

1 **The COVID HOME study research protocol:**
2 **Prospective cohort study of non-hospitalised COVID-19 patients**

3

4 **Authors:** Tami A. ^{1*}, van der Gun B.T.F. ¹, Wold K.I. ¹, Vincenti-González M.F. ¹, Veloo A.C.M. ¹, Knoester
5 M. ¹, Harmsma V.P.R. ¹, de Boer G.C. ¹, Huckriede A.L.W. ¹, Pantano D. ¹, Gard L. ¹, Rodenhuis-Zybert I. ¹,
6 Upasani V. ¹, Smit J. ¹, Dijkstra A. ¹, de Haan J. ², van Elst J. ², van den Boogaard J. ³, O' Boyle S. ⁴, Nacul L. ⁵,
7 Niesters H.G.M. ¹, Friedrich A.W. ¹

8 ¹ University of Groningen, University Medical Center Groningen, Department of Medical Microbiology and
9 Infection Prevention, Groningen, The Netherlands

10 ² University of Groningen, University Medical Center Groningen, Department of Medical Oncology, Groningen,
11 The Netherlands

12 ³ Municipal Public Health Service Groningen (GGD Groningen) – Groningen, the Netherlands

13 ⁴ Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, United
14 Kingdom

15 ⁵ University of British Columbia, Faculty of Medicine and London School of Hygiene and Tropical Medicine,
16 Department of Clinical Research.

17 * Corresponding author.

18

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

19 Funding

20 The project received funding from the Netherlands Organisation for Health Research and
21 Development (ZonMw) [1] , grant 10430012010023, and as partner of the ORCHESTRA project [2]
22 which has received funding from the European Union’s Horizon 2020 research and innovation
23 programme under grant agreement No 101016167. The views expressed in this publication are the
24 sole responsibility of the author and the Commission is not responsible for any use that may be made
25 of the information it contains. The funders had and will not have a role in study design, data collection
26 and analysis, decision to publish, or preparation of the manuscript.

27 Competing interests

28 The authors have declared that no competing interests exist.

29

30 Abstract

31 **Background.** Guidelines on COVID-19 management are developed as we learn from this pandemic.
32 However, most research has been done on hospitalised patients and the impact of the disease on non-
33 hospitalised and their role in transmission are not yet well understood.

34 The COVID HOME study conducts research among COVID-19 patients and their family members who
35 were not hospitalised during acute disease, to guide patient care and inform public health guidelines
36 for infection prevention and control in the community and household.

37 **Methods.** An ongoing prospective longitudinal observational study of COVID-19 outpatients was
38 established in March 2020 in the Netherlands. Laboratory confirmed SARS-CoV-2 infected individuals
39 of all ages that did not merit hospitalisation, and their household (HH) members, were enrolled after
40 written informed consent. Enrolled participants were visited at home within 48 hours after initial

41 diagnosis, and then weekly on days 7, 14 and 21 to obtain clinical data, a blood sample for biochemical
42 parameters/cytokines and serological determination; and a nasopharyngeal/throat swab plus urine,
43 stool and sperm or vaginal secretion (if consenting) to test for SARS-CoV-2 by RT-PCR (viral shedding)
44 and for viral culturing. Weekly nasopharyngeal/throat swabs and stool samples, plus a blood sample
45 on days 0 and 21 were also taken from HH members to determine whether and when they became
46 infected. All participants were invited to continue follow-up at 3-, 6-, 12- and 18-months post-infection
47 to assess long-term sequelae and immunological status.

48 **Preliminary Results.** A total of 256 participants belonging to 103 HH were included of which, 190
49 (74.2%) were positive for SARS-CoV-2 infection. Most individuals (183/190, 96.3%) developed mild to
50 moderate disease. At the time of writing, all participants had reached the 3 and 6 month time-points
51 of the long-term follow-up, while approximately 78% reached 12 month and 23% the 18 month time-
52 point. Preliminary analysis showed that 43% (52/121) positive individuals reported having complaints
53 at 3 months post-infection, while 42.7% (61/143) had complaints at 6 months.

54

55 **Introduction**

56 Knowledge on COVID-19, the disease caused by coronavirus SARS-CoV-2, is evolving at an enormous
57 pace since the virus emerged in Wuhan, China at the end of 2019 [3, 4]. SARS-CoV-2 was first detected
58 in the Netherlands at the end of February 2020 [5] and rapidly spread throughout the country causing
59 a high mortality and morbidity in those infected during what will be later known as the first COVID-19
60 wave [6]. At the time and still nowadays, most of the research had been focussed on hospitalised
61 infected individuals [7, 8], while much less is known about the acute and long-term clinical evolution
62 in non-hospitalised patients. Moreover, the need to understand when individuals are not infectious
63 any longer and can resume their working activities without the risk of transmitting the virus to others

64 was an urgent question then and continues to be relevant given the emergence of new SARS-CoV-2
65 variants [9].

66 To answer these questions, the COVID HOME study (<https://www.umcg.nl/-/covid-home>) was set up
67 in March 2020 by researchers from the Department of Medical Microbiology and Infection Prevention
68 of the University Medical Center Groningen (UMCG), the Netherlands in collaboration with the
69 Municipal Health Services (GGD in Dutch) of the north of the Netherlands. The project focusses on the
70 impact of COVID-19 on non-hospitalised individuals, their acute and long-term clinical presentation,
71 infectious period and transmission routes. We aimed to answer research questions that would inform
72 public health guidelines on biosafety and transmission-control measures, as well as understand
73 COVID-19 disease and its management in home-isolated patients.

74 Guidelines for quarantine and isolation measures vary by country and between regions within
75 countries. Understanding the relevance of long-term (infectious) viral shedding in infected individuals
76 [10] is key to determine when to de-escalate infection prevention and control practices and allow
77 individuals to resume activities. Viral shedding in stool has proved relevant for monitoring the
78 epidemic evolution through wastewater surveillance [11]. Moreover, the dynamics of respiratory and
79 faecal viral shedding in the same individual may be linked to differing clinical presentations and may
80 also give clues about a potential faecal-oral transmission route [12, 13]. Other pathways, such as via
81 urine or sexual intercourse, may play a role in transmission [14]. However, it is still unclear how
82 frequent and relevant is the presence of SARS-CoV-2 virus in vaginal and seminal fluids [15-17].
83 Furthermore, understanding SARS-CoV-2 transmission within restricted indoor areas, such as
84 households (HH), gives detailed insight into transmission dynamics and transmission routes amongst
85 contacts [18]. This may shed light on how HH transmission contributes to the overall community
86 transmission (critical to sustain a pandemic), becoming a potential target for intervention.

87 Characterisation of the clinical phenotypes of patients with 'mild' disease is still lacking. The WHO
88 clinical progression scale [19] classifies mild disease into asymptomatic and into two broad

89 symptomatic strata. However, COVID-19 patients have diverse clinical presentations encompassing
90 the number and duration of symptoms [20] as well as different alterations of laboratory parameters.
91 Some of these clinical phenotypes could be linked to evolution to severe disease and/or to the
92 development of long-term sequelae [21]. Several blood parameters including cytokines, have been
93 identified as markers of COVID-19 disease severity [22, 23] while other studies have characterised
94 blood parameters to predict COVID-19 diagnosis in patients presenting to hospital with symptoms [22,
95 24]. However, data collected in a systematic and longitudinal manner soon after infection will better
96 characterise the clinical features and laboratory parameters (haematological, biochemical,
97 immunological) in outpatient individuals convalescing at home and their potential as early markers of
98 disease progression. A better understanding of the disease course in outpatients is needed to guide
99 standard of care to properly manage patients at home, administer early treatment and predict
100 progression to severe or long-term disease on time.

101 Here we report the design of a prospective observational cohort study of home-isolated SARS-CoV-2
102 infected individuals and their HH members. The main strength of this study is the early, systematic,
103 prospective follow-up with an intense sampling/data collection of non-hospitalised COVID-19 patients
104 during the acute disease evolution with a medium and long-term follow-up. The comprehensiveness,
105 diverse data collection and analysis of the study described here will allow us and other researchers to
106 add knowledge to the already ongoing fast research on COVID-19.

