

1 **Detection and quantification of hepatitis A virus titers from wastewater in South Africa**
2 **and comparison with clinical data from the National Surveillance Database**

3

4 **Short title:** Detecting and quantifying Hepatitis: A virus in South African wastewater

5

6 Kathleen Subramoney^{1,2}, Sipho Gwala¹, Emmanuel Phalane¹, Mokgaetji Macheke¹, Natasha

7 Singh¹, Thabo Mangena¹, Nkosenhle Ndlovu¹, Nosihle Msomi¹, Sibonginkosi Maposa¹,

8 Chenoa Sankar¹, Fiona Els^{1,3}, Phindile Ntuli¹, Mantshali Motlounge¹, Victor Mabasa¹,

9 Kerrigan McCarthy^{1,2} ¶, Mukhlid Yousif^{1,2} ¶ *

10

11 ¹Centre for Vaccines and Immunology, National Institute for Communicable Diseases,

12 Johannesburg 2192, South Africa

13 ²School of Pathology, University of Witwatersrand, Johannesburg 2017, South Africa

14 ³Gauteng City-Region Observatory (GCRO), a Partnership of the University of

15 Johannesburg, the University of the Witwatersrand, the Gauteng Provincial Government and

16 Organised Local Government in Gauteng (SALGA), Private Bag 3, Wits 2050, South Africa

17 ¶Joint last authors

18 *Corresponding author. Email: mukhlidy@nicd.ac.za, mukhlid.yousif@wits.ac.za

19

20 **Abstract**

21 Wastewater surveillance is useful for monitoring the prevalence of hepatitis A virus (HAV).
22 We developed and optimized HAV detection and quantification methods for wastewater
23 samples, and applied them to samples collected through a national wastewater surveillance
24 program. Previously identified 5'-untranslated region-targeting primers and probes were used
25 to develop the assay. Serial dilutions of HAV-positive clinical samples were used to validate
26 and determine limits of quantification (LOQ). Retrospective testing of weekly wastewater
27 samples collected through the SARS-CoV-2 wastewater surveillance program at 26 sites in
28 Gauteng (August 2021 to March 2024) were undertaken using ultrafiltration-based
29 concentration, and nucleic acids were extracted using the KingFisher Flex purification system
30 with a wastewater isolation kit. A digital PCR assay was used for HAV detection and
31 quantification (as genome copies/ μL). Clinical data from the Surveillance Database
32 Warehouse of the National Health Laboratory Service were compared with wastewater data,
33 epidemiological week-wise and district-wise, to determine correlations between the datasets.
34 Based on the validation results, one partition on the digital PCR (dPCR) platform was
35 equivalent to an LOQ of 0.4 genome copies/ μL . In total, 2013 wastewater samples were
36 tested, of which 349 were positive for HAV (17.3%), wherein the majority (304, 87.1%) had
37 the lowest HAV concentration (2.0-2.7 gc/ μL , 1-5 partitions), followed by 20 samples (5.7%)
38 with concentrations of 2.8-3.0 gc/ μL (6-10 partitions). HAV was detected in 17.1%
39 (241/1170) of the Gauteng samples, and a 26.1% correlation between anti-HAV IgM in
40 clinical samples and HAV in wastewater samples was detected. We successfully developed
41 and optimized a dPCR method to detect and quantify HAV in wastewater samples, and
42 determined its LOQ. Further analysis is required to compare the wastewater data with clinical
43 surveillance data to facilitate appropriate interpretation of the results.

44

45 **1. Introduction**

46 Hepatitis A virus (HAV) is primarily transmitted via the fecal-oral route, and causes mild to
47 severe symptoms, including vomiting, diarrhea, fever, dark-colored urine, and jaundice, as
48 well as fulminant hepatitis [1–3]. Children under 5 years of age either present with mild
49 symptoms or are asymptomatic whereas adults present with mild to severe symptoms [1–3].
50 HAV incidence is usually higher in low-income countries because of poor sanitation,
51 contaminated water, and poor sewage disposal, with consequent contamination of food or
52 water [4–6]. Recent studies indicate a decline in HAV incidence in middle-to high-income
53 countries, and outbreaks are commonly associated with high-risk groups, including travelers
54 from highly endemic regions, men who have sex with men (MSM), the administration of
55 injectable drugs, and homeless people with poor sanitary habits [4–6]. In countries with high
56 HAV endemicity, infection occurs during early childhood, and immunity develops with
57 positivity for anti-HAV immunoglobulin (IgG) by adulthood. Unlike in countries with low
58 HAV endemicity, adults are susceptible to HAV and are often symptomatic [7]. However,
59 because HAV-infected children are usually asymptomatic and only seek medical treatment
60 and testing upon the presentation of symptoms, the prevalence of the disease may be
61 underestimated.

62 Low-income regions suffer from the most severe threat of hepatitis A, with the highest age-
63 standardized incidence rates of HAV in sub-Saharan Africa [4]. In South Africa, where HAV
64 disease contributes to approximately 0.3% of deaths annually, and the severity of the disease
65 increases with age [8], HAV infection is considered a notifiable medical condition (NMC) in
66 the surveillance system, based on passive surveillance. In this approach, reports on cases are
67 notified by clinicians and laboratories that test for the virus under the National Health
68 Laboratory Services (NHLS), and are used to determine the thresholds for public health
69 action [9]. However, the data from the private sector currently remain unreported. The

70 environmental presence of the HAV in low- and middle-income countries (LMICs) has been
71 closely studied [10]. Active surveillance of HAV in South Africa is imperative for the routine
72 monitoring of cluster influx through the implementation of public health interventions to
73 prevent outbreaks [9]. Studies in the Western Cape and Gauteng provinces reported <90%
74 HAV seropositivity in children aged 10 years, which indicates a transition from high to
75 intermediate endemicity in South Africa [11,12]. Although the World Health Organization
76 (WHO) recommends that HAV vaccination be integrated into national immunization
77 schedules for children in African populations where seroprevalence drops to <90% by age 10,
78 this approach it is not considered cost-effective for routine use in the public health sector
79 [13].

80 Wastewater surveillance has been implemented globally for the detection of viruses,
81 including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and poliovirus, in
82 conjunction with routine clinical surveillance for in-depth screening in the population of
83 interest [14, 15]. This is an effective early warning system for the detection of emerging and
84 circulating viruses, especially with the knowledge that younger children are asymptomatic
85 and will not seek healthcare services [16–18] As HAV is stable under diverse environmental
86 conditions, including high temperatures (up to 80°C) and a low pH range (1-2), its presence is
87 detectable in fresh and wastewater [2,19, 20]. In South Africa, HAV nucleic acids have been
88 detected in freshwater from selected rivers, dams, and wastewater [21, 22]. Wastewater
89 containing human excretions can enable public health authorities to identify HAV outbreak
90 hotspots, monitor currently circulating strains, study trends in several regions, and facilitate
91 early detection for more informed public health decision-making [23]. In particular, in low-
92 to middle-income countries with restricted access to appropriate resources, wastewater
93 surveillance can be used to monitor HAV incidence in asymptomatic and symptomatic
94 populations and trace the potential sources of outbreaks. However, the detection of HAV in

