively; $p=0.004$)
274 (Table S3). *Cryptosporidium* spp. specifically, was detected more often in those with low CD4+ cell
275 counts (25.7%, 4.8% and 0.0% for CD4+ cell counts of <200, 200-500, >500 cells/mm³ respectively;
276 $p=0.039$).

277 Clinical presentation was similar amongst HIV-infected cases on ART and those not on ART (Table
278 S4). Those on cotrimoxazole prophylaxis were more likely to experience fatigue (88.9% versus 69.3%,
279 $p=0.005$), weight loss (90.5% versus 76.1%, $p=0.023$) and chills (50.8% versus 15.9%, $p<0.001$) than
280 those not on cotrimoxazole. Pathogen detection was similar amongst HIV-infected cases on ART and
281 those not on ART (72.3% versus 88.5%, $p=0.082$) (Table S5). Parasites, specifically *Cystoisospora* spp.
282 were more commonly detected amongst patients on cotrimoxazole compared with those not on
283 cotrimoxazole (27.0% versus 11.4%, $p=0.014$). Other pathogens were found in similar proportions
284 across treatment groups.

285 *Clostridioides difficile* detection

286 *C. difficile* was detected in 14 cases, 13 of which were in HIV-infected patients (92.9%). Symptoms
287 were similar between cases in which *C. difficile* was detected and those without *C. difficile* detection,
288 however those with *C. difficile* detected were more likely to have had a longer duration of illness
289 before admission (6.5 (IQR 4-11) days versus 4 (IQR 2-7) days) (data not shown). It is important to

290 distinguish *C. difficile* colonisation from disease to avoid unnecessary antimicrobial therapy. The case
291 definition, as per the guidelines for diagnostic and clinical management of *C. difficile* from the South
292 African Society for Clinical Microbiology, includes onset of diarrhoea more than 48 hours after
293 admission, or diarrhoea that continues for 3 days post admission and where there is no likely
294 alternatively cause (no other pathogens detected or use of laxatives), or where the patient has had
295 an admission or antibiotics use in the past 12 weeks.⁴⁰ Three of the 14 cases did not match this case
296 definition as they did not have an admission in the past 12 weeks and had multiple pathogens
297 detected on PCR. These three cases were however HIV-infected. Five of the cases had no other
298 pathogens detected on PCR, all of which were HIV-infected. Eight of the cases did have a prior
299 admission, however the majority of these (n=6, 75.0%) had multiple pathogens detected on PCR. In
300 addition, the guidelines recommend that specimens testing positive for *C. difficile* on PCR (which is
301 highly sensitive for detection of *C. difficile*), should have further toxins A/B immunoassays performed
302 (which have high specificity for diagnosis of *C. difficile* infection). This was not done in the current
303 analysis.

304 Discussion

305 Despite advances in HIV treatment, diarrhoea amongst HIV-infected patients remains a significant
306 public health challenge¹⁹. Our study found that the majority (67.5%) of patients 5 years and older
307 with known HIV-status admitted to the study sites with diarrhoea were HIV-infected, highlighting the
308 importance of HIV-related illnesses among South African adults hospitalised with diarrhoea. The
309 Joint United Nations Programme on HIV/AIDS (UNAIDS) 90-90-90 treatment target aims to ensure
310 that 90% of the HIV-infected population are diagnosed, 90% of those diagnosed are initiated on ART
311 and 90% of those on ART should achieve viral suppression⁴¹. A recent South African analysis
312 estimated that 70.7% of diagnosed HIV-infected individuals have initiated ART, and 87.4% of these
313 are thought to be virally suppressed⁴¹. Proportions reported in the current analysis reflect these
314 numbers, as 83.3% of HIV-infected patients with known ART status were on treatment, however a

315 large proportion of HIV-infected patients included in this analysis had CD4+ cell count <200 cells/ μ l
316 (70/99, 70.7%) indicating that they are unlikely to be virally suppressed. Mpumalanga (home
317 province to the majority of the patients included in this analysis) is the only South African province in
318 which the third-90 indicator value (the percentage of those on ART that are virally suppressed) is
319 below 85%⁴¹. It is also likely that the study design was skewed towards detection of poorly
320 suppressed HIV cases as they are most likely to be hospitalised for diarrhoea. Diarrhoea in patients
321 with viral suppression will likely be seen at a clinic level. Further studies at clinic level are
322 recommended in order to include a broader spectrum of cases. The clinical presentation of
323 diarrhoea in HIV-infected cases differs to that of HIV-uninfected cases. HIV-infected diarrhoeal cases
324 were more likely to have symptoms for a longer duration before admission and to suffer from
325 weight loss and nausea. HIV-uninfected cases were more likely to suffer from dysentery. This finding
326 has not been reported in other studies, however it is likely related to aetiological differences
327 between HIV-infected and HIV-uninfected diarrhoeal patients. Amongst HIV-infected cases, patients
328 on cotrimoxazole were more likely to experience fatigue, weight loss and chills and were more likely
329 to have parasites detected, specifically *Cystoisospora* spp., compared to those not on cotrimoxazole.
330 This is an interesting finding since cotrimoxazole is the prescribed treatment for this pathogen. Since
331 all patients with CD4+ <200 cells/ μ l should receive cotrimoxazole prophylaxis, it can be used as a
332 surrogate marker for advanced HIV disease. However this finding may also point to changes in
333 *Cystoisospora* sensitivity for cotrimoxazole or may indicate that these patients are carriers or take
334 longer to clear the pathogen. The possibility of drug resistance should be further investigated in
335 future studies.