107 **Objectives**

108 The objectives of the study have been divided into two work packages (WP) and their respective sub-
109 objectives.

- 110 1. In WP1 we aim to measure the duration and routes of viral shedding, genetic diversity, and
111 development of immunity of non-hospitalised individuals with COVID-19 to improve guidelines for
112 biosafety and patient isolation by:

- 113 1.1. Determining the viral load kinetics and duration of viral shedding of SARS-CoV-2 RNA in bodily
114 specimens through real-time reverse transcription polymerase chain reaction (real-time RT-
115 PCR)
- 116 1.2. Evaluating the viability of the virus in different specimens and at various time points of
117 patient follow-up, by viral culturing
- 118 1.3. Determining the relation of viral load kinetics and the appearance and kinetics of anti-SARS-
119 CoV-2 IgM, and IgG antibodies over time
- 120 1.4. Determining the genetic diversity of SARS-CoV-2 through whole-genome sequencing (WGS)
- 121 1.5. Determining the chain of transmission within the HH and into the working environment,
122 especially for healthcare settings, through viral WGS
- 123 2. Within WP2 we expect to aid in establish guidelines for the management of (non-hospitalised)
124 COVID-19 patients at home, including early detection of predicting factors for severity and long-
125 term consequences by:
- 126 2.1. Describing the clinical spectrum of COVID-19 in non-hospitalised individuals and determine
127 potential risk factors for mild versus more severe clinical evolution
- 128 2.2. Assessing the development of long-term sequelae and post-viral fatigue syndrome (PVFS) in
129 non-hospitalised COVID-19 individuals
- 130 2.3. Correlating laboratory parameters and cytokine changes with clinical evolution
- 131 2.4. Determining the relation of viral loads & anti-SARS-CoV-2 antibody kinetics with clinical
132 evolution

133 Rationale

134 The data gathered in **WP1** will inform public health guidelines on biosafety, duration of home isolation
135 and de-escalation measures as well as improve our understanding of transmission chains in the

136 community (HH, workplace, healthcare centres). In **1.1** we aim to understand whether viral detection
137 in different body specimens in non-hospitalised individuals differs from what is being found in
138 hospitalised patients, thus understanding the contribution of outpatients to the overall transmission
139 of COVID-19. Here, we tested for SARS-CoV-2 RNA on systematically and simultaneously collected
140 clinical specimens from different body sites (respiratory, faeces, blood, urine, semen and vaginal fluid)
141 for at least 4 weeks, giving detailed data to compare to the clinical evolution and other parameters
142 such as immune response (**1.3**) and viral genotypes (**1.4**). In previous studies SARS-CoV-2 RNA was
143 found in all of these specimens with conflicting results for vaginal fluid [15, 17]; with blood, urine and
144 semen found positive in a much lower frequency than respiratory samples and faeces [14, 25, 26].
145 Secondly, in **1.2**, the RNA-positive specimens were cultured to ascertain if infectious virus could be
146 found in the samples mentioned in **1.1**, at different time points during the clinical evolution. The
147 presence and duration of infectious virus in different specimens influences infection prevention
148 measures. Recent studies showed shedding of infectious virus in respiratory samples from critical
149 hospitalised patients to last up to 20 days (median=8 days) since symptom onset [27], while other
150 studies detected infectious virus in asymptomatic or mildly affected patients up to 10 days [10] but
151 not after two weeks of diagnosis [28].

152 Follow-up studies on COVID-19 patients have shown that anti-SARS-CoV-2 IgM appears at around 4-
153 14 days from symptom onset, while IgG seroconversion is seen between days 7-28 [29, 30]. These
154 cohorts were mostly composed by hospitalised patients and IgG seroconversion was delayed in less
155 severe cases [30], while mild SARS-CoV-2 infections induced a more modest antibody response [31].
156 **In 1.3** anti-SARS-CoV-2 antibody dynamics of our cohort of patients were measured and are being
157 aligned with the viral shedding and virus culture results, to evaluate antibody kinetics, seroconversion
158 rates and duration and the relation of serological response with viral clearance and virus viability. To
159 assess long-term immunity, IgG data will be obtained up to 18 months post-infection in consenting
160 patients (see Methods).

161 **In 1.4** we perform WGS analysis to determine the SARS-CoV-2 genetic diversity in the population under
162 study. We want to establish the molecular and macro-epidemiologic context by comparing the
163 sequences from participants with known sequences in our cross-border region as well as from
164 international sequences to understand introduction and first seeding events of transmission into our
165 region. WGS studies have shown that independent introductions resulted in patterns of local or
166 national spread of viral types leading to lineage replacement [32, 33]. WGS was key to determine
167 patterns of transmission among health care workers (HCW) in NL [34]. Transmission of SARS-CoV-2
168 within healthcare institutions seems to be one of the most important drivers of the regional epidemic.
169 **In 1.5**, we combine epidemiological, diagnostic and viral genetic data to determine transmission chains
170 within the HH: if, when and how long does it take for other family members to become infected
171 (especially children) and cluster analysis using viral WGS. The transmission of SARS-CoV-2 from/to the
172 community by HCW and their family members is being addressed here.

173 In **WP2** we strive to describe the spectrum of acute clinical presentation and the interrelation with
174 paraclinical parameters in COVID-19 infected individuals that do not merit or get admitted to hospital.
175 As we have enrolled HH members of COVID-19 positive individuals (see Methods), we are able to
176 identify the whole spectrum of disease including asymptomatic individuals. Clinical phenotypes for
177 hospitalised patients have been described in recent studies [23, 35]. However, mild COVID-19 disease
178 presentation has been classified in broad categories [19]. **In 2.1** we attempt to define clinical
179 phenotypes combining clinical and laboratory parameters collected in a systematic manner during the
180 acute disease and correlate these with prognosis. Potential (risk) factors for mild versus more severe
181 clinical evolution will include age, gender, BMI and specific blood parameters that have been related
182 with severe disease progression such as C-reactive protein, hepatic transaminases, lactate
183 dehydrogenase (LDH), leptin, interleukin (IL)-6, ferritin & other [36].

184 Clinical evidence on the subacute and long-term persistence of symptoms after SARS-CoV-2 infection
185 or Long COVID, was recognised during the first few months of the COVID-19 pandemic [37, 38]. The

186 persistence of symptoms, similar to myalgic encephalomyelitis/chronic fatigue syndrome, likely result
187 from complex mechanisms involving an abnormal immune response to infection, central and
188 autonomic nervous system and metabolic dysfunction among other mechanisms [39]. Viral infections
189 are important triggers and may lead to what is often referred to as post-viral fatigue syndrome (PVFS)
190 or post-covid fatigue syndrome, in cases that follow SARS CoV-2 infection. In COVID HOME, we follow
191 consenting individuals at 3, 6, 12 and 18 months post diagnosis to evaluate fatigue and related
192 symptoms. The data collected in **2.1** is used in **2.2** to identify clinical and laboratory parameters
193 (biochemical and immunological), and clinical phenotypes, associated with the development of long-
194 term sequelae. Immunopathology appears to play a very important role in SARS-CoV-2 pathogenesis.
195 Critical disease has been found to be associated with a ‘cytokine storm’, the overproduction of
196 inflammatory cytokines, ultimately resulting in acute respiratory distress syndrome (ARDS) [40]. In
197 hospitalised patients, elevated levels of various cytokines, including IL-1 β , IL-6, IL-10, IL-12, and tumor
198 necrosis factor (TNF)- α , and chemokines, in particular CXCL10 and CCL2, have been found and the
199 levels partly correlated with disease severity [41]. Cytokines may also play an important role in less
200 severe disease courses and in Long COVID condition [42]. In **2.3** the dynamics of these cytokines will
201 be determined during the acute disease and the long-term clinical evolution. Finally, in **2.4** we test the
202 relationship of viral load dynamics and anti-SARS-CoV-2 antibody kinetics (collected in WP1 sections
203 1.1-1.3) with the clinical evolution. Viral loads of asymptomatic individuals have been reported to be
204 similar to symptomatic patients [43]. However, in the case of anti- SARS-CoV-2 antibodies, reports
205 suggest a weak immune response in mild clinical cases, as mentioned above. We aim to generate a
206 larger set of data in relation to asymptomatic, mild and more severe disease presentation to answer
207 these questions.

