

1 **Genetic epidemiology of SARS-CoV-2 transmission in renal dialysis units - a**
2 **high risk community-hospital interface**

3

4 Kathy K Li^{1*}, Y. Mun Woo^{2*}, Oliver Stirrup³, Joseph Hughes¹, Antonia Ho¹, Ana Da
5 Silva Filipe¹, Natasha Johnson¹, Katherine Smollett¹, Daniel Mair¹, Stephen
6 Carmichael¹, Lily Tong¹, Jenna Nichols¹, Elihu Aranday-Cortes¹, Kirstyn Brunker¹,
7 Yasmin A. Parr¹, Kyriaki Nomikou¹, Sarah E McDonald¹, Marc Niebel¹, Patawee
8 Asamaphan¹, Vattipally B Sreenu¹, David L Robertson¹, Aislynn Taggart¹, Natasha
9 Jesudason¹, Rajiv Shah¹, James Shepherd¹, Josh Singer¹, Alison H.M. Taylor⁴, Zoe
10 Cousland⁴, Jonathan Price⁴, Jennifer S. Lees^{4,5}, Timothy P.W. Jones⁶, Carlos Varon
11 Lopez⁷, Alasdair MacLean⁸, Igor Starinskij⁸, Rory Gunson⁸, Scott T.W. Morris², Peter
12 C. Thomson², Colin C. Geddes², Jamie P. Traynor², Judith Breuer⁹, The COVID-19
13 Genomics UK (COG-UK) consortium^{10*}, Emma C. Thomson^{1,11} †, Patrick B. Mark^{2,5} †

14

- 15 1. MRC-University of Glasgow Centre for Virus Research, Sir Michael Stoker
16 building, 464 Bearsden Road, Glasgow, G61 1QH
- 17 2. The Glasgow Renal & Transplant Unit, Queen Elizabeth University Hospital
18 University Hospital, 1345 Govan Road, Glasgow, G51 4TF, UK
- 19 3. Institute for Global Health, University College London, London
- 20 4. Renal Unit, University Hospital Monklands, Monks court Ave, Airdrie ML6 0JS
- 21 5. Institute of Cardiovascular and Medical Sciences, University of Glasgow, 126
22 University Place, Glasgow G12 8TA, UK.
- 23 6. Department of Infectious Diseases, University Hospital Monklands,
24 Monks court Ave, Airdrie ML60JS

25 7. Department of Microbiology, University Hospital Monklands, Monkscourt Ave,
26 Airdrie ML6 0JS

27 8. West of Scotland Specialist Virology Centre, Glasgow Royal Infirmary

28 9. Institute of Child Health University College London, Cruciform Building, Gower
29 Street, London, WC1E 6BT

30 10. <https://www.cogconsortium.uk>

31 11. Department of Clinical Research, London School of Hygiene and Tropical
32 Medicine

33

34 *equal contribution

35 †joint senior and corresponding authors

36 *Full list of consortium names and affiliations are listed in the appendix

37

38 Correspondence (clinical): Professor Patrick B. Mark, Patrick.Mark@glasgow.ac.uk

39 The Glasgow Renal & Transplant Unit, Queen Elizabeth University Hospital, 1345

40 Govan Road, Glasgow, G51 4TF, UK (genomics) Professor Emma Thomson

41 Emma.Thomson@glasgow.ac.uk MRC-University of Glasgow Centre for Virus

42 Research, Sir Michael Stoker building, 464 Bearsden Road, Glasgow, G61 1QH

43

44 **Abstract**

45 **Objectives:** Patients requiring haemodialysis are at increased risk of serious illness
46 with SARS-CoV-2 infection. To improve the understanding of transmission risks in
47 six Scottish renal dialysis units, we utilised the rapid whole-genome sequencing data
48 generated by the COG-UK consortium. **Methods:** We combined geographical,
49 temporal and genomic sequence data from the community and hospital to estimate
50 the probability of infection originating from within the dialysis unit, the hospital or the
51 community using Bayesian statistical modelling and compared these results to the
52 details of epidemiological investigations. **Results:** Of 671 patients, 60 (8.9%)
53 became infected with SARS-CoV-2, of whom 16 (27%) died. Within-unit and
54 community transmission were both evident and an instance of transmission from the
55 wider hospital setting was also demonstrated. **Conclusions:** Near-real-time SARS-
56 CoV-2 sequencing data can facilitate tailored infection prevention and control
57 measures, which can be targeted at reducing risk in these settings.

58 **Key words:** SARS-CoV-2, COVID-19, haemodialysis, renal dialysis unit, infection
59 control, rapid sequencing, outbreak, nosocomial

60

61 **Introduction**

62 The emergent SARS-CoV-2 virus which causes COVID-19 is associated with
63 increased morbidity and mortality in older individuals and in those with chronic
64 diseases.^{1,2} Chronic kidney disease and pre-existing conditions that may increase
65 the risk of renal failure, such as cardiovascular disease, diabetes mellitus and
66 hypertension, are significantly associated with an increased risk of death in COVID-
67 19.^{3,4} Individuals requiring haemodialysis in hospital are also at increased risk of
68 nosocomial infection due to prolonged outpatient dialysis (typically three times
69 weekly for four hours per session) and to community infection due to regular travel
70 on public or hospital transport, often with other patients.⁵⁻⁷ The case fatality rate in
71 dialysis patients has been reported as 20-30% compared with 1-2% in patients not
72 requiring haemodialysis.⁸⁻¹¹

73 Whole-genome pathogen sequencing has become increasingly accessible. Its utility
74 in the context of evolving outbreaks has been demonstrated with Ebola, Zika and
75 hospital outbreak investigations.¹²⁻¹⁵ The COVID-19 Genomics UK (COG-UK)
76 Consortium, funded by the UK Department of Health and Social Care, UK Research
77 and Innovation (UKRI) and the Wellcome Sanger Institute, was set up to enable real-
78 time sequencing at a population level and to facilitate the investigation and
79 management of hospital-associated infections, providing policymakers with
80 information on introductions and transmission events.¹⁶⁻¹⁸

81 We aimed to investigate the genetic epidemiology of COVID-19 infection from
82 patients attending six Scottish renal dialysis unit(s) (RDU) using a Bayesian
83 statistical analysis framework incorporating temporal, geographical and genetic
84 sequence data. These results were evaluated alongside traditional epidemiological
85 investigations. We further investigate the clinical impact of COVID-19 infection on

86 haemodialysis patients and incorporate the information obtained using combined
87 genetic and epidemiological data to inform future infection control strategies.