95 effluent samples was negligible over the sampling period [22] and remained undetectable in
96 the final effluent sampled immediately after chlorination [24].
97 Currently, there is no active HAV surveillance in South Africa, and data are available from
98 passive surveillance, wherein only symptomatic cases are recorded. Therefore, HAV
99 wastewater surveillance could provide added value to strengthen surveillance systems to
100 monitor trends in HAV incidence, identify outbreaks, and implement appropriate public
101 health responses and prevention measures. Similarly as for other picornaviruses, the 5'UTR is
102 commonly used for the detection of HAV because of its highly conserved nature, making it
103 less prone to nucleotide changes [25–27]. The 5'UTR is approximately 732 nucleotide bases
104 in length and comprises three pyrimidine-rich (80–90%) regions from positions 99–138, 204–
105 250, and 711–725 [27]. This constitutes a useful tool for detecting HAV in outbreak settings
106 [28].
107 South Africa is an upper-middle-income country with a population of 66 million individuals
108 within nine provinces that comprise 52 health districts. Approximately 80% of individuals
109 receive healthcare services, including laboratory diagnostic testing, from public sector clinics,
110 hospital facilities, and the National Health Laboratory Service (NHLS). The National
111 Institute for Communicable Diseases (NICD), a division of the NHLS, is responsible for
112 communicable disease surveillance through various programs, including the legally mandated
113 Notifiable Medical Condition Surveillance System (NMCSS). HAV seroprevalence has been
114 previously described [28]. Wastewater and environmental surveillance (WES) has been
115 conducted by the NICD since 2021, and commenced at national sentinel sites (two large
116 wastewater treatment plants in each province). However, limited data on the use of WES for
117 HAV monitoring have been reported in South Africa.
118 We aimed to optimize methods for HAV detection and quantification from wastewater
119 samples based on the 5' UTR region. We adopted a digital PCR method that offers sensitive

120 and absolute quantification without the need for a reference and has a greater tolerance to
121 inhibitors present in wastewater samples [29]. We applied the novel method to analyze the
122 samples collected as part of a national wastewater surveillance program, from three “case-
123 study” areas in Gauteng Province, each with 6 sampling points located within the catchment
124 area of a single large wastewater treatment site. Moreover, we compared the findings of the
125 WES data with clinical HAV surveillance data from health facilities located in the same
126 regions.

127

128 **2. Methods**

129 ***2.1. Optimization using real-time reverse-transcription PCR (RT-PCR) detection of HAV***

130 RNA was extracted using the QIAamp RNA Mini Extraction Kit (Qiagen, Hilden, Germany)
131 following the spin column method, according to the manufacturer’s instructions. A one-step
132 real-time RT-PCR assay was used to optimize the primers (HAV68F
133 TCACCGCCGTTTGCCTAG and HAV240R GGAGAGCCCTGGAAGAAAG) and probe
134 (HAV150Probe TTAATTCCTGCAGGTTTCAGG labelled with FAM-MGB) for the
135 detection of 5’UTR HAV [30]. The QuantiFast® Pathogen RT-PCR assay (Qiagen, Hilden,
136 Germany) was optimized by preparing the following reaction mix: 5 µL 5× QuantiFast
137 Pathogen Master Mix, 0.25 µL 100× QuantiFast Pathogen RT Mix, equal volumes of 10 µM
138 primers and probe were pooled together and 2.5 µL of the pool was used for the reaction mix,
139 and with 2.5 µL 10× internal control assay and internal control RNA and 8 µL RNA, made up
140 to a volume of 25 µL with nuclease-free water. Previously sequenced HAV-positive clinical
141 specimens (IgM-seropositive and successfully sequenced by Sanger sequencing) were used to
142 optimize the assay, which was performed on an ABI 7500 real-time PCR platform (Applied
143 Biosystems). The PCR cycling conditions were run according to the manufacturer’s

144 instructions, with a modification of the annealing and extension conditions (60°C for 1 min,
145 for 45 cycles).

146

147 ***2.2. Optimization using digital PCR (dPCR) from clinical specimens***

148 The QIAcuity OneStep Advanced Probe RT-dPCR assay (Catalog no. 250132, Qiagen,
149 Hilden, Germany) was performed using the QIAcuity Digital PCR system (Qiagen, Hilden,
150 Germany) according to the manufacturer's instructions, with the primers and probes used for
151 real-time PCR optimization and the previously extracted RNA. The assay was optimized by
152 preparing the following reaction mix: 3 µL 4× OneStep Advanced Probe Master Mix, 0.12
153 µL 100× OneStep Advanced RT Mix (Reverse Transcription), equal volumes of 20 µM
154 primers and probe pooled together, of which 0.6 µl was added to the reaction mix, 1.5 µL
155 enhancer GC, 5 µL RNA, and nuclease-free water to make up a total volume of 12 µL. The
156 reaction mix was added to an 8.5k (96 well) nanoplate and loaded into the digital PCR
157 instrument. The PCR cycling conditions were set as per the manufacturer's instructions, with
158 the following modifications: annealing and extension at 60°C for 1 min, for 45 cycles. The
159 master mix was evenly distributed through 8,500 partitions in each well. Nucleic acids in
160 each partition were quantified by imaging each well after the PCR. Results were interpreted
161 as copies/µL. Previously sequenced positive HAV clinical specimens (IgM seropositive and
162 successfully sequenced) were used as positive controls.

163

164 ***2.3. Limit of detection***

165 In accordance with the method described above, the limit of quantification (LOQ) of the
166 optimized assay was determined using clinical specimens that were previously confirmed as
167 IgM seropositive and were successfully sequenced. Two-fold dilutions were prepared from

168 1:2 to 1:1024, with 8 μ L RNA extracted from clinical samples used as the template. The
169 cycling conditions were identical to those described above.

170

171 ***2.4. Study site, sample collection, and processing of wastewater samples***

172 Next, 1 L grab samples of untreated raw wastewater were collected bi-weekly, weekly, or
173 monthly from sentinel and “case-study” sampling sites. Figure 1 shows the distribution of the
174 sites from which the wastewater samples were collected. The population in the study sites
175 ranged from 1200 to 950000 people. The samples included a combination of influent from
176 the wastewater treatment works (just after the screens) and street line manholes. Samples
177 from the catchment sites are collected between 9am and 11am. From all other sites samples
178 were collected anytime in the morning by trained plant operator. Samples were transported on
179 ice at 4°C to the NICD. Within 24 hours of receipt, the samples were concentrated by
180 ultrafiltration using Centricon® Plus-70 centrifugal filter units (MERCK, Germany), per the
181 manufacturer’s instructions (31). The concentrate was tested for SARS-CoV-2 in real time as
182 previously described (31). The residual concentrates were stored at –20°C. All experiments
183 were performed by the same person, using the same instrument, materials etc.

184

185 **Fig. 1 Map of South Africa displaying the wastewater sampling sites in each province.**

186 The green circles represent 29 national sampling sites and the purple circles represent 18
187 community sites within the Gauteng Province.

188

189 ***2.5. Total nucleic acid extraction from wastewater samples***

190 Nucleic acids were extracted from the residual concentrate using a MagMAX™ Wastewater
191 Ultra Nucleic Acid Isolation Kit (ThermoFisher Scientific, Massachusetts, United States) on

192 a KingFisher™ Flex instrument (ThermoFisher Scientific, Massachusetts, United States)
193 according to the manufacturer's instructions.