208 **Methods**

209 **Study population and study area**

210 The study is carried out amongst residents of the Northern Provinces of the Netherlands (Groningen,
211 Friesland, Overijssel and Drenthe) comprising around 1.7 million inhabitants. Laboratory confirmed
212 SARS-CoV-2 infected individuals of all ages (index cases) that do not merit hospitalisation, and their
213 HH members, were invited to join the study and enrolled after signing a written informed consent.
214 Consenting parents/guardians of infected children (≤ 16 years) signed the informed consent form. The
215 complete HH, i.e. all persons (contacts) living in the same home, was to join the study at enrolment.
216 If this requirement was not fulfilled, the HH was excluded from the study.

217 **Study design**

218 An ongoing prospective longitudinal observational study of SARS-CoV-2 infected individuals that did
219 not require hospitalisation and their HH members was initiated on March 2020, at the start of the
220 pandemic in the Netherlands.

221 Enrolled individuals comprised the identified laboratory COVID-19 index positive patients plus all their
222 HH members. The average number of HH inhabitants in the Netherlands and the northern region is
223 2.14 [44], thus for each enrolled patient we expected to enrol an average of 1-2 HH members in
224 addition. Patients were followed daily at home approximately during 3 weeks post-infection to
225 characterize their acute clinical phenotype, and are currently being followed at 3, 6, 12 and 18 months
226 to determine their long-term clinical and serological evolution. During the follow-up, clinical data and
227 different types of samples were/are collected, as described below.

228 **Inclusion criteria of index participant**

229 The index participant was the first laboratory-confirmed SARS-CoV-2 infected individual of any age
230 identified in a HH and invited to join the study together with the complete HH. The requirements for
231 inclusion were:

- 232 • Onset of symptoms < 5 days before inclusion

- 233 • Positive diagnostic RT-PCR for SARS-CoV-2 performed within 48h of inclusion
- 234 • Index person should supply all required samples in the protocol for the duration of the study
- 235 • All HH members should join the study
- 236 • Signing of informed consent form

237 **Data and sample collection**

238 **a. Identification of index cases**

239 Potential index cases were identified at the virology facility of the Department of Medical
240 Microbiology and Infection Prevention (MMBI, UMCG) and at the laboratories reporting to the GGD
241 of the Northern Provinces. Suspected SARS-CoV-2 infected individuals were diagnosed by real-time
242 RT-PCR performed on material from nasopharyngeal/throat (NPT) swabs. Laboratory SARS-CoV-2
243 positive individuals were contacted by phone by an infection prevention expert from either the UMCG
244 MMBI department or the GGD who explained the study objectives and procedures and invited the
245 positive individual and the HH members into the study if the inclusion criteria were met. Eligible
246 individuals were also directed to the project's website (<https://www.umcg.nl/-/covid-home>) for
247 further information. Patients were also given time to ask questions and/or to think before deciding to
248 join.

249 **b. First home visit**

250 Once a patient agreed to participate in the study, a first home visit (from now on called Day 0) was
251 planned within the next 24 hours to enrol participants into the study after signing a written informed
252 consent corresponding to the participant's age (in duplicate, dated and signed by both participant and
253 project researcher). Participants were asked to supply their individual e-mail address in the informed
254 consent and to authorise it's use to receive electronic questionnaires and information from the COVID
255 HOME researchers via the project's email. Participants were also encouraged to use the project's email

256 to ask questions or to inform of any event. A schedule was drawn so that study participants were
257 visited weekly for at least 21 days by project staff to collect data and samples according to the different
258 protocols. Those not tested or who were negative for SARS-CoV-2 RT-PCR, were enrolled into the
259 protocol for HH members (section d).

260 **c. Follow-up of SARS-CoV-2 infected individuals (acute phase of the** 261 **disease/baseline)**

262 A systematic weekly follow-up at the patient's home was set up to acquire data and samples to
263 determine kinetics, length, viability and routes of viral shedding (objectives 1.1 & 1.2), serological
264 dynamics (objective 1.3), cytokine kinetics (objective 2.3) and evolution of haematological &
265 biochemical blood parameters and their relationship with the patient's clinical progression (objectives
266 2.1, 2.2, 2.4).

267 Enrolled patients were visited at home soon after initial diagnosis (Day 0), and then weekly on days 7,
268 14 and 21 after confirmed infection (Fig 1) to obtain clinical data and a blood sample for laboratory
269 parameters, cytokine and serological determination. Research staff also obtained an NPT swab, urine
270 and stool samples; and sperm or vaginal secretion (if consenting) from adult individuals (>16 years) to
271 test for SARS-CoV-2 by RT-PCR (viral shedding) and for viral culturing (Fig 2). Children ≤16 years were
272 only required to supply an NPT swab, urine and stool samples. Different paper and electronic case
273 report forms (CRF/eCRF) were applied as described below (section d.) and in Fig 3. PCR results were
274 communicated to the participants the following working day. If still positive on Day 21 for any of the
275 specimens, the participant was invited to continue weekly sampling for RT-PCR testing for two further
276 weeks.

277 **Fig 1. Acute and long-term follow-up timeline of participants enrolled in the COVID HOME study.**

278 **Fig 2. Sampling flowchart for positive and negative SARS-CoV-2 participants.**

279 NPT= nasopharyngeal/throat

280 **Fig 3. Case report forms (CRF) applied to study participants.**

281 CRF-01 and CRF-03 are applied to all participants while CRF-10a is answered by non-infected
282 individuals. All other CRFs are applied to SARS-CoV-2 infected individuals.

283

284 Research staff visiting infected individuals at home followed safety measure protocols designed
285 according to UMCG and national guidelines. Standard operating procedures (SOP) including flow-
286 charts were designed for: i) safety procedures during home visits and at work, ii) handing/sending and
287 retrieving CRFs from participants, iii) preparation of sampling materials and transport requirements,
288 iv) sampling strategy (Fig 2), and v) aliquoting and storage conditions of samples; as well as protocols
289 for sample processing.

290 **d. Case report forms**

291 The following CRF/eCRFs were/are applied to infected individuals (Fig 3):

- 292 ▪ A self-completing clinical questionnaire (CRF-02) was given to patients at enrolment to record
293 socio-demographic data, comorbidities, treatment and daily clinical history (symptoms and signs)
294 during at least 21 days, and further if symptoms were still present. Participants still negative by
295 RT-PCR but presenting symptoms were encouraged to fill this form as well.
- 296 ▪ An individual HH questionnaire (CRF-03) was applied at enrolment to record socioeconomic
297 characteristics, preventive measures, behaviour and intradomiciliary isolation measures taken.
- 298 ▪ A phone interview (CRF-04) was carried out at 21 days, and at 3 and 6 months to assess smell and
299 taste alterations in detail in consenting individuals.
- 300 ▪ To evaluate the development of PVFS as part of the assessment of Long COVID, participants older
301 than 16 years were asked to answer on Day 14 and then at 3, 6, 12 and 18 months post-infection
302 to:

- 303 ▪ The fatigue screening questionnaire (eCRF-05) to establish presence or absence of
304 fatigue/tiredness, length of time of fatigue/tiredness, and any loss of activity
- 305 ▪ Additional fatigue questionnaire (eCRF-06) to identify presence and severity of symptoms
306 associated with development of PVFS
- 307 ▪ Health-related Quality of Life (RAND-36) questionnaire (eCRF-07) and to
308 the Hospital Depression and Anxiety Scale (HADS) questionnaire (eCRF-08)
- 309 ▪ At 3, 6, 12 and 18 months post-infection, both adults and children receive eCRF-10 to assess Long
310 COVID related symptoms, new comorbidities and re-infection.

311 CRF-02 and CRF-03 were given to participants to fill in paper while the rest were eCRFs and were
312 answered online using the REDCap™ (Research Electronic Data Capture) secure web-based software
313 platform [45]. Automatic electronic reminders were set up. In case of failure to complete the eCRFs,
314 participants were called to check if email addresses were correct or if any other issues had arisen.