88 **Methods**

89 **Study design and participants**

90 The Glasgow Renal and Transplant Unit based at the Queen Elizabeth University
91 Hospital and University Hospital Monklands serve a combined population of
92 approximately 2.16 million people across West and Central Scotland under NHS
93 Greater Glasgow and Clyde, NHS Forth Valley and NHS Lanarkshire Health Boards.
94 These institutions provide haemodialysis treatment for 828 outpatients across eight
95 RDUs (numbers extracted 1st March 2020). Use of anonymised data was approved
96 by the Local Privacy Advisory Group of NHS Greater Glasgow and Clyde 'Safe
97 Haven' on behalf of the West of Scotland Ethics Committee (approval
98 GSH/20/RE/001). Follow up was until 4th June 2020. From 2nd March 2020, patients
99 attending for dialysis with symptoms of COVID-19 were tested for SARS-CoV-2 by
100 nasopharyngeal swab. We report data on the six RDUs (RDU 1-6, number of
101 patients treated with dialysis =671) with any patients with SARS-CoV-2 infection.
102 Initially, personal protective equipment (fluid-resistant surgical masks, eye protection,
103 aprons and gloves) (PPE) was *not* recommended by UK and Scottish Governments
104 for HCWs caring for patients, unless clinical index of suspicion was high for COVID-
105 19 or for aerosol generating procedures. RDU1 instigated PPE for HCWs and
106 surgical masks for patients whilst travelling to, from and during dialysis on 23rd March
107 in response to earlier infections. From 3rd April 2020, the UK-recommendations
108 changed to include any close patient contact. Surgical masks were also given to all
109 patients, as per RDU1 and shared-patient transport discontinued. Until 5th April
110 HCWs were ineligible for SARS-CoV-2 testing unless hospitalised, being advised to
111 self-isolate for 7 days in keeping with Scottish government policy.

112 **Laboratory diagnosis**

113 Nasopharyngeal swabs in viral PCR solution (1:1 ratio of EasyMag Nuclisens
114 Extraction Buffer (BioMerieux, France) and EMEM) were extracted and tested
115 according to the availability of assays at the diagnostic laboratory.ⁱ Surplus RNA
116 extract was collected for sequencing with ethical approval from the NHSGGC
117 biorepository (16/WS/0207NHS).

118 **Sequencing**

119 Sequencing was performed on either the *ONT MinION/GridION* or the *Illumina*
120 *MiSeq* platform, as previously described.¹⁶ Briefly, libraries were prepared in
121 accordance with the ARTIC network protocols (v1 and v2) ([https://artic.network/ncov-](https://artic.network/ncov-2019)
122 [2019](https://artic.network/ncov-2019)). For nanopore reads the ARTIC-nCov-2019 bioinformatics protocol was used,
123 reads were size filtered, demultiplexed, trimmed with Porechop
124 (<https://github.com/rwick/Porechop>) and mapped against the reference strain
125 Wuhan-Hu-1 (GenBank accession number MN908947), followed by clipping of
126 primer regions. Variants were called using Nanopolish 0.11.3
127 (<https://github.com/jts/nanopolish>). For Illumina, reads were trimmed with trim_galore
128 (bioinformatics.babraham.ac.uk/projects/trim_galore/) and mapped with BWA²⁰ to
129 the Wuhan-Hu-1 reference sequence, followed by primer trimming and consensus
130 calling with iVar.²¹ A read coverage of least 10 was used for the consensus.

131 **Sequence Data.** Consensus sequences with >90% coverage were included. All
132 consensus genomes are available from the GISAID database
133 (<https://www.gisaid.org>), the COG-UK consortium website
134 (<https://www.cogconsortium.uk/data/>) and BAM files from the European Nucleotide

ⁱ: MagNA Pure 96 system (Roche, Penzberg, Germany), Abbott M2000 (Abbott, Chicago, US) or Cobas[®] 6800 Systems (Roche). SARS-CoV-2 was detected using one of three RT-PCR assays: RdRp gene/E gene.¹⁹ RdRp gene/N gene (Abbott RealTime SARS-CoV-2 assay, Chicago, US) or the ORF1a/b and E gene (Cobas[®] SARS-CoV-2, Roche).

135 Archive's Sequence Read Archive service, BioProject PRJEB37886
136 (<https://www.ebi.ac.uk/ena/data/view/PRJEB37886>).

137 **Phylogenetic and probabilistic analysis**

138 Retrospective phylogenetic analysis of whole-genome sequences was performed as
139 follows. All full genomes of SARS-CoV-2 from Scotland sequenced as part of the
140 COG-UK consortium were included and aligned using MAFFT v7.310 and with the
141 HKY+I+G4 nucleotide substitution model determined by modeltest. The global
142 lineage and UK lineage assignments for the dialysis samples were determined using
143 civet (<https://github.com/artic-network/civet>).