336 The use of molecular methods and the expanded panel of pathogens included in testing resulted in a
337 high pathogen detection rate, with pathogen detection being significantly higher in cases compared
338 with controls. Although viruses were commonly detected (18.5% of all specimens tested) and were
339 more likely to be detected in cases as compared with controls, when stratified by HIV status, viruses
340 were not significantly associated with diarrhoea in this study. Although this lack of association may

341 be partly due to asymptomatic carriage of viruses (specifically in HIV-uninfected patients), it is likely
342 that the sample size was insufficient to determine significant differences between cases and controls
343 when stratified by HIV status. Interestingly, norovirus GII was more prevalent in HIV-infected
344 individuals than HIV-uninfected individuals and detection in HIV-infected individuals was limited to
345 cases. Previous findings from South Africa have shown that HIV-infected children are more likely to
346 suffer poor outcomes related to norovirus infection⁴², however detection rates were similar amongst
347 HIV-infected and –uninfected children. This increased detection of norovirus in HIV-infected patients
348 ≥5 years of age has not been previously described. Bacteria was detected in 43.9% of specimens
349 tested and was significantly associated with diarrhoea overall and amongst HIV-uninfected patients.
350 Interestingly, prevalence of bacterial pathogens was high in both cases and controls within the HIV-
351 infected group. This was specifically due to the high prevalence of *E. coli* spp. (including ETEC, EPEC
352 and EAEC) amongst HIV-infected controls. High *E. coli* spp. colonisation rates in HIV-infected
353 patients have been documented in other settings⁴³. We found *Shigella* spp, *Salmonella* spp. and *H.*
354 *pylori* to be significantly associated with diarrhoea in the current study, specifically amongst HIV-
355 uninfected patients. It's likely that the sample size was insufficient to detect a difference in
356 prevalence between cases and controls in the HIV-uninfected group. *Campylobacter* spp. and *C.*
357 *difficile* were detected more often in HIV-infected patients as compared with HIV-uninfected
358 patients but were not significantly associated with diarrhoea (likely due to small sample sizes).
359 As expected from literature¹⁴, parasite prevalence was significantly higher amongst HIV-infected
360 compared with HIV-uninfected patients, specifically *Cystoisospora*, *Cryptosporidium* spp. and
361 *Enterocytozoon* spp., although only *Cystoisospora* was significantly associated with diarrhoea
362 amongst HIV-infected patients. The case-control analysis indicated that *Blastocystis* and *Schistosoma*
363 are unlikely to be causative agents of diarrhoea since these pathogens were detected at higher rates
364 in controls than in cases. Prevalence of *Schistosoma* among healthy individuals has been reported to
365 be as high as 51.2% in the Democratic Republic of Congo and is likely to be higher in villages and
366 rural areas where water is collected from rivers and dam⁴⁴. *Blastocystis* has been noted as an

367 important cause of diarrhoea amongst immunosuppressed patients⁴⁵, however this was not found in
368 the current study.

369 Diarrhoeal aetiology differed with HIV status, with *Shigella* spp. and *Salmonella* spp. being prevalent
370 amongst HIV-uninfected cases and *Shigella* spp., *Cystoisospora*, and *Cryptosporidium* spp. being
371 prevalent amongst HIV-infected cases. These differences should be considered during development
372 of diagnostic and treatment guidelines. This study highlights the importance of *Shigella* spp. in
373 diarrhoeal morbidity amongst adults (regardless of HIV status). According to the Global Burden of
374 Disease Study (GBD), *Shigella* spp. were the leading cause of diarrhoeal deaths amongst individuals
375 over the age of 5 years and the second leading cause of death in young children in 2016¹. While
376 there is currently no approved *Shigella* vaccine, there are several candidates which show promise for
377 efficacy testing⁴⁶. In order to introduce a candidate vaccine, burden estimates are required.

378 Although a larger sample size is required, this analysis gives an indication of the high burden of
379 *Shigella* spp. in our setting. Further work is required to determine if HIV-infected individuals should
380 be considered as a risk group for possible target should *Shigella* vaccine become available.

381 Our results are in line with clinical literature which indicates a shift in diarrhoeal aetiology with CD4+
382 cell count amongst HIV-infected patients. Pathogen detection was high in those with CD4+ cell
383 counts <200 cells/ μ l, indicating increased opportunistic infections (specifically *Cryptosporidium* spp.).
384 Pathogens could only be detected in just over a third of cases with high CD4+ cell count (>500
385 cells/ μ l), suggesting that these cases are possibly drug-related. This finding may be useful for
386 treatment guidelines. It also highlights the importance of addressing diarrhoea in patients on ART
387 especially since chronic diarrhoea has been identified as a major cause of dropout from ART care
388 services⁴⁷.

389 Detection of *C. difficile* in this study was almost exclusive to HIV-infected participants (13/14, 92.9%;
390 including 12 cases and 1 control). Data from another hospital study in Gauteng Province estimated
391 that community acquired *C. difficile* infections represent only 1.3% of cases (with the remaining

392 98.7% being hospital acquired)⁴⁸. Our study design excluded cases of hospital-acquired diarrhoea, as
393 enrolment and specimen collection was done within 48 hours of admission. The 5.4% *C. difficile*
394 prevalence reported here is lower than systematic review estimates of 15.8% amongst symptomatic
395 non-immunosuppressed in-patients from LMIC⁴⁹. This systematic review indicated similar rates
396 amongst immunosuppressed and non-immunosuppressed populations⁴⁹ whereas the current study
397 found that *C. difficile* was more commonly detected in HIV-infected cases than in –uninfected cases.
398 While not considered an opportunistic infection, *C. difficile* is associated with exposure to healthcare
399 settings and with the use of antibiotics, both of which are increased amongst HIV-patients⁵⁰. The
400 increased management of HIV patients in an outpatient setting⁵¹ may dispute the inclusion of recent
401 admission as a criterion in the case definition. There is also growing evidence to suggest the
402 possibility of *C. difficile* being an important community-acquired pathogen⁵¹, and consideration of
403 this should be given in populations with high HIV prevalence.