194

195 ***2.6. HAV detection in retained wastewater samples***

196 A duplex assay targeting hepatitis A was performed on all wastewater samples. For each
197 reaction, the mastermix comprised 3 µL 4× OneStep Advanced Probe Master Mix, 0.12 µL
198 100× OneStep Advanced RT Mix (Reverse Transcription), 0.6 µL 20 µM HAV
199 primer/probe-ROX, 1 µL enhancer GC, with the addition of 8 µL total nucleic acid. The
200 mastermix was briefly centrifuged and 13.82 µL was transferred to its respective well in the
201 8.5k 96-well dPCR nanoplate, and loaded into the dPCR instrument. The cycling conditions
202 were identical to those described above. The concentration output by the instrument,
203 presented in genome copies/µL, was converted to genome copies/mL using the following
204 formula:

$$205 \quad \frac{\text{Sample genome copies}}{\text{mL}} = \frac{\text{genome copies}}{\mu\text{L}} \times \text{reaction volume} \times \left(\frac{50}{8}\right) \times \frac{\left(\frac{1000}{200}\right)}{70}$$

206 The experiment included a negative control in which no target sample was detected.

207

208 ***2.7. Epidemiological analysis of wastewater and clinical HAV surveillance data***

209 As part of the NMC surveillance, the NICD analyzes and interprets laboratory results from
210 routine clinical hepatitis A serology testing conducted at the NHLS across South Africa, as
211 well as from private-sector laboratories. We used anonymized HAV NMC testing data from
212 July 2021 to November 30, 2023, and identified samples from health facilities located in
213 districts where wastewater surveillance testing was conducted. We grouped positive and
214 negative clinical results by epidemiological week and district of the health facility and tallied
215 the positive and negative results for each time-space combination. Similarly, we arranged

216 HAV-positive and HAV-negative wastewater samples according to the epidemiological week
217 and district of the sample collection site. We tallied the positive and negative results for each
218 time–space combination.

219 We merged the two datasets by epidemiological week and district and compared the
220 wastewater with clinical results. We classified wastewater: clinical pairs as “concordant-
221 positive” when at least one wastewater sample and at least one clinical sample tested positive
222 for HAV in that epidemiological week and in one district. Conversely, if no wastewater
223 samples and no clinical specimens tested positive for HAV in an epidemiological week and
224 district combination, we classified that week-district as ‘concordant-negative.’ The
225 discordant classification was assigned when, in specific week-district combinations, no
226 clinical specimen tested positive but HAV was detected in one or more wastewater samples
227 (discordant-wwpos) or when HAV was detected in at least one clinical specimen but all
228 wastewater samples tested negative for HAV (discordant-wwneg). Furthermore, to account
229 for the extended shedding duration of those infected with HAV both before and after the
230 epidemiological week of diagnosis (assumed to be the date of collection of the clinical
231 specimen), clinical testing from up to 2 weeks before and after the week in which a
232 wastewater sample was tested was considered.

233 Our rationale for this analytic approach is based on: 1) The epidemiology of HAV infection
234 in a high-endemicity setting, which suggests that when a single positive HAV case is
235 identified through passive surveillance, in all likelihood, there are other asymptomatic,
236 incubating, or undetected cases present in the district; and 2) On account of the inherent and
237 relative insensitivity of wastewater detection for viral RNA targets, a positive HAV result in
238 a wastewater sample indicates the presence of that pathogen in the catchment population, but
239 a negative result cannot confirm the absence of the pathogen; and 3) the need for
240 epidemiologically meaningful analytic approach, from which public health interventions

241 could potentially emerge; and 4) a district as a meaningful unit of analysis for public health
242 interventions.

243

244 **2.8. Statistical analysis**

245 Wastewater testing data were entered directly into RedCap [30, 31], and the data required for
246 hepatitis A were exported and analyzed using Excel. All graphs were generated using
247 Microsoft Excel. Clinical data from the Surveillance Data Warehouse (SDW) were analyzed
248 using the R programming language (insert version of R used). ArcGIS was used to generate a
249 map of South Africa indicating the proportion of wastewater samples that tested positive for
250 HAV and the proportion of anti-HAV IgM from the clinical data. The anti-HAV IgM data
251 from SDW were compared with wastewater data by analyzing the data of the corresponding
252 epidemiological week and district (referred to as epidemiological week-district pairs) based
253 on where the wastewater testing was conducted. A positive epidemiological week–district
254 pair was defined as the presence of at least one wastewater or clinical sample that tested
255 positive for HAV.

256

257 **2.9. Ethics**

258 The NICD conducts all routine clinical surveillance, including NMC surveillance, in a
259 protocol reviewed and approved by the University of the Witwatersrand Human Research
260 Ethics Committee (HREC) M210752 and under the legal authority of the National Health Act
261 (no. 61 of 2003). Wastewater sampling and testing were performed using an HREC system
262 (M230828). Informed consent was not required as we only used anonymized data.

263

264 **3. Results**

265 **3.1. Development of HAV detection and quantification assays for wastewater samples**

266 Real-time PCR and dPCR for the detection of the HAV 5' UTR was successfully optimized
267 (Fig. 2). The number of partitions and concentration of genome copies (gc/mL) were
268 compared to determine the LOQ, which was verified as one partition, with a Ct-value of 38.5,
269 and a concentration of 0.4 gc/ μ L (Fig. 3). When comparing the Ct value to the concentration
270 and the Ct value to the number of partitions (Fig. 3), the R² value was 0.967, which indicated
271 96% reliability.

272

273 **Fig. 2 Optimization of HAV using dPCR.** Examples of results from anti-HAV IgM-positive
274 clinical samples that were used for optimization and verification. The red line represents the
275 threshold, the gray area/dots under the threshold represent the negative partitions whereas
276 blue dots above the threshold represent the positive partitions.

277

278 **Fig. 3 Limit of quantification comparing the number of partitions and concentration to**
279 **ct-values.** Two-fold dilutions were performed and results were plotted as an exponential
280 curve. The orange points indicates the number of positive partitions and the blue points
281 indicate the concentrations (genome copies) that correspond with the ct-values based on HAV
282 detected for each dilution.

283

284 ***3.2. HAV detection in wastewater concentrates from 2021 to 2024***

285 In total, 2013 wastewater samples were tested for HAV, of which 17.3% (349/2013) tested
286 positive (Table 1). The majority of the positive samples had 1–5 partitions on dPCR (87.1%;
287 304/349) which was equivalent to 2.0–2.7 log gc/mL, followed by 6–10 partitions (5.7%;
288 20/349) which was equivalent to 2.8–3.0 log gc/mL (Fig. 4).

289

290 **Table 1. Results of wastewater samples tested for hepatitis A, from 2021 to 2023**

Results	Number of samples (N=2013)	Number of positive partitions on dPCR	Genome copies/ml	Equivalent log-transformed genome copies/μl
Negative	1664 (82.7%)	0	0	0
Positive	349 (17.3%)	1-306	109-37410	2.0 - 4.5

291

292 **Fig. 4 Genome copies/mL for samples with specified numbers of positive dPCR**
 293 **partitions, for all wastewater samples that tested positive for HAV.** The legend displays
 294 the range of partitions for each of the box and whiskers plotted and in brackets is the number
 295 of samples that had a specific range of partitions.