144 We applied a novel algorithm, the Sequence Reporting Tool (SRT) to estimate the
145 probability of healthcare- vs community-acquired infection in each case, based on
146 the statistical approach developed for the COG-UK hospital-onset COVID-19
147 infection (HOI) study (<https://clinicaltrials.gov/ct2/show/NCT04405934>).²² The
148 approach is based on Bayesian principles and involves comparison of the proportion
149 of similar viral sequences (with maximum pairwise SNP difference of two, with no
150 difference where there is an overlap in ambiguous nucleotide codes or an 'N' in
151 either sequence) observed within potential locations of infection for the case of
152 interest: i) patients' RDU and elsewhere in this hospital, ii) inpatient ward and
153 hospital (if the patient was admitted) all within the prior three weeks; along with a
154 weighted proportion of similar sequences in the local community of the patient within
155 the prior six weeks based on the outer postcode of their home address. There are 61
156 districts based on this outcode (49 in Glasgow and 12 in Lanarkshire). We assumed
157 0.5 prior probability of infection within the RDU before consideration of sequence
158 data, and the prior probability of admission-related infection among the inpatients
159 was based on the interval from admission to diagnosis and the incubation distribution
160 of COVID-19.²³ The SRT algorithm outputs two posterior probabilities in all cases:

161 that of acquiring the virus directly from the RDU (p_{RDU}) and that of acquiring the
162 infection through use of facilities within the hospital but *not* in the RDU (e.g. toilets,
163 cafes, lobbies, or shared transport, etc) (p_{hRDU}). A posterior probability (p_{RDU})
164 of 1 indicates that the transmission occurred within the RDU; if the p_{RDU} stays at
165 0.5 then it remains unclear where the transmission occurred; if p_{RDU} is 0 then the
166 infection was most likely community acquired. If the haemodialysis patient was an
167 inpatient and continued to attend the RDU at the time of COVID-19 diagnosis, the
168 SRT algorithm also gives the probability of acquiring the infection from the ward of
169 admission (p_{wADM}) and that of infection elsewhere in that hospital (p_{hADM}). The
170 SRT algorithm was coded in R version 3.6.0, using the ape v5.3 package for
171 calculation of pairwise SNP differences and PostcodesioR v0.1.1 and gmt v2.0-1
172 packages to calculate distances between postcodes.²²

173 **Survival statistics**

174 Comparisons were made between patients who lived and died following SARS-CoV-
175 2 infection. At the time of analysis, no patients who were still alive were critically ill or
176 requiring oxygen therapy. The mortality rates were calculated for patients requiring
177 dialysis expressed as deaths per 1000 patient days were calculated over the three
178 months 1st March-31st May 2020.

179 **Results**

180 **Description of cases, treatments and outcomes in patients requiring** 181 **haemodialysis with COVID-19**

182 In total, 60 of 671 (8.9%) patients requiring HD were diagnosed with COVID-19
183 infection during 1st March-31st May 2020. 16/60 patients (26.7%) died; with COVID-
184 19 as the certified cause of death. There were no statistically significant differences
185 in the clinical characteristics and associated co-morbidities between those who died
186 and those who survived (Table 1). The median time from positive SARS-CoV-2 test

187 to death was 10.5 days (range 0-29 days). Two patients required intensive care (of
188 whom one died). No patients received 'specific' therapy for COVID-19 (e.g.
189 dexamethasone, remdesivir, tocilizumab). Compared to the corresponding three-
190 month periods 2018-19 (mean deaths 44/quarter year), there were 16 more deaths
191 in patients undergoing outpatient haemodialysis in the same RDUs, equivalent to
192 0.797 deaths per 1000 patient days in all patients requiring haemodialysis during 1st
193 March-31st May 2020 compared to 0.628 deaths per 1000 patient days as mean of
194 the corresponding period during 2017-2018 (equivalent to a 27.0% increase).

195 **Genomic and epidemiological investigation**

196 Residual RNA extract from 53 of 60 patients with SARS-CoV-2 positive samples
197 were obtained for virus genome sequencing. 39 of these sequences plus one from a
198 healthcare worker were of sufficient quality and coverage for further analysis. The
199 samples belonged to 13 different UK lineages (Figure 2). Whilst a number of patients
200 had indistinguishable sequences from the same UK lineage and shared dialysis
201 sessions, some fell outwith phylogenetic lineages, providing evidence of community-
202 transmission. The recent introduction of SARS-CoV-2 into the human population and
203 its relatively low mutation rate, mean sequences in the same UK lineage and
204 phylotype cannot be interpreted as direct transmission events, with further temporal
205 and epidemiological data required to quantify the probability of transmission.
206 Conversely, sequences from different UK lineages would disprove direct
207 transmission. Epidemiological investigation identified clusters of SARS-CoV-2
208 positive patients with shared dialysis sessions and sometimes transport and this was
209 analysed with the phylogenetic data (Figure 2 and 3). We found five of the six RDUs
210 spanning two health boards in the West of Scotland had evidence of unit-linked
211 transmission events.

