

1 **The prevalence of post COVID-19 condition (PCC) and a simple risk scoring tool for PCC**
2 **screening on Bonaire, Caribbean Netherlands: a retrospective cohort study.**

3 Danytza SF Berry^{1*}, Thomas Dalhuisen^{1*}, Giramin Marchena², Ivo Tiemessen³, Eveline Geubbels¹
4 Loes Jaspers²

5 *Joint first authorship.

6 Corresponding author: Loes Jaspers, loesjaspers@gmail.com

7 **Affiliations:**

8 ¹ National Institute for Public Health and the Environment, Centre for Infectious Disease Control,
9 Bilthoven, The Netherlands

10 ² Public Health Department Bonaire, Public Entity of Bonaire, Caribbean Netherlands

11 ³ Mobilito, Bonaire, Caribbean Netherlands

12

13 **Acknowledgements:**

14 We would like to thank Adrienn Vegh (General Practitioner) who served as a practical advisor on post
15 COVID-19 condition related primary care on Bonaire. Also, we would like to thank the team of
16 interviewers from Tempo, Bonaire, who carried out the telephone interviews, and CBS Bonaire for
17 their assistance in setting up training for the group of interviewers. Furthermore, we would like to
18 thank Alicia Xuan-Krijgsman (Public Health Department Bonaire) for assistance during the
19 preliminary phase of this study's design and Karin Cox (Public Health Department Bonaire) for
20 assisting in developing the call list of cases eligible for inclusion in the study.

21

22

23

24 **ABSTRACT**

25 **Aim:** To assess the prevalence of post COVID-19 condition (PCC) on Bonaire and develop a practical
26 risk scoring tool for PCC screening, using easily obtainable characteristics.

27 **Methods:** A retrospective cohort study of symptomatic SARS-CoV-2 cases were randomly sampled
28 from Bonaire's case-registry and telephone interviewed between 15-November-2021 and 4-December-
29 2021. PCC patients had a PCR-positive SARS-CoV-2 test (1-March-2020 and 1-October-2021) and
30 self-attributed at least one symptom lasting over four weeks to their infection. Multivariate logistic
31 regression was used to derive a risk formula to develop a practical risk scoring tool.

32 **Results:** Out of 414 cases, 160 (39%) were PCC patients. Fifty-three patients were unrecovered
33 (median illness duration 250 days (IQR 34)). Of recovered patients, 35% experienced symptoms for at
34 least 3 months after disease onset. PCC prevalence was highest among females (38%), 40-59 year-olds
35 (40%), morbidly obese (31%) and hospitalized patients (80%). A PCC risk scoring tool using age, sex,
36 presence of comorbidities, and acute phase hospitalization or GP visit had an area-under-the-curve
37 (AUC) of 0.68 (95%CI 0.63-0.74). Adding smoking, alcohol use, BMI, education level, and number of
38 acute phase symptoms increased the AUC to 0.79 (95%CI 0.74- 0.83). Subgroup analyses of non-
39 hospitalized patients (n=362) resulted in similar AUCs.

40 **Conclusion:** The estimated prevalence of PCC on Bonaire was 39%. Moreover, easily obtainable
41 patient characteristics can be used to build a risk scoring tool for PCC with acceptable discriminatory
42 power. After external validation, this tool could aid the development of healthcare interventions in low
43 resource settings to identify patients at risk for PCC.

44 **Keywords:** post COVID-19 condition; long-covid; Bonaire; risk-scoring tool; prevalence, cohort
45 study, Caribbean Netherlands

46 **Key messages:**

47 **What is already known on this topic:**

- 48 • An increasing number of studies show that varying proportions of COVID-19 patients are left
49 with prolonged health issues, including persistence of symptoms such as fatigue, shortness of
50 breath, loss of muscle strength, and concentration issues following the acute phase of COVID-
51 19.
- 52 • Prevalence of this recently recognized medical syndrome, ‘long-COVID’ or ‘post COVID-19
53 condition (PCC)’, ranges between 37-49% in the European Netherlands, though there are no
54 reports of PCC prevalence in the Caribbean Netherlands to date.

55 **What this study adds:**

- 56 • We designed the first retrospective cohort study estimating the prevalence and characteristics
57 of PCC on Bonaire, Caribbean Netherlands and devised a simple risk scoring formula to
58 calculate PCC risk.
- 59 • We found a PCC prevalence of 39% on Bonaire and describe a proof-of-concept risk scoring
60 tool with good discriminatory performance.

61 **How this study might affect research, practice or policy:**

- 62 • Our study estimates PCC prevalence and describes disease and patient characteristics for
63 Bonaire. This is the first study providing these insights in the Caribbean; a region that has
64 been underrepresented in PCC research.
- 65 • Furthermore, our study highlights the added benefit of registering sufficient patient
66 characteristics at the time of consultation for acute COVID-19 disease, for improved PCC
67 screening later on.
- 68 • After external validation of our proof-of-concept study, this risk scoring tool could aid the
69 development of primary care and public health interventions and health communication
70 strategies in low resource settings for the identification of patients at risk for PCC.

71 **INTRODUCTION**

72 As of 11 January 2023, an estimated 660,378,145 confirmed COVID-19 cases including 6,691,495
73 deaths have been reported worldwide [1]. Although most cases fully recover shortly after the acute
74 infection, an increasing number of international studies suggest that varying proportions of patients are
75 left with prolonged health issues [2]. This recently recognized medical syndrome, ‘long-COVID’ or
76 ‘post COVID-19 condition (PCC)’, includes persistence of symptoms like fatigue, shortness of breath,
77 loss of muscle strength, and concentration issues following the acute phase of COVID-19 [3-5]; much
78 like other post-infectious syndromes, where prolonged symptoms can occur in a remitting pattern due
79 to persistence of an infectious agent, such as Lyme disease or Q-fever [2, 5].

80 In van der Maaden et al. (2023), a longitudinal study of PCC in the European Netherlands, 49% of
81 COVID-19 cases reported symptoms three months after SARS-CoV-2 infection [6]. A study by Nivel
82 showed 37% of Dutch COVID-19 patients visiting the GP three months after disease onset reported still
83 experiencing at least one symptom [7]. These findings are roughly in line with a recent global pooled
84 PCC prevalence estimate of 43% (95% CI, 39-46%), though untangling symptom attribution solely to
85 SARS-CoV-2 infection remains challenging [2]. Furthermore, PCC prevalence appears higher among
86 those hospitalized (54%, 95% confidence interval (CI) 44-63%) compared to non-hospitalized patients
87 (34%, 95% CI 24-46%) [2].

88 Since the start of the COVID-19 pandemic, Bonaire has seen various waves of SARS-CoV-2 infections.
89 Bonaire is an island in the Dutch Caribbean with 21,745 permanent residents in 2021 [8]. By 1 October
90 2021, 2,036 persons living or staying on the island had tested positive, of whom 57 were admitted to the
91 hospital and 19 passed away due to or with COVID-19 [9]. In May 2021, local general practitioners
92 (GPs) on the island raised concerns about the lack of knowledge on the burden and risk factors of PCC
93 among Caribbean populations and the impact of PCC on health systems in Caribbean settings. Local
94 practitioners hypothesized the prevalence and symptomatology of PCC may differ substantially from
95 that reported in Western settings due to the different health system context, demographics, and burden
96 of non-communicable disease (NCD) on Bonaire [10-12]. Various studies have shown disparities
97 between different populations in COVID-19 risk and severity [13], though little remains known of how
98 these disparities translate into PCC outcomes and care [14, 15].

99 Besides gaining insights into the prevalence of PCC on Bonaire, practical and easy-to-use risk scoring
100 tools are needed to identify cases in low resource-settings that are at an increased risk of developing
101 PCC. Formal PCC diagnostics are lacking, and PCC consequently remains an underdiagnosed condition
102 even though timely diagnosis is needed to access appropriate insured care and treatment [15, 16]. Tools
103 that may identify those at an increased PCC risk can aid in increasing the proportion of PCC patients
104 that are formally diagnosed by healthcare providers in a timely matter. Ideally, these tools should require
105 only a limited number of inputs, which can be easily obtained by GPs. Therefore, in this study we aim

106 to assess the prevalence and characteristics of PCC on Bonaire and to develop a pragmatic risk scoring
107 tool using easily obtainable patient characteristics to identify patients at risk of developing PCC.