404 A major limitation of this study was the small sample size which restricted the analysis (attributable
405 fractions and odds ratios could not be calculated as many categories had zero count). The ANDEMIA
406 study was not specifically aimed at enrolling HIV-infected cases and controls, therefore this analysis
407 relied on incidental HIV enrolments. Enrolment, specifically of controls, was also hampered by
408 COVID-19 restrictions. An extension of this analysis would benefit from enrolling controls from HIV
409 clinics. Viral load was not analysed here, since few recent results were available; thus recent CD4+
410 cell count was used as a proxy. CD4+ cell counts were not available for all HIV-infected cases at the
411 time of admission. Many of those without CD4+ cell counts available were matched from NHLS data,
412 however these tests were not always performed at the same time as the enrolment. We used a cut
413 off of 12 months on either side of enrolment for CD4+ cell counts as this was deemed to be a
414 reasonable estimation of the CD4+ cell count at admission. In cases where there were significant
415 changes in CD4+ cell counts over time, these may have been incorrectly quantified. There were also
416 limited data available for antibiotic use. Date of last dose was available, however start date for
417 antibiotic course was not collected. This has limitations specifically for the *C. difficile* analysis, since

418 we could not determine the length of antibiotic courses. Future studies should also investigate the
419 expansion of diagnostic testing for *M. avium* which was not done in the present study.

420 In conclusion, this study highlights the importance of HIV-related diarrhoea amongst South African
421 inpatients ≥ 5 years of age, and underscores the importance of research in a high HIV prevalence
422 setting in order to improve the understanding of aetiology of diarrhoea. These data are important
423 for development of guidelines and treatment protocols as well as for preventative interventions,
424 such as vaccine introduction. Our data specifically highlights the importance of *Shigella* spp. amongst
425 both HIV-infected and HIV-uninfected diarrhoeal patients. We suggest that research be expanded to
426 primary healthcare levels in order to include a broader spectrum of HIV disease, specifically to
427 investigate diarrhoea in individuals with well-controlled HIV.

428 Acknowledgments

429 We wish to acknowledge the ANDEMIA participants, funders, staff and clinical staff at the hospitals.

430 Thank you to Ann Christin Vietor for statistical assistance.

431

432

433

434

435

436

437

438

439

440 References:

- 441 1. Troeger C, Blacker BF, Khalil IA, Rao PC, Cao S, Zimsen SR, et al. Estimates of the global,
442 regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a
443 systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis.*
444 2018;18(11):1211–28.
- 445 2. Fischer Walker CL, Black RE. Diarrhoea morbidity and mortality in older children, adolescents,
446 and adults. *Epidemiol Infect.* 2010;138:1215–26.
- 447 3. Farthing M, Salam MA, Lindberg G, Dite P, Khalif I, Salazar-Lindo E, et al. Acute diarrhea in
448 adults and children: A global perspective. *J Clin Gastroenterol.* 2013;47(1):12–20.
- 449 4. Dwivedi K, Prasad G, Saini S, Mahajan S, Lal S, Baveja U. Enteric opportunistic parasites
450 among HIV infected individuals: associated risk factors and immune status. *Jpn J Infect Dis.*
451 2007;60:76–81.
- 452 5. Kelly P, Todd J, Sianongo S, Mwansa J, Sinsungwe H, Katubulushi M, et al. Susceptibility to
453 intestinal infection and diarrhoea in Zambian adults in relation to HIV status and CD4 count.
454 *BMC Gastroenterol.* 2009;9(7).
- 455 6. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H. A comparative risk
456 assessment of burden of disease and injury attributable to 67 risk factor clusters in 21
457 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010.
458 *Lancet.* 2012;380(9859).
- 459 7. Grace D. Food safety in low and middle income countries. *Int J Environ Res Public Heal.*
460 2015;12:10490–507.
- 461 8. Jansen A, Stark K, Kunkel J, Schreier E, Ignatius R, Liesenfeld O, et al. Aetiology of community-
462 acquired, acute gastroenteritis in hospitalised adults: a prospective cohort study. *BMC Infect*
463 *Dis.* 2008;8(1):1–7.

- 464 9. Bresee JS, Marcus R, Venezia RA, Keene WE, Morse D, Thanassi M, et al. The etiology of
465 severe acute gastroenteritis among adults visiting emergency departments in the United
466 States. *J Infect Dis.* 2012;205(9):1374–81.
- 467 10. Fletcher S, Stark D, Ellis J. Prevalence of gastrointestinal pathogens in Sub-Saharan Africa:
468 systematic review and meta-analysis. *J Public Heal Africa.* 2011;2(2).
- 469 11. Joint United Nations Programme on HIV/AIDS (UNAIDS). The Gap Report [Internet]. 2014.
470 Available from: [https://unaidstest.unaids.org/sites/default/files/unaidstest.unaids.org/sites/default/files/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf](https://unaidstest.unaids.org/sites/default/files/unaidstest.unaids.org/sites/default/files/unaidstest.unaids.org/sites/default/files/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf)
471
472
- 473 12. Germani Y, Minssart P, Vohito M, Yassibanda S, Glaziou P, Hocquet D, et al. Etiologies of
474 acute, persistent, and dysenteric diarrheas in adults in Bangui, Central African Republic, in
475 relation to human immunodeficiency virus serostatus. *Am J Trop Med Hyg.* 2017;59(6):1008–
476 14.
- 477 13. Acácio S, Nhampossa T, Quintó L, Vubil D, Saco C, Kotloff K, et al. The role of HIV infection
478 in the etiology and epidemiology of diarrheal disease among children aged 0–59 months in
479 Manhiça District, Rural Mozambique. *Int J Infect Dis.* 2018;73:10–7.
- 480 14. Feasey NA, Healey P, Gordon MA. Review article: the aetiology, investigation and
481 management of diarrhoea in the HIV-positive patient. *Aliment Pharmacol Ther.* 2011;34:587–
482 603.
- 483 15. Thomas P, Pollok R, Gazzard B. Enteric viral infections as a cause of diarrhoea in the acquired
484 immunodeficiency syndrome. *HIV Med.* 1999;1:19–24.
- 485 16. Wang Z, Liu Q, Liu H, Li S, Zhang L, Zhao Y, et al. Prevalence of *Cryptosporidium*, microsporidia
486 and *Isospora* infection in HIV-infected people: a global systematic review and meta-analysis.
487 *Parasites and Vectors.* 2018;11(28).