212 **RDU1**

213 In RDU1, viral sequences from seven haemodialysis patients and one HCW from the
214 same unit clustered within the UK40 lineage (Figure 3). Five of these sequences
215 were indistinguishable to each other (CVR248, CVR284, CVR495, CVR987 and
216 CVR1404), suggestive of within-unit transmission. Applying the SRT, we found the
217 probability of within-unit transmission in RDU1 was indeterminate based on
218 sequence data alone ranging from 0.53 to 0.68. Further epidemiological analysis
219 suggested transmission in some cases – for example, CVR248 and CVR1404 had
220 indistinguishable sequences and shared dialysis sessions (Figure 3). However,
221 CVR284 and CVR495, also with indistinguishable sequences, did not overlap with
222 each other or anyone else on the unit, suggesting community-acquisition. HCW,
223 CVR987, having been in direct contact with CVR248, self-isolated on 22nd March,
224 prior to testing positive seven days later. However, this viral sublineage of UK40,
225 was widespread in the community, with 63 other indistinguishable sequences
226 detected within the geographical location of RDU1 and patient communities; so this
227 transmission could not reliably be inferred. Two of the seven patients whose
228 sequences derived from the wider UK40 lineage (CVR780 and CVR1404) (Figure 3)
229 were linked epidemiologically, sharing both dialysis sessions and transport from
230 home to the unit. However, the estimated probabilities of within unit transmission for
231 CVR780 and CVR1404 were 0.63 and 0.54, respectively. This result was supported
232 by close inspection of the data. CVR780 was found to have tested positive for
233 SARS-CoV-2 six days *prior* to CVR1404 and had a single nucleotide polymorphism
234 relative to the Wuhan reference not found in CVR1404, making transmission from
235 CVR780 to CVR1404 less likely. The widespread distribution of UK40 lineages in the
236 community, the early preventative measures implemented by RDU1 and the lack of
237 definitive epidemiological evidence for transmission suggest that individual
238 transmissions were not due to infection prevention and control (IPC) challenges in

239 this unit. The final case occurred nine days after “lockdown” and the implementation
240 of enhanced PPE measures.

241 **RDU2**

242 In RDU2, there was evidence of five introductions of SARS-CoV-2 from the
243 community, of which two lineages spread within the unit or on hospital transport to
244 the unit. CVR3289 and CVR3290 had indistinguishable sequences, only seen in two
245 other non-geographically linked community cases. These patients shared the same
246 dialysis session and transport, with an estimated 100% probability of within-unit
247 transmission (Figure 2, lineage UK658). Likewise, CVR1003 and CVR1924 had
248 indistinguishable sequences and shared the same dialysis sessions (Figure 3,
249 lineage UK51). CVR2314 had a p_{RDU} of 1, but this patient had no epidemiological
250 link to other two patients, suggesting nosocomial transmission from fomites or an
251 untested staff member (staff were not routinely tested for SARS-CoV-2, at this time).

252 **RDU3, 4 and 5**

253 In RDU3, three introductions and two separate transmission events were identified.
254 Although CVR375 and CVR1511 had indistinguishable sequences, this was shared
255 with 161 other Scottish samples (Figure 2, lineage UK5098). The SRT identified
256 nosocomial infection in CVR375 due to within-hospital rather than within-dialysis unit
257 transmission (p_{hADM} 0.95). CVR1511 acquisition of infection from RDU3 (p_{RDU}
258 0.68) was less clear. CVR1817 is a close sequence match to CVR375 and CVR1511
259 based on the 2 SNP threshold, leading the SRT to estimate probable unit-based
260 transmission ($p_{\text{RDU}} = 0.73$). However, on phylogenetic analysis CVR1817 falls into
261 the separate UK501 lineage (Figure 2) and appears likely to have been community-
262 acquired on consideration of all available information. In lineage UK39, CVR937 was
263 community-acquired while the related CVR1816 was probable within-unit
264 transmission ($p_{\text{RDU}} = 0.74$).

265 In RDU4, two patients tested positive for SARS-CoV-2 but there was no linkage
266 found on sequence analysis (Supplementary Table 1) with no evidence of within unit
267 transmission of SARS-CoV-2 in this dialysis unit.

268 Within RDU5, three cases of SARS-CoV-2 infection were detected. CVR1217
269 (lineage UK5098) was community-acquired while CVR1843 (p_RDU of 0.9) was
270 highly suggestive of within RDU transmission. These related sequences differed by 2
271 SNPs. This strongly suggests that intermediary modes of transmission should ideally
272 have been investigated, including untested asymptomatic individuals, for example
273 members of staff.

274 **RDU6**

275 SARS-CoV-2 was introduced to RDU6 on at least 5 occasions with evidence of
276 onward transmission in two cases and hospital-transmission in one. Of the 26
277 (12.3% of total) SARS-CoV-2 positive patients, 16 were sequenced. Nine of these
278 patients were within the UK429 lineage (Figure 3). The SRT verified a high likelihood
279 of within-unit transmission, with p_RDU ranging from 0.96 to 1 (Figure 3). CVR3373
280 (p_RDU = 0, the first of this phylotype found in RDU6) and CVR3362 were within a
281 separate lineage, UK5098 (Figure 3). CVR3362 had been hospitalised for a month
282 prior to the positive SARS-CoV-2 test, whilst maintaining dialysis within RDU6 and
283 had a p_RDU of 1. Further discussion with the infection control team confirmed that
284 CVR3379 (non-haemodialysis patient) and another unsequenced SARS-CoV-2
285 positive haemodialysis patient, shared the same hospital bay with CVR3362. It is
286 possible that SARS-CoV-2 was brought onto this bay by the other haemodialysis
287 patient from their dialysis sessions on RDU6. This highlights one of the limitations of
288 phylogeny and the SRT algorithm in an outbreak investigation – sequencing data
289 needs to be representative of prevalent cases for the results to be interpreted.

290 Finally, CVR3732 (Figure 2, UK370) had a high p_hRDU (0.96) indicating acquisition
291 of SARS-CoV-2 from elsewhere in the hospital.

292 **Summary**

293 We found evidence of multiple introductions of SARS-CoV-2 infection into Scottish
294 dialysis units and of onward transmission within these units. There was strong
295 evidence for 15 patients acquiring SARS-CoV-2 in hospital or on shared hospital
296 transport. For a further 9 patients the source of infection was less certain although
297 most likely acquired within the hospital RDU setting (In lineage UK40: CVR284,
298 CVR495, CVR780, CVR1204, CVR1314, CVR1404 were similar but multiple
299 indistinguishable sequences were also detected in the community; lineage UK5098:
300 CVR1511; lineage UK39: CVR1816; lineage UK501: CVR1817). A further 15
301 patients most likely acquired SARS-CoV-2 in the community (Supplementary Table
302 1). RDU6 cases had a high likelihood of within-unit transmission of SARS-CoV-2 and
303 RDU6 also had one of the highest rates of infection over the longest time period (31st
304 March to 26th May, Table 2). However, in RDU1, where the rate of infection was also
305 high, there was tentative evidence of within-unit transmission; infections occurred
306 over 12 days, the incubation period of the last case was coincident with both
307 “lockdown” and enhanced PPE implementation.