108 **METHODS**

109 **Study design & setting**

110 We designed a retrospective cohort study of individuals who tested positive for SARS-CoV-2 at the
111 public health department of Bonaire between the start of the pandemic and 1 October 2021
112 (**Supplementary Figure 1**). The public health department was the main facility on the island for
113 community testing. Testing was free of charge. The type of diagnostic tests used were in accordance
114 with the guidelines from the Dutch Ministry of Health at the time of testing and included approved PCR-
115 or antigen tests. The study protocol was developed through a collaboration between researchers from
116 the Public Health Department on Bonaire and the National Institute for Public Health and the
117 Environment (RIVM, in European Netherlands). Data collection took place on Bonaire, and data
118 analysis was carried out by researchers from the RIVM in close virtual collaboration with the researchers
119 based on Bonaire.

120 **Participants**

121 Participant data was exported from HP Zone, the patient registration and outbreak management software
122 system used to register data about the notifiable cases throughout the COVID-19 pandemic on Bonaire.
123 Individuals were eligible for recruitment if they were a) permanent residents of Bonaire; b) symptomatic;
124 c) had a laboratory confirmed SARS-CoV-2 positive test result, and d) had contact details available at
125 the public health department. All eligible participants were assigned a number: 0 for patients
126 hospitalized for COVID-19 during the acute phase or randomly the number 1, 2, 3 or 4 for non-
127 hospitalized patients. All patients hospitalized for COVID-19 during the acute phase were invited for
128 study participation. Thereafter, all the eligible participants in groups 1-4 were called, starting with
129 number 1, until enough participants were included to satisfy the required power threshold of the study.
130 This approach allowed us to incorporate a random element of sample selection and reaching enough
131 power for the study by calling all patients hospitalized for COVID-19 during the acute phase first. This
132 approach was chosen, given that on Bonaire, resources for the study completion were minimal, and we
133 wanted to assure maximum power for the study.

134 **Patient and public involvement**

135 Local GPs were provided the opportunity to help design the questionnaire and to formulate research
136 priorities. Preliminary results were shared with the public by Bonaire's Public Health Department
137 through local media (radio and television) and a press briefing in the spring of 2022. Moreover, the
138 results were discussed in a group session of health care and governmental professionals, including
139 representatives of the GPs, the hospital, the health insurance office, and policy advisors of the local
140 government. Remaining research needs were addressed, and where possible, integrated into the ongoing
141 analyses.

142 Furthermore, COVID-19 patients were able to self-enroll in the study if they were not called by the
143 interviewers, but felt the need to share their experiences. They were given a separate coding in the
144 questionnaire, in order to distinguish the randomly selected from the self-selected participants. In the
145 end, no patients opted to self-enroll in the study.

146 **Case definitions and community controls**

147 A PCC patient was defined as “an individual with a laboratory confirmed SARS-CoV-2 positive test
148 result, of whom at least one symptom self-attributed to the experienced SARS-CoV-2 infection lasted
149 longer than four weeks” [17]. Cases with a laboratory confirmed SARS-CoV-2 positive test result, of
150 whom all symptoms self-attributed to the experienced SARS-CoV-2 infection lasted shorter than four
151 weeks were defined as non-PCC cases. A community control group was included to assess background
152 prevalence of symptoms. By doing so, we aimed to have a stronger basis to differentiate whether certain
153 results could be attributable to pre-existing underlying disease, to a pandemic effect, or to the experienced
154 SARS-CoV-2 infection. During the telephone interview, SARS-CoV-2 positive cases were asked to
155 provide contact details of a community control, defined as “an individual who had not been a close
156 contact of the case during the infection phase, was of the same sex and similar age, was from outside
157 their household, and who they knew had not tested positive with SARS-CoV-2 until the moment of
158 interview”. It was expected that this approach would result in more controls as opposed to recruiting
159 individuals who had tested negative at the island’s testing facilities.

160 **Study size**

161 The expected study size was calculated based on a desired alpha of 0.05, power of 80%, expected
162 response rate of 70% and a conservatively estimated expected PCC prevalence of 20% at 4 weeks after
163 acute infection. Therefore, a sample of roughly 600 COVID-19 patients would be needed to estimate
164 the expected 20% prevalence of PCC with a 95% confidence interval of 6% width [17-23%]. In addition,
165 a cohort of 200 self-reported SARS-CoV-2 negative individuals was invited in order to be able to
166 compare the distribution of risk factors for and symptoms of PCC among COVID-19 patients with that
167 of the general population. In total, we aimed to invite 800 individuals to participate in the study.

168 **Data collection**

169 A questionnaire was designed based on 2021 PCC surveys from Nivel (Netherlands Institute for Health
170 Services Research), RIVM (Netherlands National Institute for Public Health and the Environment),
171 GGD Zuid Holland Zuid (Municipal Health Service South Holland), Radboud University Medical
172 Centre (UMC), and Maastricht & Hasselt University. Our questionnaire was developed for three groups:
173 those who did not have COVID-19 (community controls), those who did get infected with SARS-CoV-
174 2 but did not develop PCC (non-PCC cases), and those who developed PCC (PCC patients).

175 Survey questions focused on demographics, health status prior to the first outbreak, pre- and post-
176 infection COVID-19 vaccination status, hospitalization for acute COVID-19 symptoms, PCC symptoms

177 at the time of the interview, and healthcare utilization pre- and post-infection. The questionnaire was
178 reviewed for practical relevance by several local GPs. It was translated into the four main languages
179 spoken on Bonaire: Dutch, English, Papiamentu, and Spanish. Thereafter, the questionnaires were
180 uploaded into an online form using the data management program ‘digitale checklisten.nl’ [18]. It was
181 expected this approach would increase the response rate, taking local health literacy into consideration.
182 No back translation procedure took place due to time constraints.

183 In collaboration with the employment agency Tempo [19], we developed a job vacancy and hired a team
184 of 30 local interviewers. In November 2021, they were trained to carry out telephone interviews by the
185 research team of the Public Health Department and the Central Bureau of Statistics (CBS) Bonaire.

186 Eligible respondents were invited to participate in the study through a WhatsApp message, which
187 included information about the study and the informed consent statement (both in writing and as a voice
188 message recorded by a well-known Bonairean spokesperson). Verbal informed consent was obtained
189 and documented by the main interviewer prior to the start of the interview. Respondents were explained
190 the purpose of the study and that they could choose to resign from the interview at any time. A second
191 interviewer verified whether consent was given, prior to the first interviewer commencing the interview.
192 The second interviewer only confirmed whether the informed consent was given by the respondent. Data
193 was collected between 15 November and 4 December 2021 through telephone interviews of
194 approximately 20 minutes for community controls, 30 minutes for non-PCC cases, and 45 minutes for
195 PCC-patients. Eligible respondents were called inside and outside office hours to increase the likelihood
196 of being reached. Interviewers made three attempts to reach an eligible respondent, otherwise they were
197 registered as not reached on the call list. Data was directly entered into the online registry system
198 ‘digitale checklisten’ by the interviewer.

199 **Ethical approval**

200 The study protocol was sent to the Central Committee on Research Involving Human Subjects (CCMO
201 Netherlands) [20], who confirmed on 8 October 2021 that the study did not require ethical approval due
202 to its observational design. Patient data export necessary for participant selection was only extracted
203 from HP zone after permission from the Dutch COVID-19 patient data protection unit at GGD GHOR
204 – the Netherlands was given.

205 **Variables**

206 For this study, we defined PCC as “the presence of at least one of the following 14 symptoms minimally
207 4 after weeks of receiving a positive test result for SARS-CoV-2”, based on the Centers for Disease
208 Control (CDC) their staging at the time: chest pain, concentration problems, cough, fatigue, headache,
209 heart palpitations, loss of appetite, reduced muscle strength, loss of sense of smell, loss of sense of taste,
210 muscle ache, reduced physical endurance, shortness of breath, and sleeping problems [5,17].