296

297 HAV was detected in wastewater samples from several provinces of South Africa (Table 2).
 298 The majority of the samples tested were from Gauteng Province (70.1%; 1411/2013), with a
 299 positivity rate of 17.1%. When determining the positivity rates within each province,
 300 Limpopo (33.3%, 4/12) and the Western Cape (32.4%, 22/68) showed the highest
 301 intraprovince positivity.

302

303 **Table 2 Wastewater samples tested for HAV across provinces in South Africa, from**
 304 **2021 to 2024.**

Provinces/Districts	Period	Total tested	Positive	Positivity rate (%)
Eastern Cape	Apr 2021-Mar 2024	173	26	15,0
Free State	Aug 2021-Mar 2024	169	32	18,9

Gauteng	Feb 2021-Mar 2024	1411	241	17,1
KwaZulu-Natal	May 2021-Mar 2024	135	15	11,1
Limpopo	Jan 2024-Mar 2024	12	4	33,3
Mpumalanga	Jan 2024-Mar 2024	10	2	20,0
North West	Oct 2023-Mar 2024	21	5	23,8
Northern Cape	Aug 2021-May 2022	14	2	14,3
Western Cape	Apr 2021-Mar 2024	68	22	32,4
All provinces	Feb 2021-Mar 2024	2013	349	17,3

305

306 We further compared the HAV data from wastewater to clinical samples (Fig. 5). Overall, the
 307 proportion of HAV-positive wastewater samples was higher than that of anti-HAV IgM-
 308 positive samples. Western Cape province (5.1-9.0%) had the highest proportion of anti-HAV
 309 IgM positives. Although fewer wastewater sites were included in this province, the
 310 proportion of HAV-positive samples was >21%. In Gauteng, the proportion of anti-HAV
 311 IgM-positive cases was greatest in the City of Tshwane and Ekurhuleni districts (<1.0%),
 312 followed by the City of Johannesburg (1.1-2.0%). The proportion of HAV that tested positive
 313 in wastewater was the highest in the cities of Johannesburg and Ekurhuleni districts (>13%).

314

315 **Fig. 5 Proportion of HAV positive across South Africa, clinical vs wastewater.** Brown
 316 circles represent the proportion of samples that tested positive for HAV, the larger the circle
 317 the higher the positivity. The shades of blue for each province (and districts within the
 318 province) represent the proportion of patients that were anti-HAV IgM positive: the darker

319 the area is, the higher is the proportion of positivity. The figure on the right is a projection of
320 the proportion positive from wastewater and clinical data in the Gauteng province.

321

322 ***3.3. Correspondence with clinical HAV surveillance data***

323 Between epidemiological Week 7, 2021, and Week 9, 2024, 442 628 clinical samples were
324 tested for anti-HAV IgM by NHLS laboratories. Two percent (9 098/ 442 628) of the samples
325 were IgM-positive. During the same period, 17.3% (349/ 2013) of wastewater samples tested
326 positive for HAV.

327 A total of 643 epidemiological week–district pairs were identified for clinical and wastewater
328 samples in the districts where wastewater testing was conducted (Table 3). Among these,
329 26.1% (168/643) had HAV detected in at least one wastewater sample and one clinical
330 sample (positive concordance). In 20.8% (134/643) of the epidemiological week–district
331 pairs, no wastewater samples or clinical specimens tested positive for HAV (negative
332 concordance). Wastewater week–district pairs were infrequently positive when clinical
333 samples from the same week and district were negative (4.4%; 28/643; discordant wastewater
334 positive). However, at least one clinical case of HAV was reported in 48.7% (313/643) of the
335 districts during an epidemiological week when no wastewater samples tested positive for
336 HAV (discordant wastewater negative; Table 3). When the definition of a positive week–
337 district pair was expanded to include clinical cases reported in the epidemiological weeks
338 before and after the wastewater sample testing, the number of concordant pairs increased
339 (Table 4).

340

341 **Table 3 HAV results from clinical and wastewater samples grouped by corresponding**
342 **epidemiological week: districts from epidemiological Week 7, 2021, to epidemiological**
343 **Week 9, 2024.**

HAV results from wastewater samples,				
classified by epidemiological week-district				
pairs				
HAV results from clinical		Positive	Negative	Total
samples, classified by	Positive	168 (26.1%)	313 (48.7%)	481
epidemiological week–district	Negative	28 (4.4%)	134 (20.8%)	162
pairs	Total	196	447	643*

344 * The denominator is the total number of epi-week district pairs.

345

346 **Table 4 HAV results from clinical and wastewater samples grouped by corresponding**

347 **epidemiological week: districts from epidemiological Week 7, 2021, to epidemiological**

348 **Week 9, 2024.**

Same Epidemiological Week + Clinical Results 1 Week After and 2 Weeks After				
Wastewater Results				
		Positive	Negative	Total
Clinical Results (IgM)	Positive (IgM)	191 (29.7%)	402 (62.5%)	593
	Negative (IgM)	5 (0.8%)	45 (7%)	50
	Total	196	447	643
Same Epidemiological Week + Clinical Results 1 Week Before and 2 Weeks Before				
Wastewater Results				
		Positive	Negative	Total
Clinical Results (IgM)	Positive (IgM)	190 (29.5%)	408 (63.5%)	598
	Negative (IgM)	6 (0.9%)	39 (6.1%)	45

Total	196	447	643
-------	-----	-----	-----

349 With a positive week–district definition expanded to include clinical cases that were reported
350 in epidemiological weeks before and after wastewater sample testing. *A positive
351 epidemiological week–district pair is defined as the presence of at least one wastewater
352 sample that tested positive in a given epidemiological week, or at least one clinical sample
353 that tested positive for HAV 1 or 2 weeks before or after the wastewater sample tested
354 positive, to accommodate for prolonged shedding (of up to 4 weeks) in HAV-infected
355 individuals.

356

357 **4. Discussion**

358 HAV is shed for up to 1–3 weeks through the stool of infected individuals and is highly
359 transmissible in the first 2 weeks before symptom onset, during which the highest viral
360 concentration is attained in the stool [14]. HAV detection in wastewater samples can enhance
361 our ability to monitor largely infected populations by collecting samples from a limited
362 number of sites, including wastewater treatment plants, because waste from multiple
363 households, clinics, and hospitals passes through a single site [18]. The incidence of HAV
364 has been decreasing, and the average age of infected individuals is increasing globally, with
365 most infections still prevalent in younger children [6]. Young children infected with the virus
366 are often asymptomatic, do not seek healthcare, and are not reported [5]. We showed that a
367 wastewater-based surveillance approach can enable the identification of HAV that goes
368 undetected and thus provide an early-detection and warning system.