- 488 17. Vyas N, Pathan N, Aziz A. Enteric pathogens in HIV-positive patients with diarrhoea and their
489 correlation with CD4+ T-lymphocyte counts. *Trop Parasitol.* 2012;2(1):29–34.
- 490 18. Asmuth D, Clay P, Blick G, Chaturvedi P, MacFarlane K. Incidence and prevalence of diarrhoea
491 in HIV clinical trials in the recent post-cART era: analysis of data from 38 clinical trials from
492 2008-2017 over 21,000 patients. In: 22nd International AIDS Conference. Amsterdam; 2018.
- 493 19. Logan C, Beadsworth M, Beeching N. HIV and diarrhoea: what is new? *Curr Opin Infect Dis.*
494 2016;29(5):486–94.
- 495 20. Harris E, Washington K. Manifestations of immunodeficiency in the GI tract. In: Odze R,
496 Goldblum J, editors. *Surgical pathology of the GI tract, liver, biliary tract, and pancreas*
497 (Second edition). 2nd ed. Elsevier Inc.; 2009. p. 81–96.
- 498 21. Awole M, Gebre-Selassie G, Kassa T, Kibru G. Isolation of potential bacterial pathogens from
499 the stool of HIV-infected and HIV-non-infected patients and their antimicrobial susceptibility
500 patterns in Jimma Hospital, South West Ethiopia. *Ethiop Med J.* 2002;40:353–64.
- 501 22. Dikman AE, Schonfeld E, Srisarajivakul NC, Poles MA. Human Immunodeficiency Virus-
502 Associated Diarrhea: Still an Issue in the Era of Antiretroviral Therapy. *Dig Dis Sci.*
503 2015;60(8):2236–45.
- 504 23. Oude Munnink B, Canuti M, Deijs M, de Vries M, Jebbink M, Rebers S, et al. Unexplained
505 diarrhoea in HIV-1 infected individuals. *BMC Infect Dis.* 2014;14(22).
- 506 24. Call S, Heudebert G, Saag M, Wilcox C. The changing etiology of chronic diarrhea in HIV-
507 infected patients with CD4 cell counts less than 200 cells/mm³. *Am J Gastroenterol.*
508 2000;95(11):3142–6.
- 509 25. Rubaihayo J, Tumwesigye N, Konde-Lule J, Wamani H, Nakku-Joloba E, Makumbi F. Frequency
510 and distribution patterns of opportunistic infections associated with HIV/AIDS in Uganda.
511 *BMC Res Notes.* 2016;9(1).

- 512 26. Rowe J, Shah S, Motlhagodi S, Bafana M, Tawanana E, Truong H, et al. An epidemiologic
513 review of enteropathogens in Gaborone, Botswana: Shifting patterns of resistance in an HIV
514 epidemic region. *PLoS One*. 2010;5(6):e10924.
- 515 27. Ahmadpour E, Safarpour H, Xiao L, Zarean M, K H-N, Barac A, et al. Cryptosporidiosis in HIV-
516 positive patients and related risk factors: A systematic review and meta-analysis. *Parasite*.
517 2020;27(27).
- 518 28. Obi CL, Bessong PO. Diarrhoeagenic bacterial pathogens in HIV-positive patients with
519 diarrhoea in rural communities of Limpopo Province, South Africa. *J Heal Popul Nutr*.
520 2002;20(3):230–4.
- 521 29. Groome MJ, Madhi SA. Five-year cohort study on the burden of hospitalisation for acute
522 diarrhoeal disease in African HIV-infected and HIV-uninfected children: Potential benefits of
523 rotavirus vaccine. *Vaccine*. 2012;30(S1):A173–8.
- 524 30. Cheung T, Teich S. Cytomegalovirus infection in patients with HIV infection. *Mt Sinai J Med*.
525 1999;66(2):113–24.
- 526 31. Azer S, Limaiem G. Cytomegalovirus colitis. Florida: StatPearls; 2022. 6 p.
- 527 32. Legenza L, Barnett S, Rose W, Bianchini M, Safdar N, Coetzee R. Epidemiology and outcomes
528 of *Clostridium difficile* infection among hospitalised patients: results of a multicentre
529 retrospective study in South Africa. *BMJ Glob Heal*. 2018;3(4):e000889.
- 530 33. Johnson L, Dorrington R, Moolla H. Progress towards the 2020 targets for HIV diagnosis and
531 antiretroviral treatment in South Africa. *South Afr J HIV Med*. 2017;18(1):a694.
- 532 34. Moshabela M, Macpherson P, Ezard N, Freat E, Mashimbye L, Elliott J, et al. Clinical and
533 social determinants of diarrhoeal disease in a rural HIV/AIDS clinic , South Africa : a case-
534 control study. *Int J STD AIDS*. 2014;23(5):346–50.