211 Our second objective was to develop a pragmatic risk score formula using variables that could be easily
212 obtained by a GP in clinical practice on Bonaire. We selected predictors with the perspective of the data
213 collection being feasible in GPs' everyday practice. Hence, these included factors and data that GPs
214 register in the patient file with regard to the COVID-19 infection consultation, factors that are generally
215 available in the GP's electronic patient database, or factors which are easily accessible upon verbal
216 inquiry with the patient.

217 Demographic factors included age, sex (male/female), and education level. We categorized age into
218 groups 0-19 years, 20-39 years, 40-59 years, and 60+ years old. Education level was categorized as low,
219 moderate or high [21]. Low education indicated no education, primary level education, lower vocational
220 and lower secondary education; moderate indicated higher secondary education, and high indicated
221 university level post-secondary education. Health-related factors included pre-pandemic smoking
222 (yes/no), alcohol use (yes/no), BMI, and comorbidities. For BMI, we removed one outlier (>70), and
223 categorized remaining values into underweight (<18), normal ($\geq 18, <25$), overweight ($\geq 25- <30$), obese
224 ($\geq 30- <35$) and morbidly obese (≥ 35).

225 Comorbidity was categorized as the presence of at least one underlying disease (yes/no). Comorbidities
226 included specific underlying diseases with a diagnosis, such as diabetes, cardiovascular disease,
227 depression, neurocognitive disorders, stroke, and more. COVID-19 related factors included acute phase
228 GP visit, hospitalized for COVID-19 during the acute phase (not stratified by ICU or general care) and
229 number of symptoms present during acute phase. Definitions of variables considered in the analyses are
230 further specified in **Supplementary Table 1**.

231 **Statistical methods**

232 First, descriptive statistics and univariate analyses were conducted comparing PCC to non-PCC cases
233 and comparing the COVID-19 cohort to community controls using Fisher's exact test or Wilcoxon test
234 where appropriate. We considered a p-value <0.05 as statistically significant. All analyses were carried
235 out in R version 4.1.3 (R Foundation for Statistical Computing).

236 *Sparse and rich models*

237 Thereafter, we developed two sets of models using variables that could be easily obtained by a GP in
238 clinical practice on Bonaire. The first set is sparse and includes patient characteristics that, based on
239 discussions with a local GP, are easily attainable from a GP practice patient file. The variables
240 included in this model are age, sex, presence of one or more comorbidities, hospitalization for
241 COVID-19 during the acute phase and GP visit during the acute phase of a SARS-CoV-2 infection.
242 The second set is richer and includes additional variables not routinely registered in GP patient files
243 but which should be obtainable with relative ease upon a patient's visit or verbal inquiry by telephone:
244 BMI, lifestyle factors (such as whether a patient smokes or uses alcohol) and the number of symptoms

245 experienced during the acute phase of their SARS-CoV-2 infection. Because having been hospitalized
246 for acute COVID-19 was shown to be a strong predictor of PCC during explorative analyses, within
247 each set we built one model for the full patient sample (hospitalized and non-hospitalized during acute
248 COVID-19) and one model for non-hospitalized patients only. Thus, four models were built:

249 Models based on the full patient sample including both hospitalized and non-hospitalized patients
250 (n=390) included the following variables:

- 251 - Model 1a (sparse model): age, sex, presence of one or more comorbidities, acute phase GP visit,
252 and acute phase hospitalization;
- 253 - Model 2a (rich model): age, sex, presence of one or more comorbidities, BMI, smoking status,
254 alcohol use, education level, number of symptoms experienced during acute phase, acute phase
255 GP visit, and acute phase hospitalization.

256 Models based on the non-hospitalized patient sample (n=362) included the following variables:

- 257 - Model 1b (sparse model): age, sex, presence of one or more comorbidities, and acute phase GP
258 visit;
- 259 - Model 2b (rich model): age, sex, presence of one or more comorbidities, BMI, smoking, alcohol
260 use, education level, number of symptoms experienced during acute phase, and acute phase GP
261 visit.

262 *Development of risk score formula*

263 Multivariate logistic regression models were fit using PCC status as the outcome variable. Regression
264 coefficients from fitted models were then multiplied by 10 and rounded to the nearest integer to derive
265 a risk score formula (**Supplementary Tables 2 and 3**). The risk score formula was then applied to
266 calculate a risk score for each COVID-19 patient. The discriminatory performance of the risk score was
267 then assessed by plotting the receiver operating curve (ROC) and calculating the area under the curve
268 (AUC) and its 95% CI using 2000 stratified bootstrap replicates implemented in the pROC package in
269 R [22]. Discriminatory performance relates to how well the risk score can distinguish between PCC
270 patients and non-PCC cases. We then assessed the risk score formula's calibration and performed
271 internal validation using the optimism bootstrapping implemented in the rms R package
272 (**Supplementary Figure 2**) [23]. Using the sensitivity and specificity values for each risk score cut-off
273 in the ROC curves, we then calculated the positive predictive value (PPV) and negative predictive value
274 (NPV) of our risk score formula for each risk score cut-off. The PPV and NPV were calculated with the
275 formulas below using this study's PCC prevalence estimate of 39%:

$$276 \quad PPV = \frac{sensitivity * prevalence}{(sensitivity * prevalence) + ((1 - specificity) * (1 - prevalence))} * 100$$

277

$$278 \quad NPV = \frac{\text{specificity} * (1 - \text{prevalence})}{(1 - \text{sensitivity}) * \text{prevalence} + (\text{specificity} * (1 - \text{prevalence}))} * 100$$

279

280 ***Missing data***

281 Prior to fitting multivariate logistic regression models, we examined the extent of missing data for all
282 variables we aimed to include. We opted for a complete case-analysis, thereby excluding n=24 (5.8%)
283 cases from the COVID-19 cohort. Of excluded cases, we confirmed they did not have disproportional
284 baseline patient characteristics.

285 ***Sensitivity analysis***

286 We investigated whether an intermediate risk scoring tool could attain a similar AUC as our rich risk
287 scoring tool. This intermediate scoring tool uses fewer patient characteristics than our rich risk scoring
288 tool, but more than our sparse risk scoring tool. The variables included were: age, sex, presence of one
289 or more comorbidities, GP visit during the acute phase, alcohol use, smoking, and BMI (**Supplementary**
290 **Tables 4 and 5**).

291 **RESULTS**

292 **Participants**

293 All cases recorded in the national case registry ‘HP Zone’ by 1 October 2021 (n=2,036) were assessed
294 for eligibility. Based on the inclusion criteria, 1,789 individuals were eligible for study participation.
295 Of the excluded persons, 133 individuals were non-residents, 95 were asymptomatic and 19 had passed
296 away at the time of inclusion. Of the eligible individuals, 357 were not reached after three attempts, 76
297 did not have time to participate, 66 did not want to participate, 47 had switched phone numbers and
298 were not reachable anymore, 15 did not see a value in participating, four did not understand or trust
299 the interview procedure, and 73 had another, unspecified, reason not to participate.