308 **Discussion**

309
310 During the first wave of the UK pandemic, we studied SARS-CoV-2 infections within
311 6 affected Scottish RDUs. Using a genomic epidemiology approach, we found that
312 transmission of SARS-CoV-2 within RDUs was common, affecting 8-9% of dialysis
313 patients with a very high associated mortality (27%) in keeping with other recent
314 studies.⁸⁻¹¹ Many guidelines have evolved for the care of dialysis patients to minimise
315 risk of infection, including nosocomial transmission of SARS-CoV-2.^{5,6,24} Less is
316 known around whether patients requiring dialysis are at greater risk of community

317 transmission of SARS-CoV-2 and how infection in the dialysis units relates to viral
318 exposure in the healthcare environment compared to that in the community. Whole-
319 genome sequencing provides high-level resolution of SARS-CoV-2 genome and in
320 combination with epidemiological data can facilitate our understanding of
321 transmission and evolution during pathogen outbreaks.^{13,14} Applying careful analysis
322 of data generated from the community and the hospital setting, we found risk was
323 present both from community and hospital settings; almost always from within the
324 RDUs themselves, but on occasion from the wider hospital. Multiple introductions
325 occurred into the dialysis units, reflecting the risks associated with individuals having
326 increased contact with the community (including need for frequent travel to hospitals)
327 as well as prolonged and regular contact within the RDUs themselves. Data from the
328 Scottish Renal Registry suggest that measures introduced in early April to reduce
329 this exposure, including the use of masks to and from dialysis and individual
330 transport, were effective; with the number of cases in people receiving dialysis falling
331 sharply two to three weeks before the rest of the general population within
332 Scotland.²⁵ In order to capitalise on the utility of such knowledge in hospital-outbreak
333 management, the results from the sequencing data need to be generated in a timely
334 manner.^{15,17} However, the feasibility of implementing this rapid sequencing response
335 in the National Health Service to date has been impeded by lack of expertise and
336 equipment for both high-throughput whole-genome sequencing and for processing
337 the data generated into a form interpretable by infection control and public health.
338 COG-UK has demonstrated that near real-time sequencing is achievable at scale.¹⁷
339 We used this framework and a novel statistical algorithm to characterise
340 transmission dynamics specifically in the haemodialysis cohort, a group both at
341 higher risk of severe outcome as well as having numerous healthcare interactions.

342 A limitation of the study is the sequences available for analysis; accurate estimation
343 of the likely source of infection depends on having sufficient sequences available
344 from the community of affected patients and of cases from the hospital setting. Here,
345 we obtained sequences from two-thirds of the lab-detected SARS-CoV-2 cases from
346 RDU1-RDU5, and 44% for RDU6. We also compared data from the RDUs with 944
347 other cases in the community, 700 inpatients and 546 samples taken from patients
348 presenting in emergency departments in the same health boards as provide care for
349 dialysis patients. Additionally, COVID-19 is asymptomatic in up to 20% of patients,
350 which may have reduced the number of infections captured.²⁶ Further, early in the
351 pandemic, HCWs were ineligible for testing. Frequent, regular testing of all HCWs
352 and all patients, regardless of symptoms is warranted. There is also mounting
353 evidence that HCWs have a higher seroprevalence for SARS-CoV-2 antibodies than
354 the general population, with a high proportion being asymptomatic.^{27,28}

355 The low substitution rate of SARS-CoV-2 limits the granularity of outbreak analysis;
356 as demonstrated in this study, indistinguishable sequences may not be part of a
357 transmission cluster if there is widespread circulation in the patients' home
358 communities. To address this limitation, we employed the SRT, which combines
359 sequence data with both temporal and geographic data to improve estimates of
360 within-unit and within-hospital versus community transmission. Based on additional
361 epidemiological evidence such as timing of haemodialysis sessions and hospital
362 transport, the SRT correctly identified a high probability of within-unit transmission for
363 RDU 2, 3, 5 and 6. Less definitive results for RDU1, not immediately apparent by
364 phylogenetic investigation alone (due to the widespread presence the lineage within
365 the community), affirm its potential as a rapid tool to aid outbreak investigations.

366 We confirm the findings of other published reports of SARS-CoV-2 in the
367 haemodialysis cohort that cases are at risk of poor outcomes, with no specific at-risk

368 group identified based on comorbid conditions.⁹ The high mortality, dearth of
369 therapeutics and likely poor response to vaccines,²⁹ emphasises the need for
370 targeted strategies to mitigate risk in this cohort. Identification of major transmission
371 risks is vital to address outbreaks in this vulnerable group where there is prolonged,
372 unavoidable contact between healthcare settings and the local community. Whilst
373 universal infection control measures are beneficial, we identified multiple community-
374 acquired infections, with RDUs being an interface for transmission. Additional
375 measures may be required to reduce infection in this setting. Longer, more extreme
376 periods of intensive social distancing ('shielding') to reduce contact with other
377 individuals may be required when the community incidence of infection is high.
378 Knowledge of the dominant site of transmission can justify and provide precision to
379 the recommendation to shield and condense the period of isolation, loneliness and
380 distress in this cohort, who already have a high incidence of depression.
381 Although we demonstrate the utility of identifying the likelihood of transmission of
382 SARS-CoV-2 infection around treatment centres for haemodialysis, our findings have
383 resonance for any group requiring frequent attendance in healthcare facilities, such
384 as patients undergoing chemotherapy, radiotherapy or outpatient rehabilitation.