- 535 35. Elfstrand L, Floren C. Management of chronic diarrhea in HIV-infected patients: current
536 treatment options, challenges and future directions. *HIV AIDS*. 2010;2:219–24.
- 537 36. Riddle MS, Dupont HL, Connor BA. ACG clinical guideline: Diagnosis, treatment, and
538 prevention of acute diarrheal infections in adults. *Am J Gastroenterol*. 2016;111(5):602–22.
- 539 37. Schubert G, Achi V, Ahuka S, Belarbi E, Bourhaima O, Eckmanns T, et al. The African Network
540 for Improved Diagnostics, Epidemiology and Management of common infectious Agents.
541 *BMC Infect Dis*. 2021;21(1):539.
- 542 38. Delpont S, Swanepoel J, Odendaal P, Roux P. Incidence of retinopathy of prematurity in very-
543 low-birth-weight infants born at Kalafong Hospital, Pretoria. *South African Med J*.
544 2002;92(12):986–90.
- 545 39. Msimang VMY, Page N, Groome MJ, Moyes J, Cortese MM, Seheri M, et al. Impact of
546 rotavirus vaccine on childhood diarrheal hospitalization after introduction into the South
547 African Public Immunization Program. *Pediatr Infect Dis J*. 2013;32(12):1359–64.
- 548 40. Nana T, Moore C, Boyles T, Brink AJ, Cleghorn J, Devenish LM, et al. South African Society of
549 Clinical Microbiology *Clostridioides difficile* infection diagnosis, management and infection
550 prevention and control guideline. *South African J Infect Dis*. 2020;35(1).
- 551 41. Marinda E, Simbayi L, Zuma K, Zungu N, Moyo S, Kondlo L, et al. Towards achieving the 90–
552 90–90 HIV targets: Results from the South African 2017 national HIV survey. *BMC Public*
553 *Health*. 2020;20(1):1–12.
- 554 42. Page NA, Groome MJ, Nadan S, Netshikweta R. Norovirus epidemiology in South African
555 children <5 years hospitalised for diarrhoeal illness between 2009 and 2013. *Epidemiol Infect*.
556 2017;
- 557 43. Garcia C, Chinchu O, Leon M, Iglesias D, Barletta F, Mercado E, et al. Short Report : High
558 Frequency of Diarrheagenic *Escherichia coli* in Human Immunodeficiency Virus (HIV) Patients

559 with and without Diarrhea in Lima , Perú. *Am J Trop Med Hyg.* 2010;82(6):1118–20.

560 44. Madinga J, Id LM, Barbe B, Linsuke S, Hermy M, Heroes A, et al. Non-typhoidal Salmonella
561 intestinal carriage in a Schistosoma mansoni endemic community in a rural area of the
562 Democratic Republic of Congo. *PLoS Negl Trop Dis.* 2020;14(10):e0007875.

563 45. Rao K, Sekar U, Iraivan KT, Abraham G, Soundararajan P. Blastocystis hominis — An Emerging
564 Cause of Diarrhoea in Renal Transplant Recipients. *J Assoc Physicians India.* 2003;51:719–21.

565 46. Hosangadi D, Smith PG, Giersing BK. Considerations for using ETEC and Shigella disease
566 burden estimates to guide vaccine development strategy. *Vaccine.* 2019;37(50):7372–80.

567 47. Asimwe SB, Kanyesigye M, Bwana B, Okello S, Muyindike W. Predictors of dropout from care
568 among HIV-infected patients initiating antiretroviral therapy at a public sector HIV treatment
569 clinic in sub-Saharan Africa. *BMC Infect Dis.* 2016;16(43).

570 48. Godinho L, Feldman C. Clostridium difficile infection in Johannesburg, South Africa. *Hosp
571 Pract.* 2018;46(5):287–94.

572 49. Forrester JD, Cai LZ, Mbanje C, Rinderknecht TN, Wren SM. Clostridium difficile infection in
573 low- and middle-human development index countries: a systematic review. *Trop Med Int
574 Heal.* 2017;22(10):1223–32.

575 50. Revolinski SL, Munoz-Price LS. Clostridium difficile in Immunocompromised Hosts: A Review
576 of Epidemiology, Risk Factors, Treatment, and Prevention. *Clin Infect Dis.* 2019;68(12):2144–
577 53.

578 51. Collini PJ, Kuijper E, Dockrell DH. Clostridium Difficile Infection in Patients with HIV/AIDS. *Curr
579 HIV/AIDS Rep.* 2013;10:273–82.

580

581

582 Supplementary tables:

583 Table S1: Pathogens included in molecular testing

Pathogen	Testing platform
Viruses	
Norovirus GI	FTD viral gastroenteritis, TaqMan Array Cards
Norovirus GII	FTD viral gastroenteritis, TaqMan Array Cards
Human astrovirus	FTD viral gastroenteritis, TaqMan Array Cards
Rotavirus	FTD viral gastroenteritis, TaqMan Array Cards
Human adenovirus	FTD viral gastroenteritis, TaqMan Array Cards
Sapovirus	FTD viral gastroenteritis, TaqMan Array Cards
Enterovirus	TaqMan Array Cards
CMV	Monoplex PCR
Bacteria	
<i>Campylobacter coli/jejuni/lari</i>	FTD bacterial gastroenteritis, TaqMan Array Cards
<i>Clostridioides difficile</i>	FTD bacterial gastroenteritis, TaqMan Array Cards
<i>Escherichia coli</i> verotoxin positives	FTD bacterial gastroenteritis
<i>Salmonella</i> spp.	FTD bacterial gastroenteritis, TaqMan Array Cards
<i>Shigella</i> spp.	FTD bacterial gastroenteritis, TaqMan Array Cards
Enteroinvasive <i>Escherichia coli</i>	FTD bacterial gastroenteritis, TaqMan Array Cards
<i>Yersinia enterocolitica</i>	FTD bacterial gastroenteritis, TaqMan Array Cards
EAEC	TaqMan Array Cards
EPEC	TaqMan Array Cards
ETEC	TaqMan Array Cards
STEC/VTEC	FTD bacterial gastroenteritis; TaqMan Array Cards
<i>E. coli</i> O157	TaqMan Array Cards
Parasites	
<i>Plesiomonas shigelloides</i>	TaqMan Array Cards
<i>Entamoeba histolytica</i>	FTD enteric parasites, TaqMan Array Cards, SeeGene Allplex
<i>Cryptosporidium</i> spp.	FTD enteric parasites, TaqMan Array Cards, SeeGene Allplex
<i>Giardia lamblia</i>	FTD enteric parasites, TaqMan Array Cards, SeeGene Allplex
<i>Vibrio cholerae</i>	TaqMan Array Cards
<i>Helicobacter pylori</i>	TaqMan Array Cards
<i>Schistosoma</i>	TaqMan Array Cards
<i>Enterocytozoon</i> spp.	TaqMan Array Cards, SeeGene Allplex
<i>Strongyloides stercoralis</i>	TaqMan Array Cards, SeeGene Allplex
<i>Cytoisospora belli</i>	TaqMan Array Cards
<i>Blastocystis hominis</i>	SeeGene Allplex
<i>Dientamoeba fragilis</i>	SeeGene Allplex
<i>Cyclospora cayentanensis</i>	SeeGene Allplex
<i>Hymenolepis nana</i>	SeeGene Allplex