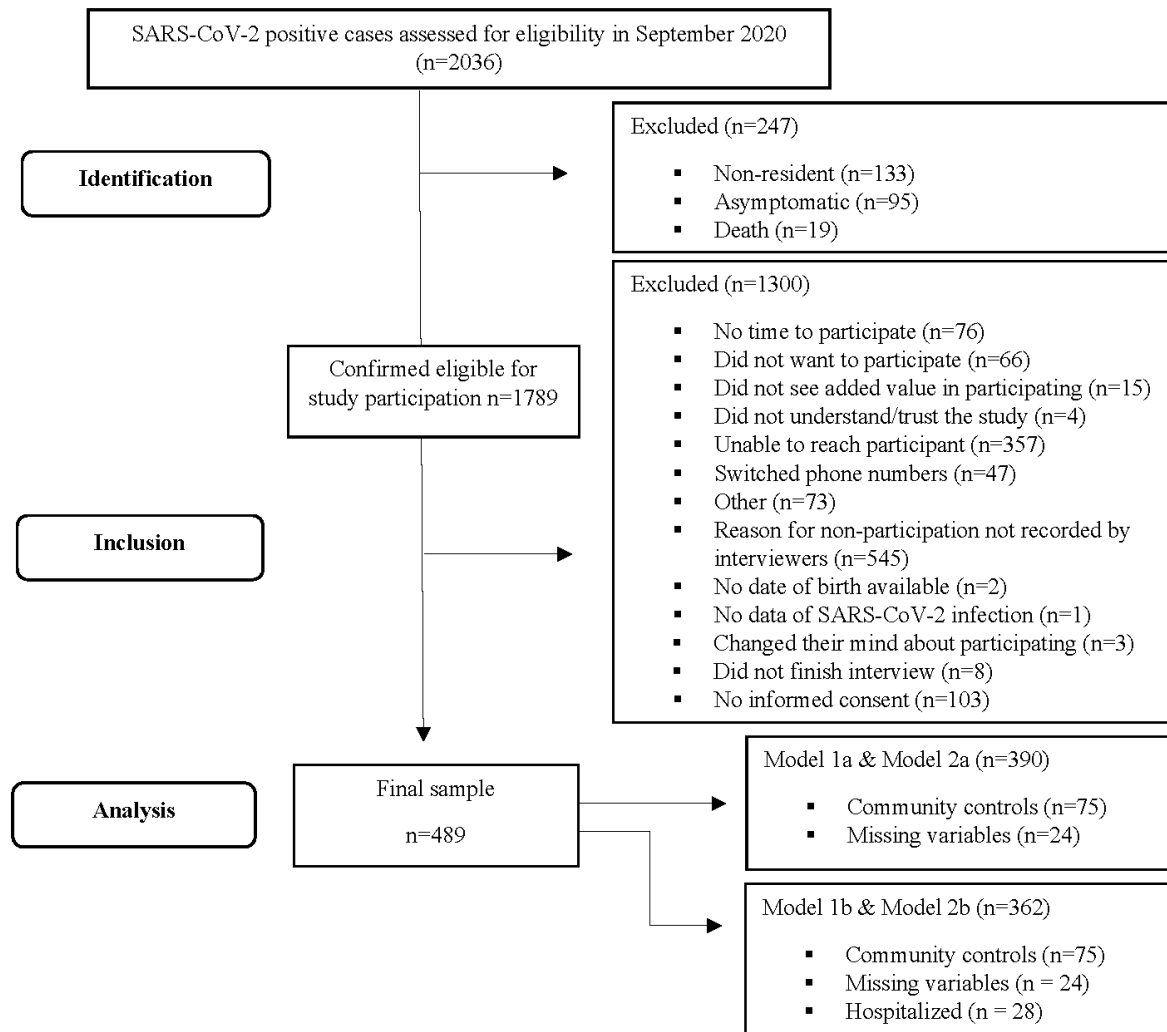
300 Of another 545 individuals, reason of non-participation was not recorded and 103 did not give
301 informed consent. Of interviewed individuals, 8 did not finish the interview, three changed their mind
302 about participating, one did not want to clarify whether they had had COVID-19, and of two, date of
303 birth was not registered. In total, 489 respondents were included in the study (**Figure 1**), of whom 414
304 were lab-confirmed SARS-CoV-2 positive cases and 75 had a self-reported SARS-CoV-2 negative
305 status (median age 43, IQR 26).

306

307 **Fig. 1** STROBE flowchart describing participant selection for the study examining prevalence and
 308 predictive factors of PCC on Bonaire, Caribbean Netherlands

309

310



311

312 Out of a total of 2,036 SARS-CoV-2 positive cases detected until October 1st, 2021, 133 cases were
 313 not eligible for inclusion because they were not a permanent resident of Bonaire, 95 because they were
 314 asymptomatic, and 19 were not eligible due to non-survival by 1 October 2021. Thus, there were 1,789
 315 cases confirmed eligible for inclusion. Of these, 57 cases had been hospitalized during the acute phase
 316 of infection. A total of 1,195 eligible participants were called during the interview period (67%), of
 317 whom 986 were reached (83%). The final sample consisted of 503 participants who had given
 318 informed consent (52%). Of these, 14 were excluded as they could not be categorized into one of the
 319 subgroups, for reasons such as missing date of onset or date of birth or choosing to stop participation
 320 during the interview. This flowchart has also been published in the sub-study by Berry et al. (2023)
 321 [‘manuscript in preparation’]

322 **Baseline characteristics**

323 160 cases (39%) fit the case definition of PCC (median age 43, IQR 21). At the time of interview, 107
 324 patients had recovered from PCC. Of recovered patients, 42 (39%) recovered within 2 months after
 325 disease onset, 27 (25%) recovered within 3 months, 16 (15%) within 4 months and 21 (20%) patients
 326 had not recovered after 4 months of disease onset. Time to recovery was unknown for 1 patient. 53
 327 patients were unrecovered at the time of interview and had a median duration of illness of 250 days
 328 (IQR 34).

329 In all three groups, there were more female participants than males (**Table 1**). PCC prevalence was
 330 higher among females (38%) than males (25%), highest in the age group 40-59 years old (40%),
 331 highest among morbidly obese patients (31%), and higher in cases hospitalized during the acute phase
 332 of infection (80%) as compared to the non-hospitalized group (35%).

333 Significant differences regarding sex, BMI, and age were observed in PCC patients compared to non-
 334 PCC cases, but these characteristics did not differ between the overall COVID-19 positive group (PCC
 335 and non-PCC cases combined) and community controls. This implies that our sampling method for
 336 community controls, aimed at a comparable distribution of main baseline characteristics between
 337 COVID-19 patients and community controls, was effective.

338 **Table 1** Baseline demographics and univariate analyses to assess between group differences

	PCC patients (n=160)	Non-PCC cases (n=254)	Overall positive cohort (n=414)	Community controls (n=75)	P value (PCC vs. non-PCC cases)	P value (Community controls vs. overall positive cohort)
Characteristics						
Median age (IQR)	43 (21)	39 (23)	40 (23)	43 (26)	<0.007^A	0.09 ^A
Age groups					<0.003^B	0.14 ^B
0-19	6 (4%)	36 (14%)	42 (10%)	3 (4%)		
20-39	61 (38%)	98 (39%)	159 (38%)	28 (37%)		
40-59	64 (40%)	85 (34%)	149 (36%)	25 (33%)		
60+	29 (18%)	34 (13%)	63 (15%)	19 (25%)		
Unknown	0	1 (0%)	1 (0%)	0		
Sex					<0.001^B	0.47 ^B
Male	46 (29%)	115 (45%)	161 (39%)	25 (33%)		
Female	114 (71%)	138 (54%)	252 (61%)	50 (67%)		
Non-binary	0	1 (0%)	1 (0%)	0		