385 (Word count: 3484)

386

387

388

389 **Figures and Tables**

390 **Table 1**

391 Demographic, laboratory and imaging data of patients with COVID-19 with
392 comparisons between patients who died compared to survivors. Laboratory test
393 results are taken from the date closest to diagnosis of SARS-CoV2 (same day in
394 85% of cases with only two cases where laboratory day >2 days from date of
395 diagnosis). All patients requiring haemodialysis are registered on the Strathclyde
396 Electronic Renal Patient Record (SERPR; Vitalpulse, UK) which records clinical,
397 laboratory and imaging data for clinical care, audit and research. Using SERPR we
398 extracted anonymised clinical data on all patients treated with haemodialysis with a
399 positive test for SARS-CoV-2 infection. The Scottish Government provides online
400 calculators allowing use of patient postcode to generate divisions of socioeconomic
401 deprivation, the Scottish Index of Multiple Deprivation (SIMD)
402 (<http://www.gov.scot/Topics/Statistics/SIMD>). Deprivation deciles were calculated
403 and categorized into most deprived deciles with 1 corresponding to most deprived
404 and 10 as least deprived. Cause of death was certified by each patient's clinical
405 team with registration with the Scottish Mortality Audit in Renal Replacement
406 Therapy (SMARRT).³⁰ Radiological imaging was coded using the British Thoracic
407 Society Classification for COVID-19.
408 ([https://www.bsti.org.uk/media/resources/files/BSTI_COVID_CXR_Proforma_v.3-
409 1.pdf](https://www.bsti.org.uk/media/resources/files/BSTI_COVID_CXR_Proforma_v.3-1.pdf)). Values are presented as means (standard deviation) or medians (inter-
410 quartile range) and comparisons between groups were made using t-test, Kruskal-
411 Wallis, Chi-square and Fisher's exact test as appropriate. Statistics were performed
412 on Minitab Version 19.2020.1.0 (Minitab, State College, Pennsylvania).

413 Abbreviations BMI – body mass index, CVD - cardiovascular disease, COPD -
414 chronic obstructive pulmonary disease, PRD - primary renal disease, DN - diabetic

415 nephropathy, GN - glomerulonephritis, AVF - arteriovenous fistula, AVG -
416 arteriovenous graft, RRT - renal replacement therapy, WCC – white cell count, Hb –
417 haemoglobin, Plts - platelets, Neut – neutrophils, Lymph – lymphocytes, NLR –
418 neutrophil to lymphocyte ratio, CXR - chest X-ray, CT –computed tomography of
419 chest, CPAP – continuous positive airway pressure

420

421 **Table 2**

422 Number of patients treated at each RDU and proportion of patients infected with
423 SARS-CoV-2 per RDU.

424

425 **Figure 1**

426 Cumulative cases of COVID-19 cases (left y-axis) with arrows demonstrating
427 additional infection control measures (narrow arrow - RDU1, wide arrow covers the
428 dates for all other RDUs). Cumulative infection numbers for Scotland are on the right
429 y-axis.

430

431 **Figure 2**

432 Phylogenetic tree showing the relationship of 39 sequences from RDU patients and
433 additional SARS-CoV-2 genomes from Scotland. Sequences are colour-coded by
434 RDU location. Dashed boxes highlight the UK lineage and are shown in more detail
435 in Figure 3. The numerical suffixes of the CVR identifier indicate the posterior
436 probability (as a percentage) of the patient acquiring SARS-CoV-2 from the RDU
437 (p_{RDU}) or from the wider hospital where dialysis takes place (p_{hRDU}). The scale
438 bar indicates substitutions per nucleotide site.

439

440 **Figure 3**

		All patients (n=60)	SD/IQR/%	Alive (n=44)	SD/IQR/%	Died (n=16)	SD/IQR	p
Age (yrs)		67.4	13.7	66.2	14.9	70.6	9.6	0.19

441 Timeline of detection of first SARS-CoV-2 positive results in haemodialysis patients
442 in RDUs with details of dialysis sessions and shared patient transport in relation to
443 the UK lineage. The phylogenetic trees are derived from the dashed boxes in Figure
444 2. Circled numbers in the phylogenetic tree represent the number of
445 indistinguishable sequences from Scotland for the given node on the phylogeny. The
446 numerical suffixes of the CVR identifier indicate the posterior probability (as a
447 percentage) of the patient acquiring SARS-CoV2 from the RDU or from another
448 healthcare-related infection (i.e., hospital where dialysis takes place and ward and/or
449 hospital they have been admitted to), respectively. The scale bar indicates
450 substitutions per nucleotide site.