315 **e. Follow-up of HH members of enrolled individuals (baseline)**

316 HH members who had not been tested or were negative for SARS-CoV-2 RT-PCR at enrolment were
317 followed with weekly testing of NPT swabs and stool samples for 3 weeks to determine if and when
318 each individual became infected. At the beginning of the study, it became apparent that some
319 individuals, especially children, stayed RT-PCR negative for NPT swabs, but became positive for SARS-
320 CoV-2 infection when stool samples were tested by RT-PCR. Therefore, both NPT swabs and stool
321 samples were collected weekly from (still) negative individuals. This information is being used to
322 understand transmission chains, timing of infection between close contacts (family/HH members) and
323 family cluster dynamics (objective 1.5). Moreover, to ascertain if RT-PCR negative individuals had been
324 previously infected or seroconvert during follow-up, blood samples were drawn at enrolment and on
325 Day 21 for IgM and IgG serology testing.

326 All HH members were asked to fill CRF-03 at enrolment and CRF-02 if presenting symptoms regardless
327 of RT-PCR results. Once an individual was confirmed by real-time RT-PCR of any of the samples (NPT
328 swabs and/or stool), the protocol for COVID-19 infected individuals (section c.) was followed.

329 **f. Long-term follow-up**

330 At 3, 6, 12 and 18 months post-infection, SARS-CoV-2 positive participants are invited to answer PVFS
331 and eCRF-10 questionnaires sent digitally to determine post-COVID long-term sequelae (section d.)
332 and provide blood samples for serologic determination.

333 Likewise, individuals that remained RT-PCR negative for SARS-CoV-2 infection have a blood sample
334 taken at 3, 6, 12 and 18 months after enrolment for serological testing and a slightly modified version
335 of eCRF-10 (eCRF-10a) is applied to attempt to differentiate long-term symptoms associated to COVID-
336 19 from other causes, such as lockdown measures.

337 **g. Substudy on smell and taste alterations**

338 Smell and taste alterations are common in the acute phase during COVID-19 infection [46]. Alterations
339 in smell and taste can have a profound impact on daily life and quality of life (QoL) by decreasing food
340 liking and appetite [47]. However, the long-term impact has not yet been studied in COVID-19
341 patients. This substudy assesses the prevalence, characteristics and the impact of smell and taste
342 alterations on daily life and QoL in COVID-19 patients at 3 weeks, and at 3 and 6 months after
343 diagnosis. Knowledge on these alterations could lead to more personalized guidance of COVID-19
344 patients by their physician or a dietician.

345 Consenting study participants were called three weeks after COVID-19 diagnosis to fill in a
346 questionnaire (CRF-04) regarding smell and taste alterations by phone. The questionnaire was based
347 on a previously used questionnaire which was designed to study taste and smell disturbances in
348 patients with gastrointestinal stromal tumors treated with tyrosine-kinase inhibitors [48]. The
349 questionnaire was adjusted to study COVID-19 patients. It contains four parts where participants

350 report the characteristics and impact of smell and taste alterations. The first and second part was
351 completed by all participants, the third and fourth part was completed only by patients who reported
352 changes in smell and taste respectively.

353 **Laboratory data and sample processing**

354 Protocols and SOPs for sample processing were designed and updated according to the evolving
355 knowledge on SARS-CoV-2 diagnosis and clinical evolution.

356 **SARS-CoV-2 real-time quantitative Reverse Transcription PCR (qRT-PCR)**

357 Viral load kinetics and duration of viral shedding of SARS-CoV-2 RNA in bodily specimens of all origins
358 (respiratory, faeces, vaginal swab, urine, semen) were determined using an in-house protocol (in-
359 house qPCR) [49]. Nucleic acids were extracted from 190µl of sample material, in addition to 10µl of
360 internal control (phocine distemper virus (PDV)), using the NucliSense EasyMag or eMAG with the
361 Specific A protocol (bioMerieux, Lyon, France), according to the manufacturer's instructions. The in-
362 house qPCR targets the SARS-CoV-2 E gene following the description by Corman *et al* [50] with minor
363 modifications. Analysis is performed using a middleware software referred to as FlowG (LabHelp
364 Labautomation), which can be used for interpretation of the results and communicating them back to
365 the laboratory information system. Samples with a cycle time value (Ct) lower than 34 are considered
366 positive while a Ct of 34-39 is considered inconclusive and the sample is repeated, which is routine
367 clinical practice in our hospital for any new cases. Finally, Ct values above 40 are considered negative.
368 The protocol was adapted to target new SARS-CoV-2 viral variants.

369 **SARS-CoV-2 virus isolation and culture**

370 The presence of infectious virus in different SARS-CoV-2 positive samples (nasopharyngeal/ throat,
371 stool, urine, sperm/vaginal specimens) was ascertained through virus isolation in the BSLII facility of
372 the UMCG. Samples cleared were spun down and filtered through 0.2 µm filter and characterised by
373 a confirmatory RT-PCR positive test for COVID-19 and negative for a panel of other respiratory viruses.

374 The presence of infectious SARS-CoV-2 virus was monitored observing for cytopathic effect (CPE)
375 according to protocol for 7 days, and at day 2 and 7 using an in-house flow cytometry-based assay
376 detecting viral antigen. Samples were considered positive for the presence of infectious SARS-CoV-2
377 virus if the flow cytometry-based assay showed more than two standard deviations from the mock-
378 infected cells. Samples scoring positive in CPE but negative in flow cytometry-based assay were
379 considered negative.

380 **Anti-SARS-CoV-2 IgM, and IgG antibodies**

381 Serological measurements (objective 1.3) are performed on the Abbott Architect platform, using a
382 chemiluminescent magnetic microparticle immunoassay performed at the virology facility of the
383 UMCG. Samples obtained during the acute disease were tested with the Semi-quantitative Architect
384 IgM (AdviseDx SARS-CoV-2 IgM_{SP} assay), and Semi-quantitative Architect IgG (Abbott Alinity i SARS-
385 CoV-2 anti-nucleocapsid protein IgG assay), while long-term samples at 1, 3, 6, 12 & 18 months are
386 tested with the Quantitative Architect IgG (AdviseDx SARS-CoV-2 IgG II) [51]. These results will be
387 aligned per patient with the RT-PCR results, to evaluate the mean time points of IgM and IgG
388 seroconversion, the range of IgM and IgG titres, and if seroconversion is related to viral clearance. We
389 will determine the relation between Ct-values, viability of the virus and seroconversion. Patients are
390 followed up to 18 months post-infection to assess the potential waning of immunity, the effect of
391 vaccination and potential re-infections. The results of the serological testing are sent to the
392 participants by post.

393 **SARS-CoV2 whole-genome sequencing**

394 To fulfil objective 1.4, nucleic acid extraction, purification and concentration from positive samples
395 was performed after which, cDNA was prepared using a sequence-independent single-primer
396 amplification (SISPA) approach, based on the round A/B methodology. Amplified cDNA was purified
397 using a 1:0.8 ratio of AMPure XP beads (Beckman Coulter, Brea, CA), and quantified using a Qubit high-
398 sensitivity double-stranded DNA (dsDNA) kit (Thermo Fisher), both according to the manufacturers'

399 instructions. Multiplex sequencing libraries was prepared using 250 ng of cDNA from up to 5 samples
400 as input to the SQK-LSK109 and barcoded individually using the EXP-NBD104 Native barcodes (Oxford
401 Nanopore Technologies). Libraries were sequenced on FLO-MIN106 flow cells on the MinION Mk1b
402 device (Oxford Nanopore Technologies), with sequencing proceeding for up to 24 h.

403 **Phylogenetic analysis to determine viral diversity and potential chains of** 404 **transmission (Objectives 1.4 & 1.5)**

405 Consensus sequences will be aligned with relevant sequences from SARS-CoV-2 lineages using MAFFT
406 v7.313 [52]. Maximum likelihood phylogenies will be estimated using RAxML v8.2.12 (Stamatakis
407 2014) under a GTR-CAT nucleotide substitution. Viral lineage diversity and transmission clusters will
408 be evaluated from the estimated phylogenies.