Education level					0.15 ^B	0.78 ^B
Low	119 (74%)	208 (82%)	327 (79%)	57 (76%)		
Moderate	23 (14%)	29 (11%)	52 (13%)	12 (16%)		
High	17 (11%)	14 (6%)	31 (8%)	6 (8%)		
Unknown	1 (1%)	3 (1%)	4 (1%)	0		
Pre-pandemic factors						
Median BMI (IQR)	30.2 (11.3)	28.1 (8.71)	29.0 (9.38)	28.7 (8.32)	<0.001 ^A	0.78 ^A
BMI category					<0.001 ^B	0.92 ^B
Underweight	2 (13%)	4 (2%)	6 (1%)	1 (1%)		
Normal	32 (20%)	80 (32%)	112 (27%)	19 (25%)		
Overweight	41 (26%)	73 (29%)	114 (28%)	22 (29%)		
Obese	36 (23%)	53 (21%)	89 (21%)	19 (25%)		
Morbidly obese	49 (31%)	42 (17%)	91 (22%)	1 (1%)		
Unknown	0	2 (1%)	2 (0%)	0		
Comorbidities					0.05 ^B	0.40 ^B
Yes, one or more	33 (21%)	33 (13.0%)	66 (16%)	15 (0%)		
No	127 (79%)	221 (87.0%)	348 (84%)	60 (80%)		
Smoking					0.45 ^B	0.66 ^B
Yes	18 (11%)	36 (14%)	54 (13%)	14 (19%)		
No	142 (89%)	215 (85%)	357 (86%)	61 (81%)		
Unknown	0	3 (1%)	3 (1%)	0		
Alcohol use					0.29 ^B	0.73 ^B
Yes	109 (68%)	158 (62%)	267 (65%)	44 (59%)		
No	51 (32%)	94 (37%)	145 (35%)	30 (40%)		
Unknown	0	2 (1%)	2 (0.5%)	1 (1%)		
Factors during acute COVID-19						
Median number of symptoms	10 (5)	5 (7)	7 (8)	-	<0.001 ^A	-

present (IQR)						
GP visit					0.02^B	-
Yes	14 (9%)	8 (3%)	22 (5%)	-		
No	146 (91%)	233 (92%)	379 (92%)	-		
Unknown	0	13 (5%)	13 (3%)	-		
Hospitalization					<0.001^A	-
Yes	24 (15%)	6 (2%)	30 (7%)	-		
No	136 (85%)	235 (93%)	371 (90%)	-		
Unknown	0	13 (5%)	13 (3%)	-		

339 ^A Wilcoxon test. ^B Fisher's exact test.

340 Post COVID-19 Condition (PCC); Inter Quartile Range (IQR); Body Mass Index (BMI); General
341 Practitioner (GP).

342 **Symptomatology**

343 Among the SARS-CoV-2 positive cohort, presence of shortness of breath (OR=2.27 [1.07-4.94]) or
344 reduced physical endurance (OR=3.34 [1.13-11.09]) prior to infection were significantly associated
345 with PCC. All symptoms except chest pain were significantly more often experienced during the acute
346 phase by patients that developed PCC than patients who did not (**Supplementary Table 6**). Among
347 PCC patients, the most prevalent acute symptoms included fatigue (85%), reduced physical endurance
348 (83%), headache (78%), and reduced muscle strength (76%) (**Supplementary Table 7**). Of non-PCC
349 cases, 48% reported fatigue, 46% worsened physical endurance, 61% headache, and 39% reduced
350 muscle strength in the acute phase.

351 In the post-acute phase, fatigue (67%) and reduced physical endurance (66%) were the most persisting
352 symptoms. Persisting chest pain (37%) and sleeping problems (46%) were more prevalent among
353 female PCC patients (vs. 22% and 28% among male patients, respectively). Common symptoms
354 among patients over 60 years old included reduced physical endurance (79%), fatigue (76%) and loss
355 of appetite (76%). Concentration problems, reduced physical endurance, shortness of breath, and
356 reduced muscle strength were predominant among patients who had been hospitalized in the acute
357 phase, as opposed to patients who had not been hospitalized (**Supplementary Table 7**). There was no
358 significant association between presence of one or more comorbidities, sex, age group, or
359 hospitalization during the acute phase and each of the fourteen symptoms in the post-acute phase
360 (**Supplementary Table 7**).

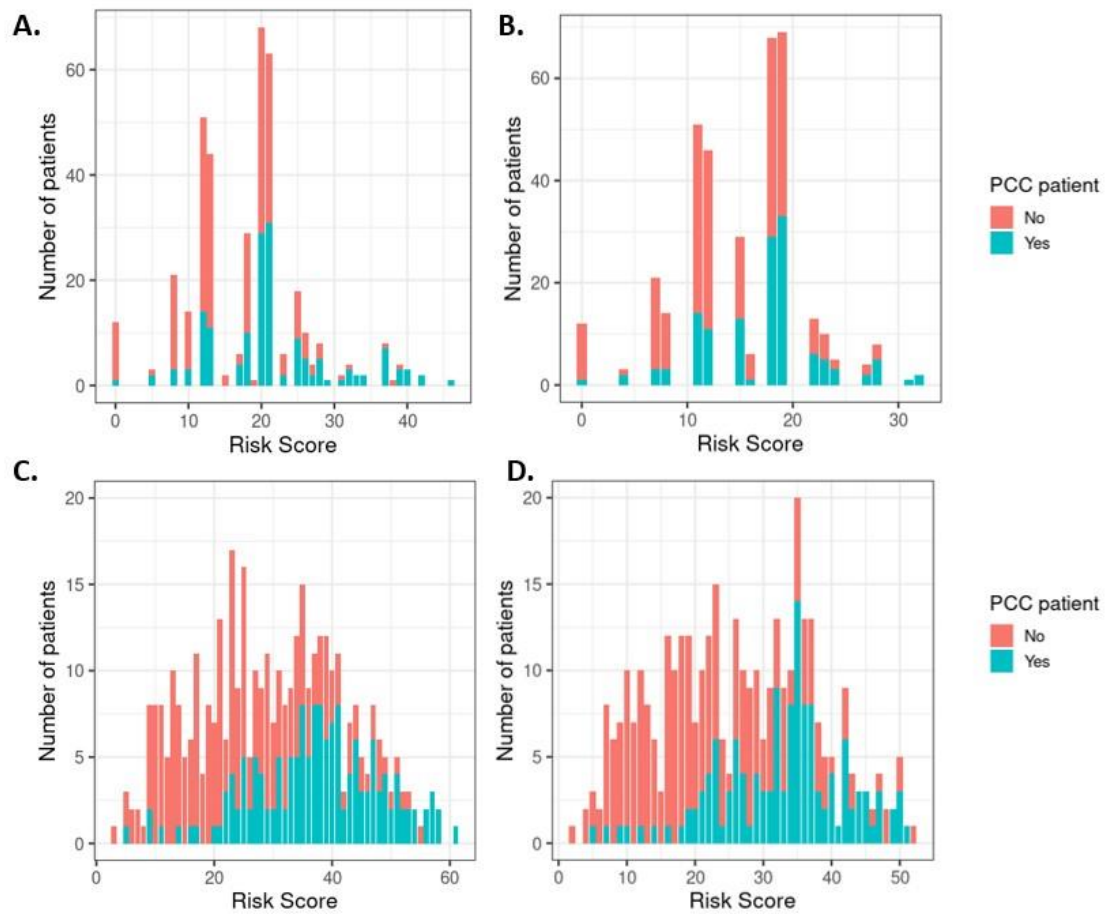
361 **PCC screening based on a simple risk score**

362 In total, 390 respondents were included in risk score analyses of which 40% (n =156) were PCC
363 patients. The distributions of the risk scores are shown in **Figure 2** and the risk score formulas used to
364 calculate a risk score per patient in **Supplementary Table 3**. Baseline characteristics for PCC patients

365 and non-PCC cases included are shown in **Supplementary Table 8**. The risk score ranges calculated
366 using the sparse risk score formulas are narrower compared to the risk score ranges calculated from
367 the rich risk score formulas (**Figure 2**). The sparse risk scoring tool based on the full patient sample
368 had an Area Under the Curve (AUC) of 0.68 (bootstrapped 95 CI% 0.63 – 0.74), whereas the sparse
369 risk scoring tool based on the non-hospitalized patient sample had an AUC of 0.65 (bootstrapped 95%
370 CI 0.59 – 0.71). The rich risk scoring tool based on the full patient sample had an AUC of 0.79 (95
371 CI% 0.74 – 0.83) and the rich risk scoring tool based on the non-hospitalized patient sample had an
372 AUC of 0.77 (95CI% 0.72 -0.82). ROC curves are shown in **Supplementary Figure 3**. The graphs in
373 **Figure 3** can be used to assess the trade-off between time and resource investment and identifying a
374 greater number of people at risk of developing PCC. Generally, a higher risk score cut-off results in a
375 higher PPV and lower NPV.