451

Male (%)		33	(55)	22	(36.7)	11	(18.3)	0.20
BMI (kg m ⁻²)		30.3	10.9	31.5	11.8	26.9	7.2	0.19
Race	White	56	(93.3)	42	(70.0)	15	(25.0)	-
	S. Asian	1	(1.7)	1	(1.7)	0	-	
	Black	1	(1.7)	1	(1.7)	0	-	
	Other	2	(3.4)	1	(1.7)	1	(1.7)	
SMID Decile		3	1-6	3	1-6	3	2-6	0.43
Diabetes		29	(48.3)	20	(33.3)	9	(15.0)	0.46
Cancer		7	(11.7)	5	(8.3)	2	(3.3)	1.0
CVD		23	(38.3)	15	(25.0)	8	(13.3)	0.26
COPD		5	(8.3)	2	(3.3)	3	(5.0)	0.11
PRD	DN	22	(36.7)	14	(23.3)	8	(13.3)	-
	GN	9	(15.0)	7	(11.6)	2	(3.3)	
	Interstitial	16	(26.7)	14	(23.3)	2	(3.3)	
	Multisystem	8	(13.3)	5	(8.3)	3	(5.0)	
	Other	5	(8.3)	4	(6.7)	1	(1.7)	
Dialysis vascular access	AVF	26	(43.3)	18	(30.0)	8	(13.3)	-
	AVG	1	(1.7)	0	-	1	(1.7)	
	Line	33	(55.0)	26	(43.3)	7	(11.6)	
RRT time (yrs)		3.3	1.2-5.8	3.4	1.1-5.6	2.8	1.2-6.1	0.89
Prev transplant		16	(26.7)	13	(21.7)	3	(5.0)	0.32
WCC		6.3	3.5	5.6	1.9	8.4	5.5	0.06
Hb (g L ⁻¹)		104.0	17.3	105.1	17.0	101.1	18.2	0.46
Plts (10 ⁹ L ⁻¹)		173.9	66.1	177.3	58.6	164.3	85.1	0.58
Neut (10 ⁹ L ⁻¹)		4.9	3.3	4.2	1.8	6.9	5.3	0.06
Lymph(10 ⁹ L ⁻¹)		0.75	0.35	0.75	0.34	0.76	0.38	0.92
NLR		7.8	6.9	6.5	3.9	11.5	11.1	0.10
CRP (mg L ⁻¹)		46.0	-	44.5	26.3-77.3	69.0	35.0-104.0	0.32
CXR/CT	Normal	7	(15.9)	7	(15.9)	0	-	-
	Equivocal	22	(50.0)	13	(29.5)	9	(20.5)	
	Classical	14	(31.8)	9	(20.5)	5 (11.4)	(11.4)	
	Alt. diagnosis	1	(2.3)	1	(2.3)	0	-	
Ventilated	No	58	(96.7)	43	(71.7)	15	(25.0)	-
	CPAP	1	(1.7)	0	-	1	(1.7)	
	Invasive	1	(1.7)	1	(1.7)	0	-	

452 Table 1
453
454
455
456
457
458
459
460 Table 2
461

	Number of patients with SARS-CoV2	Total number of patients treated	Percentage diagnosed with SARS-CoV2 infection	First case to final case detected (days)
RDU1	9	60	15.0	12
RDU2	11	109	10.1	38
RDU3	7	104	6.7	66
RDU4	2	50	4.0	9
RDU5	3	137	2.2	12
RDU6	28	211	13.2	56

462

References

463

- 464 1. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory
465 disease in China. *Nature* 2020; **579**(7798): 265-9.
- 466 2. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia
467 in China, 2019. *N Engl J Med* 2020; **382**(8): 727-33.
- 468 3. Docherty AB, Harrison EM, Green CA, et al. Features of 202133 UK patients in
469 hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective
470 observational cohort study. *BMJ* 2020; **369**: m1985.
- 471 4. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With
472 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* 2020.

- 473 5. Brioni E, Leopaldi D, Magnaghi C, et al. [Covid-19 in patients on dialysis: infection
474 prevention and control strategies]. *G Ital Nefrol* 2020; **37**(2).
- 475 6. Ikizler TA, Klinger AS. Minimizing the risk of COVID-19 among patients on dialysis. *Nat*
476 *Rev Nephrol* 2020; **16**(6): 311-3.
- 477 7. Rombolà G, Brunini F. COVID-19 and dialysis: why we should be worried. *J Nephrol*
478 2020; **33**(3): 401-3.
- 479 8. Cho JH, Kang SH, Park HC, et al. Hemodialysis with Cohort Isolation to Prevent
480 Secondary Transmission during a COVID-19 Outbreak in Korea. *J Am Soc Nephrol* 2020;
481 **31**(7): 1398-408.
- 482 9. Ravanan R, Callaghan CJ, Mumford L, et al. SARS-CoV-2 infection and early mortality
483 of wait-listed and solid organ transplant recipients in England: a national cohort study. *Am J*
484 *Transplant* 2020.
- 485 10. Valeri AM, Robbins-Juarez SY, Stevens JS, et al. Presentation and Outcomes of
486 Patients with ESKD and COVID-19. *J Am Soc Nephrol* 2020; **31**(7): 1409-15.
- 487 11. Wu J, Li J, Zhu G, et al. Clinical Features of Maintenance Hemodialysis Patients with
488 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *Clin J Am Soc Nephrol* 2020;
489 **15**(8): 1139-45.
- 490 12. Brown AC, Bryant JM, Einer-Jensen K, et al. Rapid Whole-Genome Sequencing of
491 Mycobacterium tuberculosis Isolates Directly from Clinical Samples. *J Clin Microbiol* 2015;
492 **53**(7): 2230-7.
- 493 13. Quick J, Grubaugh ND, Pullan ST, et al. Multiplex PCR method for MinION and
494 Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat*
495 *Protoc* 2017; **12**(6): 1261-76.