<i>Ascaris lumbricoides</i>	SeeGene Allplex
<i>Taenia</i> spp.	SeeGene Allplex
<i>Trichuris trichiura</i>	SeeGene Allplex
<i>Ancylostoma duodenale</i>	SeeGene Allplex
<i>Enterobius vermicularis</i>	SeeGene Allplex
<i>Necator americanus</i>	SeeGene Allplex

584

585 Table S2: Clinical presentation for HIV-infected cases, stratified by CD4+ cell counts

	CD4+ cell count ^a			p-value
	<200 cells/ μ l (n=70)	200-500 cells/ μ l (n=21)	>500 cells/ μ l (n=8)	
Duration of symptoms before admission – median (IQR)	6 (3-10)	7 (3.5-12)	5.5 (3-7)	0.822
Chronic/persistent diarrhoea	6 (8.6%)	2 (9.5%)	1 (12.5%)	0.855
Weight loss	64 (91.4%)	17 (81.0%)	5 (62.5%)	0.032
Fatigue	58 (82.9%)	16 (76.2%)	5 (62.5%)	0.293
Nausea	6 (75.0%)	15 (71.4%)	6 (75.0%)	>0.99
Vomiting	48 (68.6%)	14 (66.7%)	6 (75.0%)	>0.99
Fever (current or history)	46 (65.7%)	14 (66.7%)	5 (62.5%)	>0.99
Abdominal pain	41 (58.6%)	14 (66.7%)	4 (50.0%)	0.667
Respiratory symptoms	33 (47.1%)	12 (57.1%)	1 (12.5%)	0.097
Chills	32 (45.7%)	7 (33.3%)	3 (37.5%)	0.634
Headache	32 (45.7%)	7 (33.3%)	4 (50.0%)	0.635
Dermatological symptoms	13 (18.6%)	3 (14.3%)	0 (0.0%)	0.494
Arthralgia	11 (15.7%)	6 (28.6%)	0 (0.0%)	0.171
Myalgia	11 (15.7%)	1 (4.8)%	0 (0.0%)	0.334
Neurological symptoms	5 (7.1%)	0 (0.0%)	1 (12.5%)	0.347
Dysentery ^b	2 (2.9%)	0 (0.0%)	0 (0.0%)	>0.99
Painful swollen glands	2 (2.9%)	0 (0.0%)	0 (0.0%)	>0.99

586 ^a CD4+ count within 12 months of enrolment only known for 99 of the 164 HIV-infected patients included in
587 molecular testing; ^b Dysentery defined as self-reported blood in the stool.

588

589

590

591

592

593

594

595 Table S3: Pathogen detection in specimens of HIV-infected cases, stratified by CD4+ cell count

	CD4+ count ^a - n (%)			p-value
	>500 cells/ μ l (n=8)	200-500 cells/ μ l (n=21)	<200 cells/ μ l (n=70)	
Any pathogen	5 (62.5)	12 (57.1)	59 (84.3)	0.022
Virus	2 (25.0)	5 (23.8)	19 (27.1)	>0.99
Adenovirus	0 (0.0)	5 (23.8)	8 (11.4)	0.170
Adenovirus 40/41	0 (0.0)	1 (4.8)	0 (0.0)	0.293
Norovirus	0 (0.0)	1 (4.8)	6 (8.6)	>0.99
Norovirus GI	0 (0.0)	0 (0.0)	0 (0.0)	-
Norovirus GII	0 (0.0)	1 (4.8)	6 (8.6)	>0.99
Enterovirus	2 (25.0)	0 (0.0)	7 (10.0)	0.068
CMV	0 (0.0)	0 (0.0)	2 (2.9)	>0.99
Astrovirus	0 (0.0)	0 (0.0)	2 (2.9)	>0.99
Rotavirus	0 (0.0)	0 (0.0)	1 (1.4)	>0.99
Sapovirus	0 (0.0)	0 (0.0)	1 (1.4)	>0.99
>1 virus detected	0 (0.0)	1 (4.8)	5 (7.1)	>0.99
Bacteria	4 (50.0)	8 (38.1)	35 (50.0)	0.638
Shigella spp.	1 (12.5)	1 (4.8)	12 (17.1)	0.389
Salmonella spp.	0 (0.0)	0 (0.0)	7 (10.0)	0.322
C. difficile	0 (0.0)	1 (4.8)	9 (12.9)	0.450
Campylobacter	0 (0.0)	0 (0.0)	4 (5.7)	0.695
STEC	0 (0.0)	0 (0.0)	0 (0.0)	-
ETEC	0 (0.0)	0 (0.0)	2 (2.9)	>0.99
EPEC	0 (0.0)	0 (0.0)	6 (8.6)	0.464
EAEC	1 (12.5)	4 (19.1)	8 (11.4)	0.622
O157	0 (0.0)	0 (0.0)	2 (2.9)	>0.99
Plesiomonas	0 (0.0)	1 (4.8)	0 (0.0)	0.293
Helicobacter pylori	2 (25.0)	2 (9.5)	4 (5.7)	0.109
>1 bacteria detected	0 (0.0)	1 (4.8)	14 (20.0)	0.131
Parasite^b	2 (25.0)	8 (38.1)	39 (55.7)	0.145
Cystoisospora	1 (12.5)	4 (19.1)	12 (17.1)	>0.99
Cryptosporidium spp.	0 (0.0)	1 (4.8)	19 (27.1)	0.028
Blastocystis	1 (12.5)	2 (9.5)	5 (7.1)	0.563
Giardia spp.	0 (0.0)	1 (4.8)	4 (5.7)	>0.99
Enterocytozoon spp.	0 (0.0)	1 (4.8)	6 (8.6)	>0.99
Schistosoma	0 (0.0)	0 (0.0)	1 (1.4)	>0.99
>1 parasite detected	0 (0.0)	1 (4.8)	8 (11.4)	0.611
Mixed infections				
Virus-bacteria	1 (12.5)	4 (19.1)	9 (12.9)	0.793
Virus-parasite	1 (12.5)	3 (14.3)	12 (17.1)	>0.99
Bacteria-parasite	2 (25.0)	5 (23.8)	21 (30.0)	0.032