376

377 **Fig. 2** PCC risk score distributions



378

379 Histogram showing the distribution of PCC risk scores calculated with risk score formulas. (A) Risk
380 score distribution using the sparse risk score formula based on the full patient sample. (B) Risk score
381 distribution using the sparse risk score formula on non-hospitalized sample. (C) Risk score distribution
382 using the rich risk score formula based on the full patient sample. (D) Risk score distribution using the
383 rich risk score formula based on non-hospitalized sample.

384

385

386

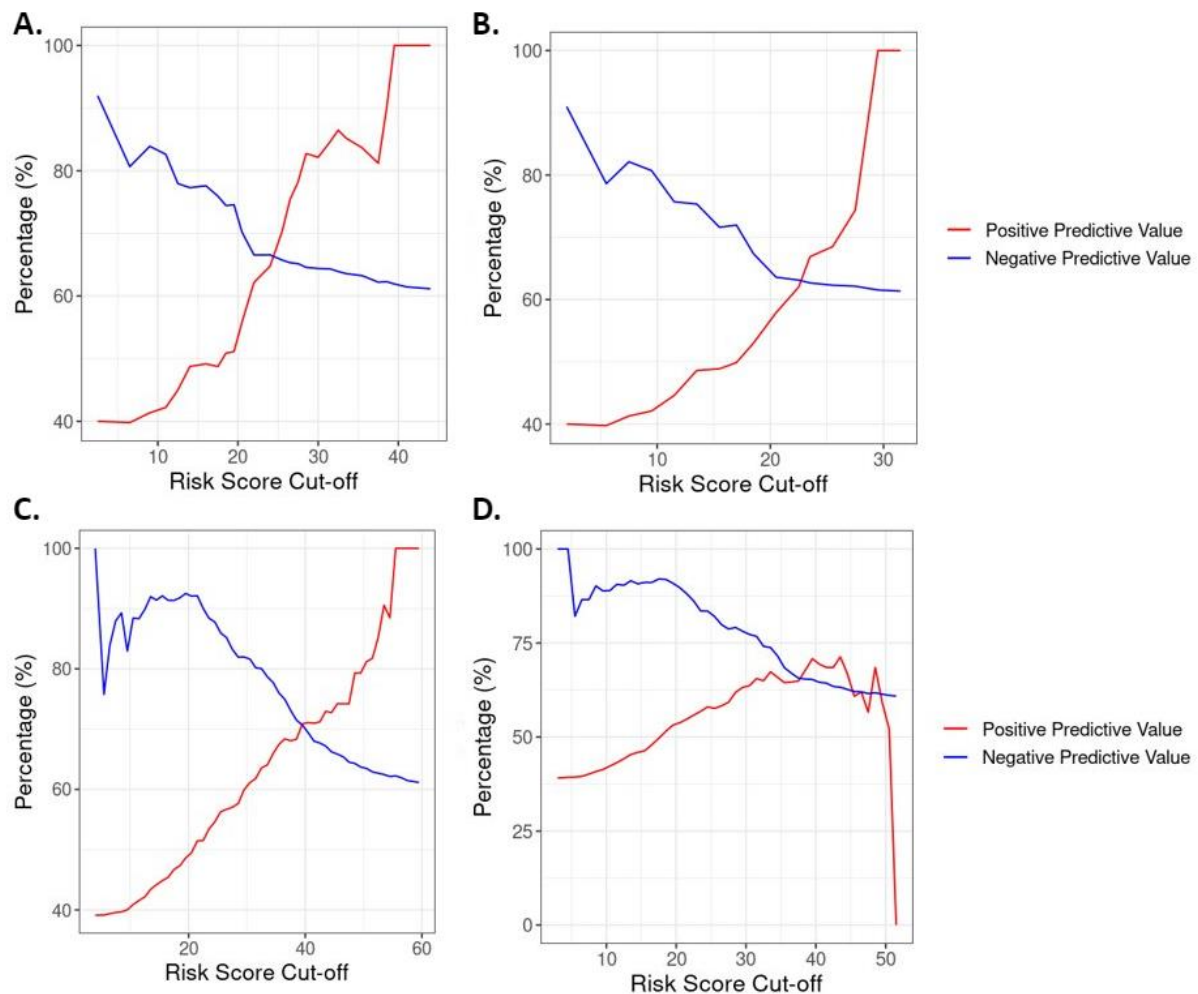
387

388

389

390

391 **Fig. 3** Positive and negative predictive values per risk score cut-off



392

393 Positive predictive value (PPV) and negative predictive values (NPV) per risk score cut-off. (A)

394 Sparse risk scoring tool based on full patient sample. (B) Sparse risk scoring tool based on non-

395 hospitalized sample. (C) Rich risk scoring tool based on full patient sample. (D) Rich risk scoring tool

396 based on non-hospitalized sample.

397

398 **Worked example of using risk score for PCC screening**

399 We provide worked examples for two scenarios in **Table 2**, using the graphs in **Figure 3**. The first

400 scenario reflects a low-resource setting, where a high PPV and high risk score cut-off are desirable given

401 the limited number of people that would have to be contacted. All those who would be contacted in this

402 scenario would have a high probability of being true PCC patients. A disadvantage, however, is that in

403 this scenario many other true PCC patients may miss out on being contacted. Contrarily, in the second,

404 high resource scenario, a high NPV and low risk score cut-off may be desirable. Here, many patients
 405 would be contacted which would increase the number of false positives; however most true PCC patients
 406 would be reached.

407 To illustrate the above points, we provide worked examples based on the rich risk scoring tool based on
 408 the full sample and the rich risk scoring tool based on the non-hospitalized sample for a low- and high-
 409 resource setting. Taking the low-resource setting as an example, our aim would be to identify as many
 410 true PCC patients whilst minimizing capacity going towards contacting efforts. Therefore, a high risk
 411 score cut-off and PPV is desirable; yet care must be taken not to choose a risk score that is at the tail
 412 end of the distribution, given few people will have such a score (**Figure 2**). Based on **Figure 3**, we chose
 413 a risk score of 40 with a PPV around 70%. Applying the risk scoring tool to our data, we found that 92
 414 patients had a risk score of 40 or greater, of whom 66 (72%) were true PCC patients, yet 90 other true
 415 PCC patients would be missed as they would not be contacted.

416 **Table 2** Number of patients who would be contacted (Call +) who truly have or do not have
 417 PCC for different risk score cut-offs

418 **A. Low-resource setting**

	PCC +	PCC -	Total
Hospitalized and non-hospitalized PCC patients ^a			
Call +	66	26	92
Call -	90	208	298
	156	234	390
Non-hospitalized PCC patients ^b			
Call +	64	36	100
Call -	70	192	262
	134	228	362

419 ^a Risk score cut-off 40, using our rich risk scoring formula based on full patient sample.

420 ^b Risk score cut-off 35, using our rich risk scoring formula based on non-hospitalized sample.

421 **B. High resource setting**

	PCC +	PCC -	Total
Hospitalized and non-hospitalized PCC patients ^a			
Call +	151	183	334
Call -	5	51	56
	156	234	390
Non-hospitalized PCC patients ^b			
Call +	124	119	243
Call -	70	109	119
	134	228	362

422 ^a Risk score cut-off 35, using our rich risk scoring formula based on full patient sample.

423 ^b Risk score cut-off 20, using our rich risk scoring formula based on non-hospitalized sample.

424 **Sensitivity analysis**

425 Our intermediate risk scoring tool for the full patient sample had an AUC of 0.71 (95% CI, 0.65 –
426 0.76) and the AUC was 0.67 (95% CI, 0.61-0.73) for the non-hospitalized patient sample
427 (**Supplementary Table 5**).

428 **DISCUSSION**

429 PCC is a debilitating condition and evidence regarding prevalence, symptomatology, and risk
430 screening tools are lacking in the Caribbean context. In our study, 39% of COVID-19 cases on
431 Bonaire experienced at least one persisting symptom for at least four weeks after disease onset. Most
432 common persisting symptoms included fatigue (67%) and reduced physical endurance (66%).
433 Additionally, we describe proof-of-concept risk scoring tools using easily obtainable patient
434 characteristics with good discriminatory power, which, after external validation in other study
435 populations, could aid in the timely identification of those at increased risk of PCC. This is important
436 because long-COVID rehabilitation care on Bonaire needs to be applied for within a specified time
437 frame in order to be eligible for health insurance coverage [16].