409 **Haematological and biochemical blood parameters**

410 As part of objectives 2.1-2.4, blood samples were processed at the routine UMCG laboratory to
411 determine full blood count, haematocrit, C-reactive protein, creatinine, glycaemia, albumin, hepatic
412 transaminases (AST, ALT), bilirubin, creatine kinase (CK), CK-MB isoenzyme , LDH, Prothrombin time
413 (PT-INR), troponin, ferritin, procalcitonin, D-dimer, leptin and IL-6.

414 **Cytokine/chemokine profiling**

415 A panel of cytokines and chemokines (including IL-1 β , IL-6, IL-10, IL-12, and TNF α , and chemokines, in
416 particular CXCL10 and CCL2) is being measured using Luminex-based multiplex technology. Serum
417 samples and associated clinical and Luminex/ELISA data will be combined and shared between
418 different UMCG-COVID-19 research teams working on complementary studies addressing different
419 COVID-19-related research questions. These markers will be used to assess prediction of disease
420 severity and of the development of long-term symptoms.

421

422 **Data analysis**

423 Due to the exploratory and mostly descriptive nature of the analysis, we did not conduct an exact
424 sample size calculation. However, we focused on the prevalence of long-term sequelae as initial
425 observations indicated that close to half of the infected individuals continue to suffer from symptoms
426 after 4 weeks from infection [53, 54]. Therefore, we estimated that if the prevalence of Long COVID
427 in the outpatient infected population was 40%, then a sample size of 256 individuals would give 80%
428 power to estimate this prevalence with a +/- 3% margin of error. Given the intense follow-up, we
429 expected to enrol a minimum of 275 individuals, accounting for a 10% loss to follow-up, for a final
430 sample size of 250 participants. Data is analysed anonymously using STATA (Stata Statistical Software:
431 Release 14. College Station, TX,), R (version 3.6.3) and/or SPSS (IBM SPSS Statistics 23 (IBM Corp.,
432 Armonk, NY)) softwares.

433 ***Acute disease***

434 Clinical presentation during the acute disease will be firstly presented in a descriptive analysis where
435 the systematic prospective data collection will be used to better characterise mild clinical cases. We
436 will attempt to define clinical phenotypes as endpoints and factors related to each phenotype.
437 Principal component analysis-based cluster analysis will be used to group relevant variables of interest
438 such as age, gender, comorbidities, signs, symptoms, haematological and biochemical parameters into
439 phenotypes. Multinomial logistic regression models will be fitted to estimate associations of
440 phenotypes with exposure variables such as viral loads, genetic sequence of SARS-CoV-2, anti-SARS-
441 CoV-2 antibody kinetics, laboratory parameters and cytokine/chemokine levels. Clinical phenotypes
442 will be categorised from milder to more severe disease presentation and potential predicting factors
443 (measured at presentation) will be identified using generalized linear mixed models.

444 Generalised mixed models will be also used to assess the associations between the duration and
445 severity of signs/symptoms of COVID-19 disease and viral loads, anti-SARS-CoV-2 antibody kinetics

446 and change in laboratory parameters and cytokines, after adjusting for age and gender. The likelihood
447 ratio test will assess statistical significance of the estimates at the 5% level. Univariate and multivariate
448 analyses of potential risk factors for disease severity (mild versus more severe) will be performed using
449 logistic regression comparing crude and adjusted odds ratios. Variables with a P-value ≤ 0.2 after
450 adjusting by age and gender will be fitted into multivariate logistic regression models and adjusted for
451 further confounders. Effect modification will be analysed and resulting models compared by likelihood
452 ratio test.

453 A generalised linear regression model (Poisson distribution and log-link function) will be fitted to
454 analyse the relationship between the duration of viral/RNA shedding and the exposure variables (viral
455 loads, genetic diversity, antibody titres, laboratory parameters). Significance will be determined at the
456 5% level.

457 Phylogenetic analysis will be performed to determine viral diversity and potential chains of
458 transmission within HHs and into the community as explained under Rationale above.

459 ***Long COVID***

460 The presence of long-term sequelae at 3,6,12 & 18 months will be defined as reporting any symptoms
461 and/or not feeling fully recovered at each of these time-points, as well as assessing the health-related
462 quality of life using the SF-36 score at the different time points. These three definitions will be used
463 as endpoints. A descriptive analysis will be performed to compare symptom presentation between
464 time points and symptoms characteristics, such as number of symptoms, type of symptoms, symptom
465 presentation (recurrent, constant, or both), and severity. Univariate and multivariate analyses will be
466 carried out to assess risk factors for Long COVID from relevant variables of the acute disease such as
467 comorbidities, number and duration of symptoms, type of symptoms and socio-demographic
468 variables of interest. Analysis strategy will be as described above.

469 **Ethics, data management plan and biobanking**

470 This study has been approved by the Medical Ethical Review Committee of the UMCG (METC
471 2020/158) and follows international standards for the ethical conduct of research involving human
472 subjects. All procedures employed in the clinical and related laboratory studies comply with national
473 and European legislation in respect of research involving human subjects.

474 Informed consents including participant information sheets (available in Dutch and English) were
475 obtained before the collection of participants' data and samples. At enrolment, participants invited to
476 join the study received a written and oral explanation of what the research entails. Adult informed
477 consents were applied to individuals older than 16 years, while a parent/guardian informed consent
478 was signed by the child's parent/guardian and together with their child if aged between 12-16 years.
479 Individuals may withdraw from the study or may not consent to parts of the study, without any
480 consequences.

481 A detailed data management plan (DMP) has been prepared by the project data steward and approved
482 by the principal investigator describing how data will be handled both during research and after the
483 project is completed. The UMCG has a comprehensive ISO27001 certificate and is compliant covering
484 all of its strategic activities: Healthcare, Research and Education. The UMCG Research Code was drawn
485 based on this certification. The COVID HOME DMP complies with the UMCG quality policy, UMCG and
486 University of Groningen (RUG) Research Code and national guidelines on ethical research.

487 All information obtained during the course of this study, including hospital records, personal data and
488 research data will be kept strictly confidential according to the General Data Protection Regulation
489 (GDPR) of the Netherlands. Data protection and privacy is coordinated by the data protection officer
490 which supervises that research conducted within the UMCG is performed in accordance with the
491 GDPR. Study data were/are collected and managed using REDCap electronic data capture tools hosted
492 at UMCG. Electronic web-based eCRFs have been developed using REDCap (Research Electronic Data
493 Capture) which is a secure, web-based software platform designed to support data capture for
494 research studies [45]. Participants enter data online by means of an electronic data entry interface.
495 Data collected in paper CRFs were entered by research staff into the REDCap eCRFs. Inventory of

496 samples collected and related SOPs were/are recorded into eLABJournal (eLabNext, Bio-ITech BV), a
497 secured electronic lab notebook for biobanking purposes available only to dedicated UMCG
498 researchers. Clinical data and laboratory results were entered into electronic databases and merged
499 for analysis. To protect privacy, all data and sample records are entered as pseudonymized records
500 (using a study identification number), not allowing the personal identification of study participants.
501 Identifiable information is linked to stored data or samples only by a protected Master list.

502 All data (including clinical, laboratory and viral genetic data) are stored in password protected
503 databases and archived in institutional research drives secured under encryption by our UMCG
504 Information and Communication Technology department that can only be accessed by authorized
505 project personnel. Data storage is backed-up automatically every day on the servers of the UMCG
506 ensuring data safety. Samples are stored and managed locally according to UMCG SOPs. Patient
507 information sheets and informed consents are kept in a locked cabinet. The data will be stored at the
508 UMCG for 15 years. Data that may be reported in scientific journals will not include any information
509 that identifies the participants in this study. We follow a FAIR (Findable, Accessible, Interoperable and
510 Re-useable) data policy. After an embargo period of approximately 12 months (subject to change)
511 from the project culmination and publication of main results, the processed, pseudonymised data will
512 be made available for re-use. Requests for re-use of data will be evaluated by the principal
513 investigator, project manager and data steward who will check whether the research question falls
514 within the scope of the informed consent. If data will be shared it will not include directly identifiable
515 information and can only be (re-)used within the scope of the informed consent. Third party use of
516 data is governed in part by intellectual property rules and agreements, which will become available
517 after publication of the main results.

518 **Preliminary Results**

519 A total of 276 participants belonging to 108 HH were enrolled in the study by the end of May 2021.
520 Twenty (7.2%) were excluded since 16 had a false positive RT-PCR test result and 4 were non-compliant.