596 ^a CD4+ count within 12 months of enrolment only known for 99 of the 164 HIV-infected cases; ^b The following
597 parasites were screened for but not detected: *Vibrio cholerae*, *Yersinia enterocolitica*, *E. histolytica*,
598 *Strongyloides* spp., *Cyclospora*, *Hymenolepis*, *Ascaris*, *Taenia*, *Trichuris*, *Ancylostoma*, *Enterobius*, *Necator*,
599 *Dientamoeba*.

600 Table S4: Clinical presentation of HIV-infected cases, stratified by treatment

	Antiretroviral treatment ^a - n (%)			Cotrimoxazole prophylaxis ^a - n (%)		
	Yes (n=130)	No (n=26)	p-value	Yes (n=63)	No (n=88)	p-value
Duration of symptoms before admission – median (IQR)	5 (3-8)	4.5 (2-10)	0.489	5 (3-9)	4 (3-8)	0.410
Chronic/persistent diarrhoea	10 (7.7)	3 (11.5)	0.456	6 (9.5)	7 (8.0)	0.774
Weight loss	108 (83.1)	19 (73.1)	0.231	57 (90.5)	67 (76.1)	0.023
Fatigue	98 (75.4)	19 (73.1)	0.804	56 (88.9)	61 (69.3)	0.005
Nausea	92 (70.8)	19 (73.1)	0.813	43 (68.3)	65 (73.9)	0.451
Fever (current or self-reported history in the past 10 days)	89 (68.5)	20 (76.9)	0.391	43 (68.3)	61 (69.3)	0.889
Vomiting	89 (68.5)	22 (84.6)	0.153	42 (66.7)	67 (76.1)	0.200
Abdominal pain	83 (63.9)	18 (69.2)	0.600	42 (66.7)	59 (67.1)	0.961
Respiratory symptoms	59 (45.4)	13 (50.0)	0.667	29 (46.0)	41 (46.6)	0.946
Headache	50 (38.5)	12 (46.2)	0.464	30 (47.6)	33 (37.5)	0.214
Chills	39 (30.0)	6 (23.1)	0.477	32 (50.8)	14 (15.9)	<0.001
Arthralgia	32 (24.6)	2 (7.7)	0.069	13 (20.6)	21 (23.9)	0.640
Dermatological symptoms	21 (16.2)	2 (11.5)	0.768	10 (15.9)	14 (15.9)	0.995
Myalgia	14 (10.8)	1 (3.9)	0.469	10 (15.9)	6 (6.8)	0.075
Dysentery ^b	6 (4.6)	1 (3.9)	>0.99	3 (4.8)	4 (4.6)	>0.99
Neurological symptoms	5 (3.9)	1 (3.9)	>0.99	4 (6.4)	2 (2.3)	0.236
Painful swollen glands	3 (2.3)	0 (0.0)	>0.99	2 (3.2)	1 (1.1)	0.571

601 ^a ART status known for 156 and cotrimoxazole use known for 151 of 164 HIV-infected cases; ^b Dysentery
602 defined as self-reported blood in the stool.