369 We adopted an ultrafiltration technique to concentrate and enrich the wastewater samples.

370 Although relatively expensive, this approach requires no preconditioning steps and has been
371 previously used to detect HAV in Africa [32]. A systematic review of the global prevalence
372 of HAV in wastewater from 1986 to 2020 reported that the overall HAV prevalence was

373 31.4% in untreated wastewater, with viral concentrations of up to 3.7×10^{10} gc/L, which was
374 higher than concentrations detected in other water sources [33]. In South Africa, HAV was
375 detected in 43.1% of samples collected from various sites along the Buffalo River [21].
376 Another study described high levels of HAV detected at concentrations of 1.34×10^5 to $3.7 \times$
377 10^{10} gc/L in sewage with an 80% detection rate in five wastewater treatment plants in South
378 Africa during 2015 and 2016 [34]. The prevalence of HAV in South African-treated sewage
379 and surface water samples ranges from 4% to 37% and 16% to 76%, with an estimated
380 average detection rate of 15% (95% CI 1–29) and 51% (95% CI 21–80), respectively [10].
381 Detecting HAV in water matrices involves various detection methods, including the most
382 frequently used conventional RT-PCR (54.5%), followed by real-time PCR (37.5%) [33].
383 dPCR offers greater precision in comparison to qPCR and is far simpler to use for copy-
384 number quantification owing to its binary nature wherein the partitions are counted as
385 positive or negative [35]. To detect and quantify HAV from wastewater samples, we
386 successfully optimized a dPCR method that was developed using previously published
387 primers to detect and quantify HAV from wastewater samples. Similarly, although our
388 wastewater samples had lower HAV concentrations, HAV was detected in the untreated
389 water sources. This could be attributed to the time of viral shedding or dilution of the viral
390 concentration diluted during flow from the tributaries to the wastewater site. Our results
391 showed that the detected viral concentrations of HAV are a good representation of the
392 epidemiological distribution of HAV infection in the surrounding population, where the virus
393 is excreted from infected individuals .
394 Takuissu et al. [2023], in a systematic review, showed that low-income economies had the
395 highest HAV prevalence (29.0%), and Africa had the highest HAV prevalence. Varying
396 proportions of HAV have been detected in wastewater samples collected from different
397 provinces of South Africa. In our study, although the number of wastewater samples

398 collected varied across the different provinces in South Africa, in the samples collected from
399 the Eastern Cape, Free-State, Gauteng, Kwa-Zulu Natal and Western Cape provinces the
400 proportion of HAV in wastewater samples was generally higher compared to the proportion
401 of anti-HAV IgM reported from clinical data. The primary wastewater sites were located in
402 Gauteng Province, in the city of Tshwane, Johannesburg, and Ekurhuleni districts. In these
403 districts, the proportion of HAV in wastewater samples was >10 times greater than the
404 proportion of anti-HAV IgM from clinical samples. Our findings suggested that a large
405 proportion of HAV clinical cases were undetected, which could be attributable to
406 asymptomatic HAV-infected patients who would not have sought healthcare services [18].
407 Further comparison of the clinical data to wastewater data by epidemiological week and
408 district showed some concordance (26.1%) between the positive HAV results. However,
409 <50% of the samples showed discordant results on comparing HAV-positive results from
410 clinical samples to HAV-negative results from wastewater samples. This finding could be a
411 result of the larger number of clinics and hospitals tested across South Africa during the study
412 period compared with the number of sites tested for wastewater.

413

414 ***4.1. Limitations***

415 First, a limitation of this study was the low number of samples collected from wastewater
416 sites in provinces other than Gauteng Province, which made it difficult to determine the true
417 prevalence of HAV in South Africa from wastewater samples. Second, the number of sites
418 with clinical cases was much more extensive than that of wastewater sites. Our study
419 demonstrated the feasibility of using wastewater to complement clinical surveillance in South
420 Africa. Further comparisons with clinical data are required to establish the significance of
421 HAV detection in wastewater. Third, the enrichment method may have contributed to lower
422 viral concentrations, and alternative methods should be tested to improve viral concentration.

423 Lastly, the concentration of HAV detected in the wastewater samples may be lower than
424 expected because the samples were tested retrospectively, and freeze/thawing may have
425 resulted in RNA degradation and reduced viral loads. Prospective testing of the samples will
426 eliminate this factor and enable the detection of accurate virus concentrations.

427

428 **5. Conclusion**

429 Low- to middle-income countries with restricted access to appropriate resources can benefit
430 from wastewater surveillance for monitoring HAV incidence in asymptomatic and
431 symptomatic populations and identifying outbreaks and potential sources of transmission. To
432 expand on our current data, future work is required to compare wastewater data with clinical
433 surveillance data to facilitate the appropriate interpretation of the results; next-generation
434 sequencing should be performed to determine the epidemiological significance of HAV in
435 South Africa.

436

437 **Acknowledgments**

438 Editorial support, in the form of medical writing, assembling tables and creating high-
439 resolution images based on authors' detailed directions, collating author comments,
440 copyediting, fact checking, and referencing, was provided by Editage, Cactus
441 Communications.

442

443 **References**

444 1. World Health Organization. https://www.who.int/health-topics/hepatitis#tab=tab_2
445 (accessed 29 May 2023). 2023. Hepatitis.

- 446 2. Nainan OV, Xia G, Vaughan G, Margolis HS. Diagnosis of hepatitis a virus infection: a
447 molecular approach. *Clin Microbiol Rev.* 2006;19: 63-79. doi:10.1128/CMR.19.1.63-
448 79.2006.
- 449 3. Shin EC, Jeong SH. Natural history, clinical manifestations, and pathogenesis of
450 hepatitis A. *Cold Spring Harb Perspect Med.* 2018;8: a031708.
451 doi:10.1101/cshperspect.a031708.
- 452 4. Cao G, Jing W, Liu J, Liu M. The global trends and regional differences in incidence
453 and mortality of hepatitis A from 1990 to 2019 and implications for its prevention.
454 *Hepatology Int.* 2021;15: 1068-1082. doi:10.1007/s12072-021-10232-4.
- 455 5. Abutaleb A, Kottiril S. Hepatitis A: Epidemiology, natural history, unusual clinical
456 manifestations, and prevention. *Gastroenterol Clin North Am.* 2020;49: 191-199.
457 doi:10.1016/j.gtc.2020.01.002.
- 458 6. Jacobsen KH. Globalization and the changing epidemiology of hepatitis A virus. *Cold*
459 *Spring Harb Perspect Med.* 2018;8: a031716. doi:10.1101/cshperspect.a031716.
- 460 7. NICD. Communicable Diseases Communique. January 2023, vol.22(1). Vol. 22. 2023.
- 461 8. South African Department of Health. 2019. Hepatitis A. National guidelines for the
462 management of viral hepatitis.
463 [https://knowledgehub.health.gov.za/system/files/elibdownloads/2023-](https://knowledgehub.health.gov.za/system/files/elibdownloads/2023-04/SA%2520NDOH_Viral%2520Hepatitis%2520guidelines%2520final.pdf)
464 [04/SA%2520NDOH_Viral%2520Hepatitis%2520guidelines%2520final.pdf](https://knowledgehub.health.gov.za/system/files/elibdownloads/2023-04/SA%2520NDOH_Viral%2520Hepatitis%2520guidelines%2520final.pdf).
465 (Accessed{24/09/2024}).
- 466 9. Prabdi-Sing N, Motaze V, Manamela J, McCarthy K, Suchard M. Establishment of
467 Outbreak Thresholds for Hepatitis A in South Africa Using Laboratory Surveillance,
468 2017-2020. *Viruses.* 2021;13: 2470. doi:10.3390/v13122470.
- 469 10. Kuodi P, Patterson J, Silal S, Hussey GD, Kagina BM. Characterisation of the
470 environmental presence of hepatitis A virus in low-income and middle-income