521 The final study population comprised 256 participants (103 HH), of which 30 (13 HH) were enrolled
522 during the first COVID-19 wave (March-May 2020) and 226 (90 HH) during the second and third waves
523 (October 2020 until May 2021). Of the final study population, 190 participants (74.2%) had a positive
524 RT-PCR test result for SARS-CoV-2 by the end of the 21-day acute disease follow-up period. The
525 proportion of SARS-CoV-2 infected individuals was fairly similar for both enrolment periods, 73.3%
526 (n=22) for the first wave versus 74.3% (n=168) for the second and third waves. Most individuals
527 (183/190, 96.3%) developed mild to moderate disease while 7 patients were hospitalised between 7-
528 9 days post-infection and one of these died.

529 At the time of writing this manuscript, all participants had reached the 3 and 6 month time-points of
530 the long-term follow-up with data and serological sampling being collected. However, due to
531 budgetary constraints, the first wave participants were not included in the 3 month follow-up.
532 Approximately 78% of participants have been sampled/data collected at 12 month and 23% at 18
533 month post-enrolment. Preliminary analysis has shown that 43% (52/121) positive individuals
534 reported having complaints at 3 months post-infection, while 42.7% (61/143) had complaints at 6
535 months.

536 Discussion

537 This study aims to close a knowledge gap on the impact and consequences of COVID-19 on non-
538 hospitalised patients, in contrast to most studies that naturally, have concentrated on hospitalised
539 and critically ill patients. Enrolment of participants was initiated soon after the first cases were
540 detected in the Netherlands [5], however, the north of the country experienced a low incidence during
541 the first wave of the pandemic. Therefore, most of the enrolment took place during the second and
542 third waves (October 2020 – June 2021) of COVID-19 in the Netherlands [55].

543 Our COVID HOME study addresses key questions of relevance to guide patient care and inform public
544 health guidelines for infection prevention and control in the community and HH situation in a single

545 study. The enrolment of patients right after diagnosis will allow the early identification of parameters
546 potentially indicative for progression to severe disease. This early identification cannot be achieved in
547 hospitalised individuals who are usually admitted a week or more after disease onset. The parameters
548 that will be collected in this study, such as viral load, antibody response, cytokine changes and clinical
549 and laboratory parameters of patient evolution will consent the identification of independent factors
550 associated with disease evolution, insight in other transmission routes, and immune response
551 dynamics in non-hospitalised patients. Moreover, as HH members of COVID-19 positive individuals
552 were also enrolled, we have been able to determine if other family members become infected
553 (especially children), how long does it take to get infected and which are the transmission dynamics
554 within the HH. It will be important to determine transmission chains between the HH and workplace of
555 adults, especially for healthcare settings to inform biosafety guidelines. A better understanding of the
556 clinical evolution of non-hospitalised patients and the factors influencing disease progress (viral,
557 immunological and other) can be rapidly turned into “home standard of care” guidelines for patients
558 in home isolation and their HH members, as well as shed light on potential new therapies and/or
559 strategies preventing progression to severe COVID-19 disease. Determining the length of
560 infectiousness of COVID-19 patients will either confirm current guidelines or prompt their change in
561 relation to quarantine and isolation measures, contributing to limit the spread of COVID-19 in the
562 Netherlands and elsewhere. Hence the importance of the collaboration with public health institutions
563 (National Institute for Public Health and the Environment [RIVM]/GGD) and policy makers.

564 In the Netherlands, the first indications of long-term consequences of SARS-CoV-2 infections (what is
565 now known as Long COVID or Post-COVID syndrome) appeared though a survey carried out by the
566 Dutch Lung Foundation in June 2020 [37]. We have found a high prevalence of SARS-CoV-2 infected
567 individuals reporting still having complaints at 3 and 6 months post-infection, similar to those found in
568 hospitalised patients [56]. However, there is uncertainty about the prevalence of Long COVID, the risk
569 factors for this condition in adults and children and the role of new variants. A recent review on post-
570 acute COVID-19 patients [38] highlight the importance of understanding the long-term evolution of

571 these patients and the identification of factors that could predict Long COVID manifestations. The
572 COVID HOME study is following participants up to 18 months after enrolment in order to contribute
573 knowledge on the impact of this incapacitating condition at individual and societal level.

574 **Acknowledgements**

575 We would like to firstly thank all the participants of the COVID HOME study for their ongoing
576 contribution to the project and to the understanding of COVID-19. Our thanks also go to the Clinical
577 Virology (Molecular and Serology) unit and to the Infection Prevention section, in special to Jet van
578 der Weerd (Head of Infection Prevention), all part of the Department of Medical Microbiology &
579 Infection Prevention, UMCG; as well as Lars van Heerden (manager) and the staff from the Laboratory
580 of General Haematology and Chemistry, Laboratory Medicine, UMCG; and to Myke Mol and Paul
581 Koenes from the Service Desk Clinical Research Office, UMCG, the Netherlands. Lastly, we would like
582 to thank the collaboration of the Municipal Public Health Services (GGD) of Groningen and Drenthe
583 with the recruitment of study participants in this region.

584

585 **References**

- 586 1. ZonMw. Prospective cohort study of non-hospitalised COVID-19 patients: determining length
587 of isolation and patient clinical development at home (COVID-HOME study): Netherlands
588 Organisation for Health Research and Development (ZonMw), 2020. Available from:
589 [https://www.zonmw.nl/nl/over-zonmw/coronavirus/programmas/project-detail/covid-19-](https://www.zonmw.nl/nl/over-zonmw/coronavirus/programmas/project-detail/covid-19-programma/prospective-cohort-study-of-non-hospitalised-covid-19-patients-determining-length-of-isolation-and/)
590 [programma/prospective-cohort-study-of-non-hospitalised-covid-19-patients-determining-](https://www.zonmw.nl/nl/over-zonmw/coronavirus/programmas/project-detail/covid-19-programma/prospective-cohort-study-of-non-hospitalised-covid-19-patients-determining-length-of-isolation-and/)
591 [length-of-isolation-and/](https://www.zonmw.nl/nl/over-zonmw/coronavirus/programmas/project-detail/covid-19-programma/prospective-cohort-study-of-non-hospitalised-covid-19-patients-determining-length-of-isolation-and/).
- 592 2. ORCHESTRA. Connecting European Cohorts to Increase Common and Effective Response to
593 SARS-CoV-2 Pandemic: ORCHESTRA, 2020. Available from: <https://orchestra-cohort.eu/>.
- 594 3. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with
595 Pneumonia in China, 2019. *N Engl J Med.* 2020;382(8):727-33.
- 596 4. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan,
597 China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med.* 2020;382(13):1199-207.
- 598 5. RIVM (National Institute for Public Health and the Environment). Patient with novel
599 coronavirus COVID-19 in the Netherlands. 2020 [cited 27-02-2020]. Available from:
600 <https://www.rivm.nl/node/152811>.
- 601 6. Coronavirus Dashboard - Deaths [Internet]. Rijksoverheid (Government of the Netherlands).
602 2022. Available from: <https://coronadashboard.government.nl/landelijk/sterfte>.
- 603 7. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical Characteristics of Coronavirus
604 Disease 2019 in China. *N Engl J Med.* 2020;382(18):1708-20.
- 605 8. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al.
606 Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized
607 With COVID-19 in the New York City Area. *Jama.* 2020;323(20):2052-9.
- 608 9. World Health Organization. Tracking SARS-CoV-2 variants 2022 [updated 3 May 2022].
609 Available from: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>.