438 **Prevalence**

439 Symptom attribution to SARS-CoV-2 infection has shown to be challenging and there is substantial
440 heterogeneity in prevalence estimates both within and between studies. This is caused by differences
441 in case definition (such as duration of illness and type of symptoms) and differences in region,
442 characteristics, and hospitalization status of the study populations [2, 6, 24]. For example, our estimate
443 of 39% of symptomatic acute COVID-19 cases having symptoms for more than 4 weeks is in line with
444 a recent global PCC pooled prevalence estimate of 43% (95% CI, 39-46) [2], however, it is lower than
445 an estimate from a prospective cohort study in the European Netherlands of 49% of cases reporting at
446 least one symptom three months after SARS-CoV-2 infection [6].

447 Our PCC case definition is based on the (CDC) working definition at the time of interview, using a
448 cutoff of at least four weeks after disease onset [17]. The WHO case definition at the time of writing,
449 on the other hand, indicates that a symptom should persist for at least 12 weeks after infection, of
450 which at least 8 weeks cannot be explained by an alternative diagnosis [25]. This difference in follow-
451 up time is important as its been shown to affect PCC prevalence estimates. Results of a meta-analysis
452 analyzing global PCC estimates by Chen et al. (2022) reveal that PCC prevalence was 37 % when
453 using a 30 day follow-up definition (95% CI, 26-49%), which decreased to 25% (95% CI, 15-38%)
454 when a 60 day follow-up definition was used [2]. Surprisingly, prevalence across studies increased
455 again to 32% (95% CI, 14- 57%) and 49% (95% CI , 40-59%) using follow-up definitions of 90 and
456 120 days; possibly due to oversampling of hospitalization patients in the study populations of studies
457 with longer follow up [2]. Using the WHO case definition, we observed a PCC prevalence 23% among
458 our sample.

459 Consistent with other studies, we found that fatigue (67%) was the most common PCC symptom,
460 followed by reduced physical endurance (66%) [26]. Among a sample of European Dutch and Belgian
461 patients, Goertz et al. (2020) found fatigue (87%) and shortness of breath (71%) were most prevalent
462 PCC symptoms [27]. Van der Maaden et al. (2023) found a significantly higher prevalence of fatigue
463 (31.1%), dyspnea (16.4%), concentration problems (15.0%) and several other symptoms among cases
464 from the European Netherlands as compared to control groups [6].

465 **Risk scoring tool**

466 Besides gaining insight into PCC prevalence on Bonaire, practical and easy-to-use risk scoring tools
467 for public and primary care are needed in low resource-settings to screen and identify persons that are
468 at risk of developing PCC. Few PCC prediction tools exist and the ones that do, do not necessarily
469 have practicality of use in mind – making them less usable for low-resource settings – nor are they
470 applied to Caribbean populations. Our sparse risk scoring tool, using patient characteristics that are
471 easily obtainable by GPs in practice (i.e. age, sex, presence of one or more comorbidities,
472 hospitalization or GP visit for acute COVID-19); resulted in moderate discriminative power for the
473 prediction of PCC. Adding smoking, alcohol use, BMI, education level, and number of symptoms in
474 the acute phase improved the discriminative power considerably in the full patient sample
475 (hospitalized plus non-hospitalized) as well as for the non-hospitalized patient sample.

476 Our results highlight the importance of obtaining certain patient characteristics for designing risk scoring
477 tools either at intake or upon verbal inquiry with the patient. The moderate discriminatory performance
478 of our sparse risk scoring tools – suggest that solely using baseline characteristic is not optimal in
479 identifying patients at risk for PCC. Our rich risk scoring tools, on the other hand, showed a marked
480 increase in discriminatory performance allowing for better delineation of PCC patients. Importantly, our
481 risk scoring tools excluding hospitalized patients still obtained a good predictive performance aiding in
482 the identification of PCC patients, which in absolute numbers is the largest patient population at risk for
483 PPC. Sensitivity analysis showed a moderate performance of our intermediate risk scoring tool when
484 including hospitalized patients. However, when excluding hospitalized patients, its discriminatory
485 performance was poor-to-moderate and similar to our sparse risk scoring tool. This indicates that
486 education level, and in particular the number of symptoms during the acute phase, are important
487 predictors for PCC risk in the non-hospitalized population.

488 Although our rich risk scoring tools had a better discriminatory performance, calibration plots also show
489 they overpredict PCC risk at higher predicted probabilities and underpredict at lower predicted
490 probabilities (**Supplementary Figure 2**). This behavior is not necessarily problematic if risk scores are
491 used as a PCC screening tool. In practice overoptimistic PCC risk prediction will result in a greater
492 number of false positives. Although more false positives in this case would result in an increased use of

493 capacity to contact patients, there would be negligible negative consequences for patients, as they could
494 simply inform their GP whether they are experiencing persisting symptoms or not.

495 **Strengths & Limitations**

496 We designed the first retrospective cohort study estimating the prevalence and characteristics for PCC
497 in the Dutch Caribbean context and devised a simple risk scoring formula to estimate PCC risk. Our
498 study adds valuable insight into PCC prevalence for the Caribbean region, which has been
499 underrepresented in PCC research [15]. A major strength of our study was that we included a negative
500 control group, age-sex matched to the COVID-19 cases, to assess baseline health status, and only few
501 studies have done this thus far [24]. We found no significant difference in baseline demographic
502 characteristics between the COVID-19 cohort and community controls, allowing an unbiased
503 comparison of pre-existing health.

504 Contrary to most studies [24], we collected data on symptoms present before the SARS-CoV-2 infection
505 for PCC patients, non-PCC cases, and community controls. Prior to the pandemic, symptoms were
506 experienced to a similar or slightly higher degree by PCC patients than community controls, though we
507 had limited statistical power to assess the significance of these differences. However, considering the
508 vastly higher prevalence of symptoms post COVID-19 among PCC patients compared to their pre-
509 pandemic experience and compared to the community controls, it is not likely that all PCC
510 symptomatology can be explained as previously existing symptoms now being labeled as post-acute
511 symptoms.

512 We did find that COVID-19 patients who reported already experiencing shortness of breath (OR=2.27
513 [1.07-4.94]) and reduced physical endurance (OR=3.34 [1.13-11.09]) prior to infection were more likely
514 to suffer from post-acute symptoms than COVID-19 patients without these two pre-existing symptoms.
515 A recent study using the European Netherlands Lifelines cohort data matched SARS-CoV-2 positive
516 cases with a control group of SARS-CoV-2 negative participants and found that one in eight COVID-
517 19 cases had persistent symptoms for 90-150 days after their acute COVID19 infection that could be
518 attributed to their infection [2]. Our inclusive sampling strategy, i.e., inviting all residents on Bonaire
519 who had received a SARS-CoV-2 positive test result before 1 October 2021 and experienced a
520 symptomatic infection to participate in the study, supports the generalizability of our findings to the
521 COVID-19 patient population on Bonaire.

522 As community controls were included based on self-reported SARS-CoV-2 negative status, PCC
523 prevalence may be affected if these respondents were instead untested, SARS-CoV-2 positive cases. We
524 opted to use patient reported data rather than using validated instruments for patient characteristics,
525 previously registered patient data, or diagnostic tools for underlying disease, which would require more
526 capacity and resources to obtain. Using self-reported data on presence of these symptoms may have led
527 to information bias, though we believed our approach would result in more data considering the low-

528 resource setting of our study. We did not expect potential information bias to impact our results, as we
529 assumed this would not differ across subgroups. It is important to note that our risk score formula has
530 not been externally validated. Our risk score formulas should therefore be interpreted as a proof-of-
531 concept, and external validation would be needed before they can be utilized for PCC screening in
532 general practice or public health settings.