- 471 countries: a systematic review and meta-analysis. *BMJ Open*. 2020;10: e036407.
472 doi:10.1136/bmjopen-2019-036407.
- 473 11. Enoch A, Hardie DR, Hussey GD, Kagina BM. Hepatitis A seroprevalence in Western
474 Cape Province, South Africa: Are we in epidemiological transition. *S Afr Med J*.
475 2019;109: 314-318. doi:10.7196/SAMJ.2019.v109i5.13410.
- 476 12. du Plessis NM, Haeri Mazanderani A, Motaze NV, Ngobese M, Avenant T. Hepatitis A
477 virus seroprevalence among children and adolescents in a high-burden HIV setting in
478 urban South Africa. *Sci Rep*. 2022;12: 20688. doi:10.1038/s41598-022-25064-x.
- 479 13. Haeri Mazanderani A, Motaze NV, McCarthy K, Suchard M, du Plessis NM. Hepatitis
480 A virus seroprevalence in South Africa - Estimates using routine laboratory data, 2005-
481 2015. *PLoS One*. 2019;14: e0216033. doi:10.1371/journal.pone.0216033.
- 482 14. EMRO WHO. WHO. Polio Eradication Initiative, Surveillance.
- 483 15. World Health Organization [Internet]. 2022. WHO. Available from:
484 <https://www.who.int/health-topics/coronavirus>
- 485 16. Rachida S, Matsapola PN, Wolfaardt M, Taylor MB. Genetic characterization of a novel
486 hepatitis a virus strain in irrigation water in South Africa. *J Med Virol*. 2016;88: 734-
487 737. doi:10.1002/jmv.24370.
- 488 17. Street R, Malema S, Mahlangeni N, Mathee A. Wastewater surveillance for Covid-19:
489 An African perspective. *Science of the Total Environment*. 2020;743: 1-3. doi:
490 10.1016/j.scitotenv.2020.140719
- 491 18. Hellmér M, Paxéus N, Magnus L, Enache L, Arnholm B, Johansson A, et al. Detection
492 of pathogenic viruses in sewage provided early warnings of hepatitis A virus and
493 norovirus outbreaks. *Appl Environ Microbiol*. 2014;80: 6771-6781.
494 doi:10.1128/AEM.01981-14.

- 495 19. Sattar SA, Jason T, Bidawid S, Farber J. Foodborne spread of hepatitis A: Recent
496 studies on virus survival, transfer and inactivation. *Can J Infect Dis*. 2000;11: 159-163.
497 doi:10.1155/2000/805156.
- 498 20. Vaughan G, Goncalves Rossi LM, Forbi JC, de Paula VS, Purdy MA, Xia G, et al.
499 Hepatitis A virus: host interactions, molecular epidemiology and evolution. *Infect Genet*
500 *Evol*. 2014;21: 227-243. doi:10.1016/j.meegid.2013.10.023.
- 501 21. Chigor VN, Okoh AI. Quantitative RT-PCR detection of hepatitis A virus, rotaviruses
502 and enteroviruses in the Buffalo River and source water dams in the Eastern Cape
503 Province of South Africa. *Int J Environ Res Public Health*. 2012;9: 4017-4032.
504 doi:10.3390/ijerph9114017.
- 505 22. Adefisoye MA, Nwodo UU, Green E, Okoh AI. Quantitative PCR detection and
506 characterisation of human adenovirus, rotavirus and hepatitis A virus in discharged
507 effluents of two wastewater treatment facilities in the Eastern Cape, South Africa. *Food*
508 *Environ Virol*. 2016;8: 262-274. doi:10.1007/s12560-016-9246-4.
- 509 23. World Health Organization. *World Health Statistics 2020*. 2022. 1-77 p.
- 510 24. Osulale O, Okoh A. Incidence of human adenoviruses and Hepatitis A virus in the final
511 effluent of selected wastewater treatment plants in Eastern Cape Province, South Africa.
512 *Virol J*. 2015;12: 98. doi:10.1186/s12985-015-0327-z.
- 513 25. CDC. *National Wastewater Surveillance System (NWSS)*. 2020.
- 514 26. Silveira Carneiro J, Equestre M, Pagnotti P, Gradi A, Sonenberg N, Perez Bercoff R. 5'
515 UTR of hepatitis A virus RNA: mutations in the 5'-most pyrimidine-rich tract reduce its
516 ability to direct internal initiation of translation. *J Gen Virol*. 1995;76 (Pt 5): 1189-1196.
517 doi:10.1099/0022-1317-76-5-1189.
- 518 27. Liu GD, Hu NZ, Hu YZ. Full-length genome of wild-type hepatitis A virus (DL3)
519 isolated in China. *World J Gastroenterol*. 2003;9: 499-504. doi:10.3748/wjg.v9.i3.499.

- 520 28. Wu S, Nakamoto S, Kanda T, Jiang X, Nakamura M, Miyamura T, et al. Ultra-deep
521 sequencing analysis of the hepatitis A virus 5'-untranslated region among cases of the
522 same outbreak from a single source. *Int J Med Sci.* 2014;11: 60-64.
523 doi:10.7150/ijms.7728.
- 524 29. Prajapati B, Rathore D, Joshi C, Joshi M. Digital PCR: A partitioning-based application
525 for detection and surveillance of SARS-CoV-2 from sewage samples. *Methods Mol*
526 *Biol.* 2023;2967: 1-16. doi:10.1007/978-1-0716-3358-8_1.
- 527 30. Costafreda MI, Bosch A, Pintó RM. Development, evaluation, and standardization of a
528 real-time TaqMan reverse transcription-PCR assay for quantification of hepatitis A virus
529 in clinical and shellfish samples. *Appl Environ Microbiol.* 2006;72: 3846-3855.
530 doi:10.1128/AEM.02660-05.
- 531 31. Iwu-Jaja C, Ndlovu NL, Rachida S, Yousif M, Taukobong S, Macheke M, et al. The role
532 of wastewater-based epidemiology for SARS-CoV-2 in developing countries:
533 Cumulative evidence from South Africa supports sentinel site surveillance to guide
534 public health decision-making. *Sci Total Environ.* 2023;903: 165817.
535 doi:10.1016/j.scitotenv.2023.165817.
- 536 32. Osunmakinde CO, Selvarajan R, Sibanda T, Mamba BB, Msagati TAM. Overview of
537 trends in the application of metagenomic techniques in the analysis of human enteric
538 viral diversity in Africa's environmental regimes. *Viruses.* 2018;10: 429.
539 doi:10.3390/v10080429.
- 540 33. Takuissu GR, Kenmoe S, Ebogo-Belobo JT, Kengne-Ndé C, Mbaga DS, Bowo-Ngandji
541 A, et al. Occurrence of hepatitis A virus in water matrices: a systematic review and
542 meta-analysis. *Int J Environ Res Public Health.* 2023;20: 1054.
543 doi:10.3390/ijerph20021054.

- 544 34. Rachida S, Taylor MB. Potentially infectious novel hepatitis A virus strains detected in
545 selected treated wastewater discharge sources, South Africa. *Viruses*. 2020;12: 1468.
546 doi:10.3390/v12121468.
- 547 35. Whale AS, Huggett JF, Cowen S, Speirs V, Shaw J, Ellison S, et al. Comparison of
548 microfluidic digital PCR and conventional quantitative PCR for measuring copy number
549 variation. *Nucleic Acids Res*. 2012;40: e82. doi:10.1093/nar/gks203.