- 496 14. Quick J, Loman NJ, Duraffour S, et al. Real-time, portable genome sequencing for
497 Ebola surveillance. *Nature* 2016; **530**(7589): 228-32.
- 498 15. Houlihan CF, Frampton D, Ferns RB, et al. Use of Whole-Genome Sequencing in the
499 Investigation of a Nosocomial Influenza Virus Outbreak. *J Infect Dis* 2018; **218**(9): 1485-9.
- 500 16. Filipe ADS, Shepherd J, Williams T, et al. Genomic epidemiology of SARS-CoV-2
501 spread in Scotland highlights the role of European travel in COVID-19 emergence. *medRxiv*
502 2020: 2020.06.08.20124834.
- 503 17. Meredith LW, Hamilton WL, Warne B, et al. Rapid implementation of SARS-CoV-2
504 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic
505 surveillance study. *Lancet Infect Dis* 2020.
- 506 18. Oude Munnink BB, Nieuwenhuijse DF, Stein M, et al. Rapid SARS-CoV-2 whole-
507 genome sequencing and analysis for informed public health decision-making in the
508 Netherlands. *Nat Med* 2020.
- 509 19. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-
510 nCoV) by real-time RT-PCR. *Euro Surveill* 2020; **25**(3).
- 511 20. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler
512 transform. *Bioinformatics* 2010; **26**(5): 589-95.
- 513 21. Grubaugh ND, Gangavarapu K, Quick J, et al. An amplicon-based sequencing
514 framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar.
515 *Genome Biol* 2019; **20**(1): 8.
- 516 22. Stirrup OT, Hughes J, Parker M, et al. Rapid feedback on hospital onset SARS-CoV-2
517 infections combining epidemiological and sequencing data. *medRxiv* 2020.

518 23. Lauer SA, Grantz KH, Bi Q, et al. The Incubation Period of Coronavirus Disease 2019
519 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern*
520 *Med* 2020; **172**(9): 577-82.

521 24. Rombolà G, Heidempergher M, Pedrini L, et al. Practical indications for the
522 prevention and management of SARS-CoV-2 in ambulatory dialysis patients: lessons from
523 the first phase of the epidemics in Lombardy. *J Nephrol* 2020; **33**(2): 193-6.

524 25. Bell S, Campbell J, McDonald J, et al. COVID-19 in patients undergoing chronic kidney
525 replacement therapy and kidney transplant recipients in Scotland: findings and experience
526 from the Scottish renal registry. *BMC Nephrol* 2020; **21**(1): 419.

527 26. Buitrago-Garcia DCE-G, D.; Counotte, M.J.; Hossmann, S.; Imeri, H; Ipekci, A.M.;
528 Salanti, G.; Low, N. Asymptomatic SARS-CoV-2 infections: a living systematic review and
529 meta-analysis. <https://doi.org/10.1101/2020042520079103> 2020.

530 27. Abo-Leyah H, Gallant S, Cassidy D, et al. Seroprevalence of SARS-COV-2 Antibodies in
531 Scottish Healthcare Workers. *medRxiv* 2020: 2020.10.02.20205641.

532 28. Houlihan CF, Vora N, Byrne T, et al. Pandemic peak SARS-CoV-2 infection and
533 seroconversion rates in London frontline health-care workers. *Lancet* 2020; **396**(10246): e6-
534 e7.

535 29. Gangappa S, Wrammert J, Wang D, et al. Kinetics of antibody response to influenza
536 vaccination in renal transplant recipients. *Transpl Immunol* 2019; **53**: 51-60.

537 30. Findlay MD, Donaldson K, Doyle A, et al. Factors influencing withdrawal from
538 dialysis: a national registry study. *Nephrol Dial Transplant* 2016; **31**(12): 2041-8.

539

540 **Author contributions:** Each named author has substantially contributed to
541 conducting the underlying research and drafting this manuscript. Additionally, to the
542 best of our knowledge, the named authors have no conflict of interest, financial or
543 otherwise.

544

545 **Ethical approval:** Use of anonymised data was approved by the Local Privacy
 546 Advisory Group of NHS Greater Glasgow and Clyde ‘Safe Haven’ on behalf of the
 547 West of Scotland Ethics Committee (approval GSH/20/RE/001). Surplus RNA extract
 548 was collected for sequencing with ethical approval from the NHSGGC biorepository
 549 (16/WS/0207NHS).

550

551 **Funding:** Funding: COG-UK is supported by funding from the Medical Research
 552 Council (MRC) part of UK Research & Innovation (UKRI), the National Institute of
 553 Health Research (NIHR) and Genome Research Limited, operating as the Wellcome
 554 Sanger Institute. Also MRC (MC UU 1201412), UKRI/Wellcome (COG-UK),
 555 Wellcome Trust Collaborator Award (206298/Z/17/Z – ARTIC Network; TCW
 556 Wellcome Trust Award 204802/Z/16/Z).

557

	Data sharing statement
Will individual participant data be available?	Yes
What data in particular will be shared?	Individual participant data that underlie the results reported in this article, after de-identification. Sequences and de-identified metadata is available on MRC-CLIMB through COG-UK. Sequences are also available on GISAID.
What other documents will be available?	Study protocol, statistical analysis, analytic code
When will data be available?	Immediately following publication, no end date
With whom?	Researchers who provide a methodologically sound proposal
For what types of analysis?	To achieve aims in the approved proposal

<p>By what mechanism will data be made available?</p>	<p>Proposals should be directed to contact@cogconsortium.uk</p> <p>(Statement from COG-UK consortium: <i>“We are committed to open science, and sharing all data that we can as rapidly as possible. This includes sharing data for use by Public Health authorities internationally, to support COVID-19 response, and sharing data in such a way that the academic community can access and use the data and analysis according to FAIR data principles.”</i>)</p>
---	--

558

CVR1816 74+3
CVR937 0+1

UK39

CVR157 0+0 UK14

UK40

UK51

UK5098

medRxiv preprint doi: <https://doi.org/10.1101/2021.03.24.21268667>; this version posted March 26, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

CVR1817 73+3 UK501

CVR1438 0+0 UK2913

CVR3839 0+7 UK36

CVR3289 0+0 UK658
CVR3290 100+0

CVR3842 0+13 UK370

CVR3732 0+96

CVR2228 0+0 UK5

CVR536 0+0

UK429

CVR2124 0+0 UK498

7.0E-5

Dialysis unit

- RDU1
- RDU2
- RDU3
- RDU4
- RDU5
- RDU6