- 610 10. Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, et al. Duration of
611 infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19,
612 England, January to May 2020. *Euro Surveill.* 2020;25(32).
- 613 11. Goncalves J, Torres-Franco A, Rodriguez E, Diaz I, Koritnik T, Silva PGD, et al. Centralized and
614 decentralized wastewater-based epidemiology to infer COVID-19 transmission - A brief
615 review. *One Health.* 2022;15:100405.
- 616 12. Tian Y, Rong L, Nian W, He Y. Review article: gastrointestinal features in COVID-19 and the
617 possibility of faecal transmission. *Aliment Pharmacol Ther.* 2020;51(9):843-51.
- 618 13. Laszkowska M, Faye AS, Kim J, Truong H, Silver ER, Ingram M, et al. Disease Course and
619 Outcomes of COVID-19 Among Hospitalized Patients With Gastrointestinal Manifestations.
620 *Clin Gastroenterol Hepatol.* 2021;19(7):1402-9 e1.
- 621 14. Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, et al. Detection of SARS-CoV-2 in Different Types of
622 Clinical Specimens. *JAMA.* 2020;323(18):1843-4.
- 623 15. Aslan MM, Uslu Yuvaci H, Kose O, Toptan H, Akdemir N, Koroglu M, et al. SARS-CoV-2 is not
624 present in the vaginal fluid of pregnant women with COVID-19. *J Matern Fetal Neonatal Med.*
625 2022;35(15):2876-8.
- 626 16. Pike JFW, Polley EL, Pritchett DY, Lal A, Wynia BA, Roudebush WE, et al. Comparative analysis
627 of viral infection outcomes in human seminal fluid from prior viral epidemics and Sars-CoV-2
628 may offer trends for viral sexual transmissibility and long-term reproductive health
629 implications. *Reprod Health.* 2021;18(1):123.
- 630 17. Khoiwal K, Kalita D, Kumari R, Dhundi D, Shankar R, Kumari R, et al. Presence of SARS-CoV-2
631 in the lower genital tract of women with active COVID-19 infection: A prospective study. *Int J*
632 *Gynaecol Obstet.* 2022;157(3):744-7.
- 633 18. Curmei M, Ilyas A, Evans O, Steinhardt J. Constructing and adjusting estimates for household
634 transmission of SARS-CoV-2 from prior studies, widespread-testing and contact-tracing data.
635 *Int J Epidemiol.* 2021;50(5):1444-57.

- 636 19. WHO. Working Group on the Clinical Characterisation Management of Covid-infection. A
637 minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis.*
638 2020;20(8):e192-e7.
- 639 20. Sahanic S, Tymoszuk P, Ausserhofer D, Rass V, Pizzini A, Nordmeyer G, et al. Phenotyping of
640 acute and persistent COVID-19 features in the outpatient setting: exploratory analysis of an
641 international cross-sectional online survey. *Clin Infect Dis.* 2021.
- 642 21. da Silva JF, Hernandez-Romieu AC, Browning SD, Bruce BB, Natarajan P, Morris SB, et al.
643 COVID-19 Clinical Phenotypes: Presentation and Temporal Progression of Disease in a Cohort
644 of Hospitalized Adults in Georgia, United States. *Open Forum Infect Dis.* 2021;8(1):ofaa596.
- 645 22. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133
646 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol:
647 prospective observational cohort study. *Bmj.* 2020;369:m1985.
- 648 23. Gutierrez-Gutierrez B, Del Toro MD, Borobia AM, Carcas A, Jarrin I, Yllescas M, et al.
649 Identification and validation of clinical phenotypes with prognostic implications in patients
650 admitted to hospital with COVID-19: a multicentre cohort study. *Lancet Infect Dis.*
651 2021;21(6):783-92.
- 652 24. Porto LC, Costa CH, Nunes AS, Bouzas I, Ferreira TF, Porto VM, et al. Clinical and laboratory
653 characteristics in outpatient diagnosis of COVID-19 in healthcare professionals in Rio de
654 Janeiro, Brazil. *J Clin Pathol.* 2022;75(3):185-92.
- 655 25. Li D, Jin M, Bao P, Zhao W, Zhang S. Clinical Characteristics and Results of Semen Tests Among
656 Men With Coronavirus Disease 2019. *JAMA Netw Open.* 2020;3(5):e208292.
- 657 26. Jeong HW, Kim SM, Kim HS, Kim YI, Kim JH, Cho JY, et al. Viable SARS-CoV-2 in various
658 specimens from COVID-19 patients. *Clin Microbiol Infect.* 2020;26(11):1520-4.
- 659 27. van Kampen JJA, van de Vijver D, Fraaij PLA, Haagmans BL, Lamers MM, Okba N, et al. Duration
660 and key determinants of infectious virus shedding in hospitalized patients with coronavirus
661 disease-2019 (COVID-19). *Nat Commun.* 2021;12(1):267.

- 662 28. Sohn Y, Jeong SJ, Chung WS, Hyun JH, Baek YJ, Cho Y, et al. Assessing Viral Shedding and
663 Infectivity of Asymptomatic or Mildly Symptomatic Patients with COVID-19 in a Later Phase. *J*
664 *Clin Med*. 2020;9(9).
- 665 29. To KK, Tsang OT, Leung WS, Tam AR, Wu TC, Lung DC, et al. Temporal profiles of viral load in
666 posterior oropharyngeal saliva samples and serum antibody responses during infection by
667 SARS-CoV-2: an observational cohort study. *Lancet Infect Dis*. 2020;20(5):565-74.
- 668 30. Xiang F, Wang X, He X, Peng Z, Yang B, Zhang J, et al. Antibody Detection and Dynamic
669 Characteristics in Patients With Coronavirus Disease 2019. *Clin Infect Dis*. 2020;71(8):1930-4.
- 670 31. Rijkers G, Murk JL, Wintermans B, van Looy B, van den Berge M, Veenemans J, et al.
671 Differences in Antibody Kinetics and Functionality Between Severe and Mild Severe Acute
672 Respiratory Syndrome Coronavirus 2 Infections. *J Infect Dis*. 2020;222(8):1265-9.
- 673 32. Dellicour S, Durkin K, Hong SL, Vanmechelen B, Martí-Carreras J, Gill MS, et al. A Phylodynamic
674 Workflow to Rapidly Gain Insights into the Dispersal History and Dynamics of SARS-CoV-2
675 Lineages. *Mol Biol Evol*. 2021;38(4):1608-13.
- 676 33. Worobey M, Pekar J, Larsen BB, Nelson MI, Hill V, Joy JB, et al. The emergence of SARS-CoV-2
677 in Europe and North America. *Science*. 2020;370(6516):564-70.
- 678 34. Sikkema RS, Pas SD, Nieuwenhuijse DF, O'Toole A, Verweij J, van der Linden A, et al. COVID-19
679 in health-care workers in three hospitals in the south of the Netherlands: a cross-sectional
680 study. *Lancet Infect Dis*. 2020;20(11):1273-80.
- 681 35. Rodríguez A, Ruiz-Botella M, Martín-Loeches I, Jimenez Herrera M, Solé-Violan J, Gómez J, et
682 al. Deploying unsupervised clustering analysis to derive clinical phenotypes and risk factors
683 associated with mortality risk in 2022 critically ill patients with COVID-19 in Spain. *Crit Care*.
684 2021;25(1):63.
- 685 36. Hariyanto TI, Japar KV, Kwenandar F, Damay V, Siregar JI, Lugito NPH, et al. Inflammatory and
686 hematologic markers as predictors of severe outcomes in COVID-19 infection: A systematic
687 review and meta-analysis. *Am J Emerg Med*. 2021;41:110-9.

- 688 37. Long Fonds. Peiling onder groep met Long Covid: driekwart nog niet volledig aan het werk -
689 2021 [cited 10-05-2021]. Available from: <https://www.longfonds.nl/nieuws/peiling-onder->
690 [groep-met-long-covid-driekwart-nog-niet-volledig-aan-het-werk](https://www.longfonds.nl/nieuws/peiling-onder-groep-met-long-covid-driekwart-nog-niet-volledig-aan-het-werk).
- 691 38. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute
692 COVID-19 syndrome. *Nat Med*. 2021;27(4):601-15.
- 693 39. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue
694 syndrome: a comprehensive approach to its definition and study. International Chronic
695 Fatigue Syndrome Study Group. *Ann Intern Med*. 1994;121(12):953-9.
- 696 40. Coperchini F, Chiovato L, Croce L, Magri F, Rotondi M. The cytokine storm in COVID-19: An
697 overview of the involvement of the chemokine/chemokine-receptor system. *Cytokine Growth*
698 *Factor Rev*. 2020;53:25-32.
- 699 41. Han H, Ma Q, Li C, Liu R, Zhao L, Wang W, et al. Profiling serum cytokines in COVID-19 patients
700 reveals IL-6 and IL-10 are disease severity predictors. *Emerg Microbes Infect*. 2020;9(1):1123-
701 30.
- 702 42. Fernandez-Castaneda A, Lu P, Geraghty AC, Song E, Lee MH, Wood J, et al. Mild respiratory
703 COVID can cause multi-lineage neural cell and myelin dysregulation. *Cell*. 2022;185(14):2452-
704 68 e16.
- 705 43. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 Viral Load in Upper
706 Respiratory Specimens of Infected Patients. *N Engl J Med*. 2020;382(12):1177-9.
- 707 44. CBS. Statistics Netherlands. StatLine-Bevolking, huishoudens en bevolkingsontwikkeling;
708 vanaf 1899. [updated 09-04-2021. Available from:
709 <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37556/table?ts=1646300534892>.
- 710 45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture
711 (REDCap)--a metadata-driven methodology and workflow process for providing translational
712 research informatics support. *J Biomed Inform*. 2009;42(2):377-81.