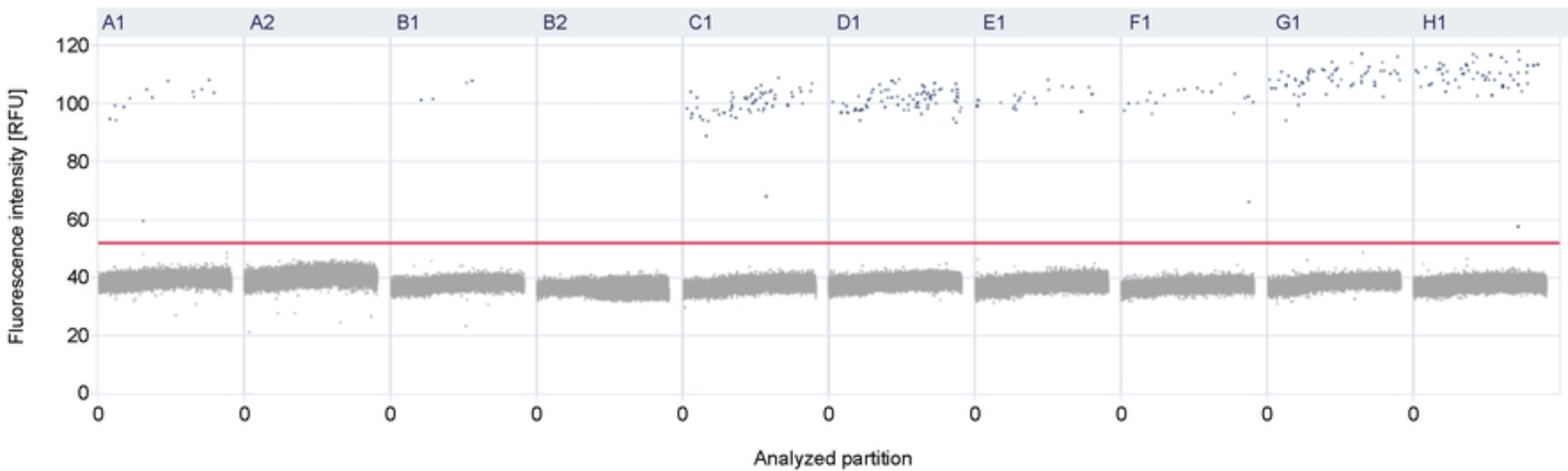


Figure 2

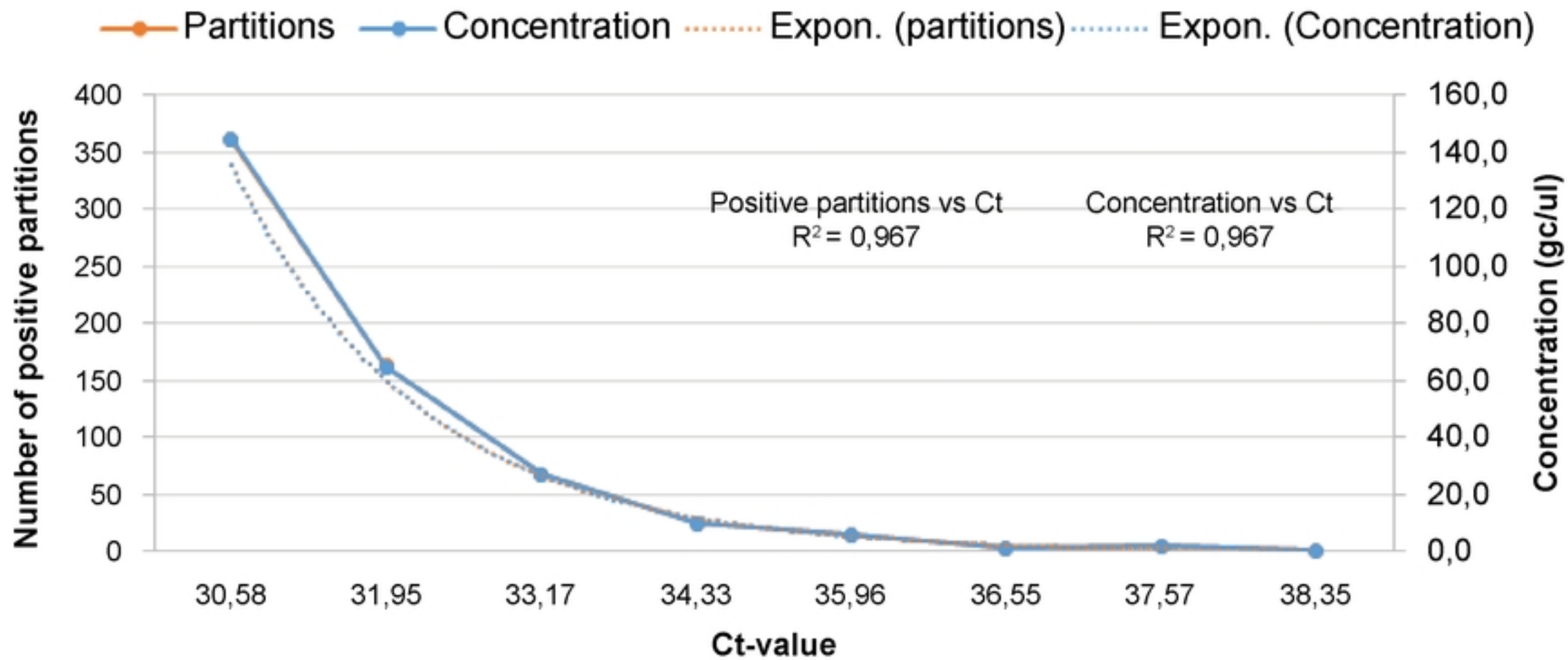


Figure 3

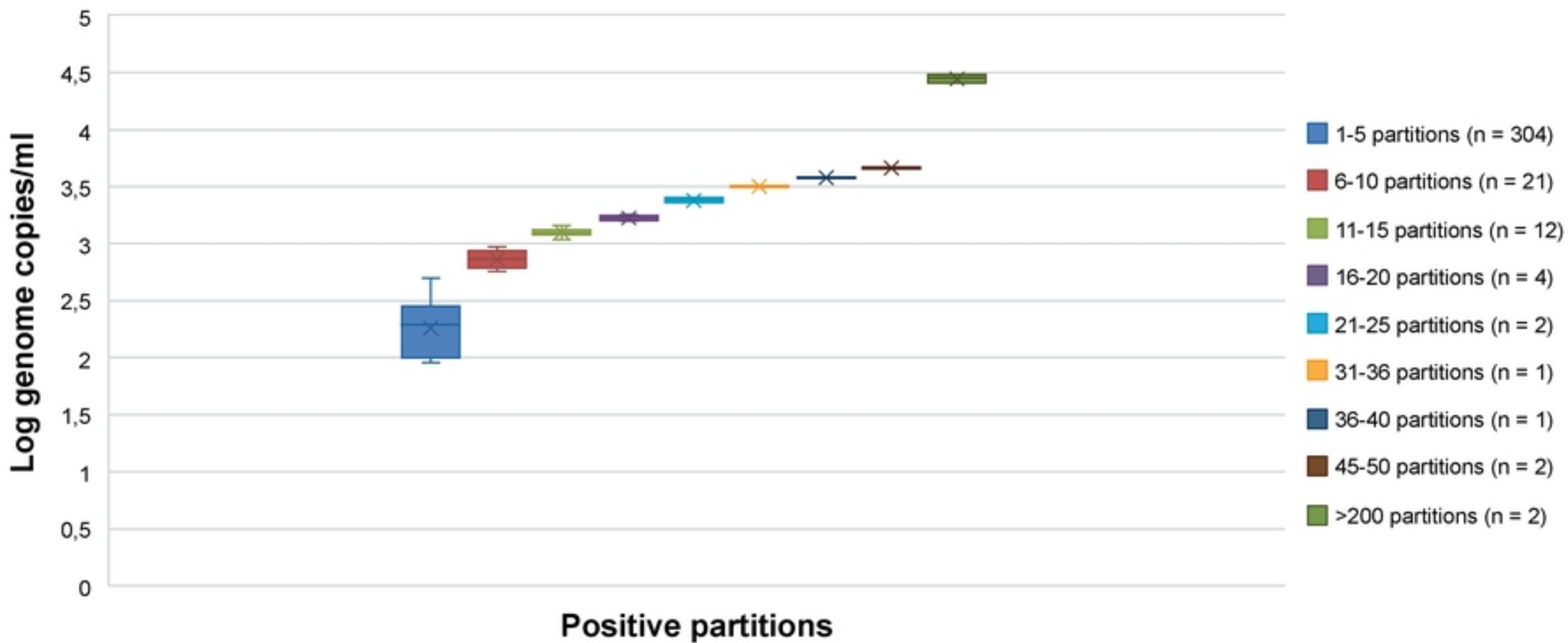
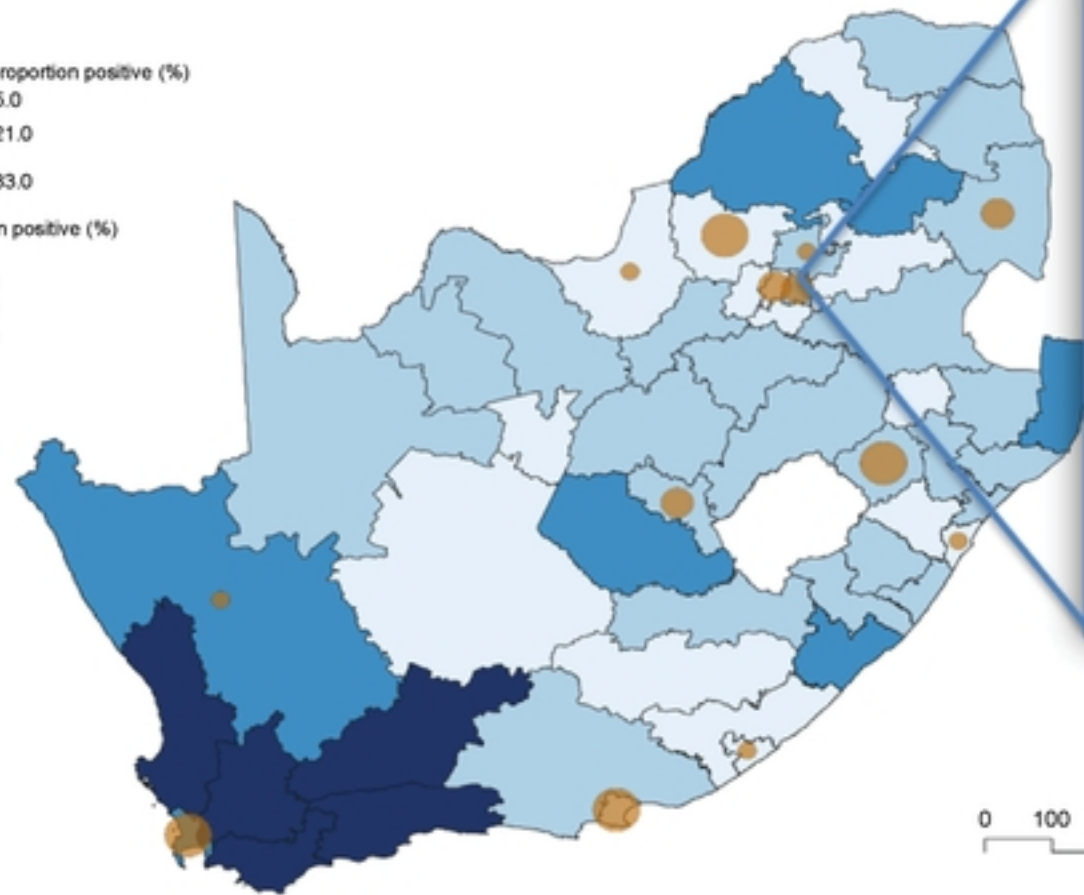