603
604
605
606
607
608
609
610
611
612
613

614 Table S5: Pathogen detection in specimens of HIV-infected cases, stratified by treatment

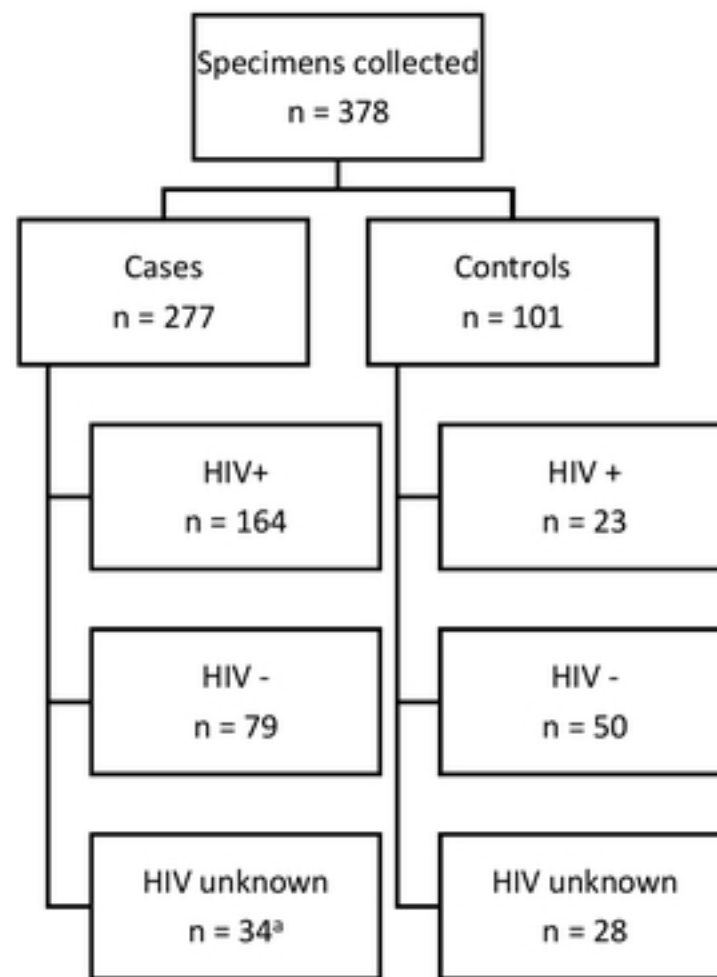
	Antiretroviral treatment ^a - n (%)			Cotrimoxazole prophylaxis ^a - n (%)		
	No (n=26)	Yes (n=130)	p-value	No (n=88)	Yes (n=63)	p-value
Any pathogen	23 (88.5)	94 (72.3)	0.082	66 (75.0)	49 (77.8)	0.693
Virus	6 (23.1)	32 (24.6)	0.868	20 (22.7)	16 (25.4)	0.704
Adenovirus	3 (11.5)	14 (10.8)	>0.99	9 (10.2)	8 (12.7)	0.795
Adenovirus 40/41	0 (0.0)	1 (0.8)	>0.99	1(1.1)	0 (0.0)	>0.99
Norovirus	3 (11.5)	10 (7.7)	0.456	6 (6.8)	7 (11.1)	0.354
Norovirus GI	0 (0.0)	1 (0.8)	>0.99	1 (1.1)	0 (0.0)	>0.99
Norovirus GII	3 (11.5)	9 (6.9)	0.423	5 (5.7)	7 (11.1)	0.240
Enterovirus	1 (3.9)	8 (6.2)	>0.99	4 (4.6)	4 (6.4)	0.720
CMV	0 (0.0)	3 (2.3)	>0.99	2 (2.3)	1 (1.6)	>0.99
Astrovirus	0 (0.0)	3 (2.3)	>0.99	0 (0.0)	2 (3.2)	0.172
Rotavirus	0 (0.0)	1 (0.8)	>0.99	0 (0.0)	1 (1.6)	0.417
Sapovirus	0 (0.0)	0 (0.0)	-	0 (0.0)	0 (0.0)	-
>1 virus detected	1 (3.9)	5 (3.9)	>0.99	1 (1.1)	5 (7.9)	0.083
Bacteria	16 (61.5)	62 (47.7)	0.197	44 (50.0)	32 (50.8)	0.923
Shigella spp.	5 (19.2)	24 (18.5)	>0.99	18 (20.5)	9 (14.3)	0.329
Salmonella spp.	2 (7.7)	10 (7.7)	>0.99	6 (6.8)	5 (7.9)	>0.99
C. difficile	2 (7.7)	10 (7.7)	>0.99	6 (6.8)	6 (9.5)	0.544
Campylobacter	1 (3.9)	10 (7.7)	0.692	4 (4.6)	7 (11.1)	0.202
STEC	0 (0.0)	0 (0.0)	-	0 (0.0)	0 (0.0)	-
ETEC	1 (3.9)	3 (2.3)	0.522	3 (3.4)	1 (1.6)	0.641
EPEC	2 (7.7)	8 (6.2)	0.673	6 (6.8)	5 (7.9)	>0.99
EAEC	2 (7.7)	18 (13.9)	0.531	10 (11.4)	8 (12.7)	0.803
O157	0 (0.0)	3 (2.3)	>0.99	1 (1.1)	2 (3.2)	0.571
Plesiomonas	0 (0.0)	1 (0.8)	>0.99	0 (0.0)	1 (1.6)	0.417
Helicobacter pylori	4 (15.4)	8 (6.2)	0.117	6 (6.8)	6 (9.5)	0.544
>1 bacteria detected	2 (7.7)	24 (18.5)	0.252	11 (12.5)	13 (20.6)	0.178
Parasite^b	12 (46.2)	59 (45.4)	>0.99	33 (37.5)	37 (58.7)	0.010
Cystoisospora	4 (15.4)	24 (18.5)	>0.99	10 (11.4)	17 (26.9)	0.014
Cryptosporidium spp.	5 (19.2)	21 (16.2)	0.773	12 (13.6)	14 (22.2)	0.168
Blastocystis	4 (15.4)	6 (4.6)	0.063	8 (9.1)	2 (3.2)	0.195
Giardia spp.	1 (3.9)	5 (3.9)	>0.99	4 (4.6)	2 (3.2)	>0.99
Enterocytozoon spp.	0 (0.0)	10 (7.7)	0.215	4 (4.6)	6 (9.5)	0.321
Schistosoma	0 (0.0)	4 (3.1)	>0.99	3 (3.4)	1 (1.6)	0.641
>1 parasite detected	2 (7.7)	10 (7.7)	>0.99	7 (8.0)	5 (7.9)	>0.99
Mixed infections						
Virus-bacteria	3 (11.5)	20 (15.4)	0.768	11 (12.5)	11 (17.5)	0.394
Virus-parasite	2 (7.7)	19 (14.6)	0.531	8 (9.1)	11 (17.5)	0.126

Bacteria- parasite	8 (30.8)	34 (26.2)	0.633	19 (21.6)	22 (34.9)	0.069
-------------------------------	----------	-----------	-------	-----------	-----------	-------

615 ^a ART status known for 156 and cotrimoxazole use known for 151 of 164 HIV-infected cases; ^b The following
616 parasites were screened for but not detected: *E. histolytica*, *Strongyloides* spp., *Cyclospora*, *Hymenolepis*,
617 *Ascaris*, *Taenia*, *Trichuris*, *Ancylostoma*, *Enterobius*, *Necator*, *Dientamoeba*.

618

619



^a Due to limited TAC tests available, cases with unknown HIV status were excluded from the analysis. Controls with unknown HIV status were included to increase sample size for the overall case-control analysis, but were not included in the HIV specific analysis.

Figure 1: Specimens included in the analysis