- 713 46. Parma V, Ohla K, Veldhuizen MG, Niv MY, Kelly CE, Bakke AJ, et al. More Than Smell-COVID-
714 19 Is Associated With Severe Impairment of Smell, Taste, and Chemesthesis. *Chem Senses*.
715 2020;45(7):609-22.
- 716 47. Miwa T, Furukawa M, Tsukatani T, Costanzo RM, DiNardo LJ, Reiter ER. Impact of olfactory
717 impairment on quality of life and disability. *Arch Otolaryngol Head Neck Surg*.
718 2001;127(5):497-503.
- 719 48. van Elst JM, NS IJ, Mathijssen RHJ, Steeghs N, Reyners AKL, de Haan JJ. Taste, smell and
720 mouthfeel disturbances in patients with gastrointestinal stromal tumors treated with
721 tyrosine-kinase inhibitors. *Support Care Cancer*. 2022;30(3):2307-15.
- 722 49. Gard L, Fliss MA, Bosma F, Ter Veen D, Niesters HGM. Validation and verification of the
723 GeneFinder COVID-19 Plus RealAmp kit on the ELITe InGenius(R) instrument. *J Virol Methods*.
724 2022;300:114378.
- 725 50. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel
726 coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill*. 2020;25(3).
- 727 51. Narasimhan M, Mahimainathan L, Araj E, Clark AE, Markantonis J, Green A, et al. Clinical
728 Evaluation of the Abbott Alinity SARS-CoV-2 Spike-Specific Quantitative IgG and IgM Assays
729 among Infected, Recovered, and Vaccinated Groups. *J Clin Microbiol*. 2021;59(7):e0038821.
- 730 52. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements
731 in performance and usability. *Mol Biol Evol*. 2013;30(4):772-80.
- 732 53. Chopra V, Flanders SA, O'Malley M, Malani AN, Prescott HC. Sixty-Day Outcomes Among
733 Patients Hospitalized With COVID-19. *Ann Intern Med*. 2021;174(4):576-8.
- 734 54. Garrigues E, Janvier P, Kherabi Y, Le Bot A, Hamon A, Gouze H, et al. Post-discharge persistent
735 symptoms and health-related quality of life after hospitalization for COVID-19. *J Infect*.
736 2020;81(6):e4-e6.

- 737 55. Coronavirus Dashboard. Number of confirmed cases. [Internet]. RIVM (National Institute for
738 Public Health and the Environment). 2022. Available from:
739 <https://coronadashboard.government.nl/landelijk/positief-geteste-mensen>.
- 740 56. Groff D, Sun A, Ssentongo AE, Ba DM, Parsons N, Poudel GR, et al. Short-term and Long-term
741 Rates of Postacute Sequelae of SARS-CoV-2 Infection: A Systematic Review. JAMA Netw Open.
742 2021;4(10):e2128568.

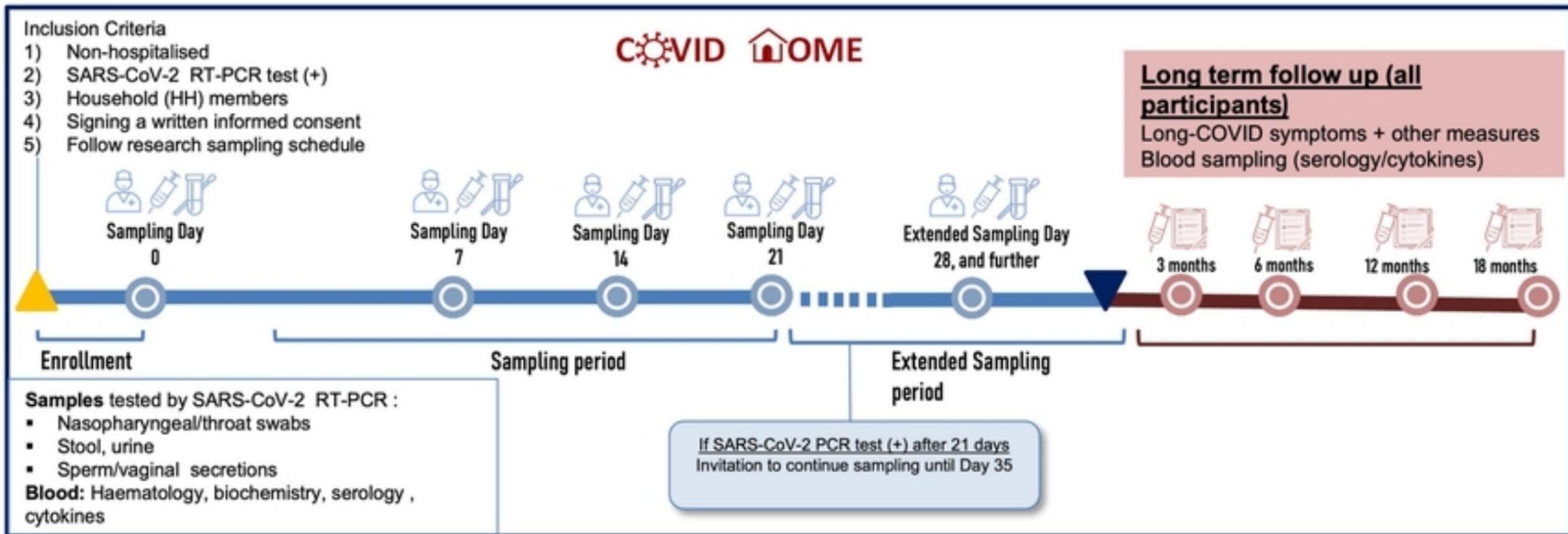


Figure 1

SARS-CoV-2 Positive

Adults (>16 years)

Children (≤16 years)

Day 0/1 Day 7 Day 14 Day 21 Day ...

Day 0/1 Day 7 Day 14 Day 21 Day ...

medRxiv preprint doi: <https://doi.org/10.1101/2022.08.14.22278762>; this version posted August 16, 2022. 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 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Samples
 1- NPT swab
 2- Blood sample
 3- Stool (faeces) & urine
 4- Semen or vaginal swab

Samples
 1- NPT swab
 2- Blood sample
 3- Stool (faeces) & urine

SARS-CoV-2 Negative

Adults (>16 years)

Children (≤16 years)

Day 1 Day 7 Day 14 Day 21

Day 1, 7, 14, 21 samples
 1- NPT swab
 2- Stool (faeces)

Day 1 & 21 samples
 Blood sample (Serology)

Figure 2

CRF-01: Enrollment data

CRF-02 : Daily symptoms follow-up for 21 days & comorbidities

CRF-03 : Socio-demographic and behavioural variables (applied to all)

CRF-04: Smell & taste alterations

Post-viral fatigue syndrome (PVFS):

eCRF-05 : extreme fatigue screening

eCRF-06 : extended fatigue

eCRF-07 : Health-related quality of life (SF-36)

eCRF-08 : Hospital anxiety and depression scale (HADS)

eCRF-10 : Long-term sequelae/symptoms assessment

eCRF-10a: modified version applied to uninfected individuals

Acute disease

Acute disease

Month 3 – 6 – 12

& 18

Figure 3