Figure 4

Wastewater proportion positive (%)



IgM proportion positive (%)



IgM proportion positive (%)



IgM proportion positive (%)

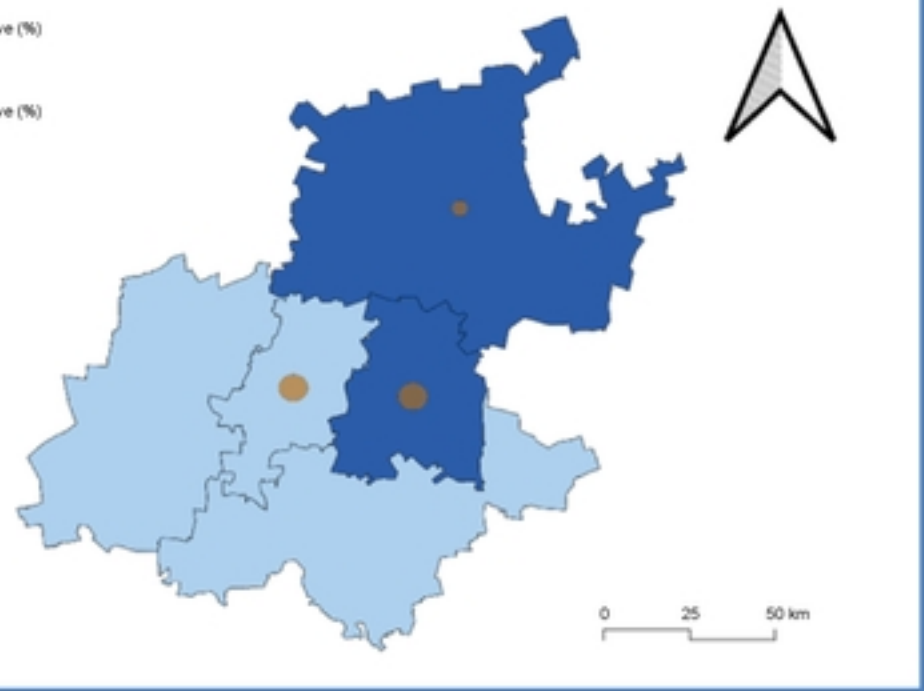


Figure 5