533 **Implications for research, clinicians and policymakers**

534 Our study adds valuable insight into PCC prevalence for the Caribbean region, which has been an
535 underrepresented region for PCC research. Since the start of the pandemic, Bonaire reported 9,822
536 SARS-CoV-2 cases, of whom 98.5% concerned non-hospitalized cases with mild or asymptomatic
537 infections (up to 8 January 2023) [9]. Applying our PCC prevalence of 39%, we expect that a
538 substantial group of (former) COVID-19 patients may be experiencing persisting symptoms without
539 seeking care at their GP, leaving the condition undiagnosed or with vast diagnostic delay among the at
540 risk population [28]. GPs will require adequate tools to identify at risk patients, preferably using easy
541 to use screening tools. Here, we highlight the added benefit of obtaining extra patient characteristics
542 either at intake or upon verbal inquiry for improved PCC screening and stress the need for further
543 research into external validation of risk scoring formulas.

544 **CONCLUSION**

545 We estimated a PCC prevalence of 39% on Bonaire four weeks after disease onset. Moreover, easily
546 obtainable patient characteristics can be used to build a risk scoring tool with moderate-to-good
547 discriminatory power to identify those with PCC. If used for PCC screening, our risk scoring tool
548 could aid public health in low-resource settings, however it needs to be externally validated first.

549 REFERENCES

- 550 1. World Health Organization. WHO coronavirus (COVID-19) dashboard [online]. 2023.
551 Available from: <https://covid19.who.int/> (accessed 12 January 2023).
- 552 2. Chen C, Hauptert SR, Zimmermann L, et al. Global Prevalence of Post-Coronavirus Disease
553 2019 (COVID-19) Condition or Long COVID: A Meta-Analysis and Systematic Review. *J*
554 *Infect Dis* 2022;226:1593-1607; <https://doi.org/10.1093/infdis/jiac136>
- 555 3. Aiyegbusi OL, Hughes SE, Turner G, et al. Symptoms, complications and management of
556 long COVID: a review. *J R Soc Med* 2021;114:428-442;
557 <https://doi.org/10.1177/01410768211032850>
- 558 4. Banda JM., Adderley N, Ahmed W, et al. Characterization of long-term patient-reported
559 symptoms of COVID-19: An analysis of social media data. *MedRxiv* 2021.07.13.21260449
560 [Preprint]. 2021; <https://doi.org/10.1101/2021.07.13.21260449>
- 561 5. Pergolizzi, JV, LeQuang JA, Magnusson P, et al. It's not over till it's over: a narrative review
562 of Long COVID. *Signa Vitae* 2021;17:21-30; <https://doi.org/10.22514/sv.2021.067>
- 563 6. van der Maaden T, Mutubuki EN, de Bruijn S, et al. Prevalence and Severity of Symptoms 3
564 Months After Infection With SARS-CoV-2 Compared to Test-Negative and Population
565 Controls in the Netherlands. *J Infect Dis*. 2023;227(9):1059-1067;
566 <https://doi.org/10.1093/infdis/jiac474>
- 567 7. Bosman L, van den Hoek R, van Waarden W, et al. Het Post-COVID-syndroom: hoe
568 definiëren we het en hoe vaak komt het voor? [online]. 2022. Available from:
569 [https://www.nivel.nl/nl/publicatie/het-post-covid-syndroom-hoe-definieren-we-het-en-hoe-](https://www.nivel.nl/nl/publicatie/het-post-covid-syndroom-hoe-definieren-we-het-en-hoe-vaak-komt-het-voor)
570 [vaak-komt-het-voor](https://www.nivel.nl/nl/publicatie/het-post-covid-syndroom-hoe-definieren-we-het-en-hoe-vaak-komt-het-voor) (accessed 16 August 2022).
- 571 8. Central Bureau of Statistics (CBS). Caribbean Netherlands; population, births, deaths,
572 migration. 2022. Available from: [https://www.cbs.nl/en-](https://www.cbs.nl/en-gb/figures/detail/83774ENG?q=bonair)
573 [gb/figures/detail/83774ENG?q=bonair](https://www.cbs.nl/en-gb/figures/detail/83774ENG?q=bonair) (accessed 16 August 2022).
- 574 9. National Institute for Public Health and the Environment. COVID-19 weekly epidemiological
575 update for the Caribbean part of the Kingdom [online]. 2021. [https://www.rivm.nl/en/novel-](https://www.rivm.nl/en/novel-coronavirus-covid-19/current-information/covid-19-monthly-epidemiological-update-for-caribbean-part-of-kingdom)
576 [coronavirus-covid-19/current-information/covid-19-monthly-epidemiological-update-for-](https://www.rivm.nl/en/novel-coronavirus-covid-19/current-information/covid-19-monthly-epidemiological-update-for-caribbean-part-of-kingdom)
577 [caribbean-part-of-kingdom](https://www.rivm.nl/en/novel-coronavirus-covid-19/current-information/covid-19-monthly-epidemiological-update-for-caribbean-part-of-kingdom) (accessed 15 October 2021).
- 578 10. United Nations Office for Disaster Risk Reduction. Regional Office for the Americas and the
579 Caribbean Economic Commission for Latin America and the Caribbean. The coronavirus
580 disease (covid-19) pandemic - an opportunity for a systemic approach to disaster risk for the
581 Caribbean [online]. 2021. [https://www.undrr.org/publication/undrr-eclac-report-coronavirus-](https://www.undrr.org/publication/undrr-eclac-report-coronavirus-disease-covid-19-pandemic-opportunity-systemic-approach)
582 [disease-covid-19-pandemic-opportunity-systemic-approach](https://www.undrr.org/publication/undrr-eclac-report-coronavirus-disease-covid-19-pandemic-opportunity-systemic-approach) (accessed 16 August 2022).

- 583 11. Openbaar Lichaam Bonaire. Directie Samenleving en Zorg. Afdeling Publieke Gezondheid.
584 Publieke Gezondheid op Bonaire 2020-2023. Een blik op de gezondheidssituatie van de
585 Bonairiaanse bevolking en de weg vooruit naar een gezondere toekomst. Bonaire: Openbaar
586 Lichaam Bonaire; 3 October 2022. 28.
- 587 12. Centers for Disease Control and Prevention. Risk for COVID-19 infection, hospitalization,
588 and death by Race/Ethnicity. 2022. Available from: [https://www.cdc.gov/coronavirus/2019-](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html)
589 [ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html](https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html)
590 (accessed 12 January 2023).
- 591 13. Kaiser Family Foundation. Covid-19 cases and deaths by Race/ethnicity: Current data and
592 changes over time. [place unknown], [publisher unknown]. 2023. Available from: from
593 [https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-cases-and-deaths-by-race-](https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-cases-and-deaths-by-race-ethnicity-current-data-and-changes-over-time/)
594 [ethnicity-current-data-and-changes-over-time/](https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-cases-and-deaths-by-race-ethnicity-current-data-and-changes-over-time/) (accessed 12 January 2023).
- 595 14. Pressley A, Beyer D. Congressional Letter on Long Covid. 25 January 2022. Congress of the
596 United States. House of Representatives. Available from:
597 [https://pressley.house.gov/sites/pressley.house.gov/files/Pressley%20Beyer%20Congressional](https://pressley.house.gov/sites/pressley.house.gov/files/Pressley%20Beyer%20Congressional%20Letter%20on%20Long%20COVID.pdf)
598 [%20Letter%20on%20Long%20COVID.pdf](https://pressley.house.gov/sites/pressley.house.gov/files/Pressley%20Beyer%20Congressional%20Letter%20on%20Long%20COVID.pdf) (accessed 13 January 2023).
- 599 15. Michelen M, Manoharan L, Elkheir N, et al. Characterising long COVID: a living systematic
600 review. *BMJ Glob Health* 2021;6:e005427; <https://doi.org/10.1136/bmjgh-2021-005427>
- 601 16. Regeling Aanspraken Zorgverzekering Bes wetten.nl - Regeling - Regeling aanspraken
602 zorgverzekering BES - BWBR0029442. [place unknown]. Available from:
603 <https://wetten.overheid.nl/BWBR0029442/2023-01-20> (accessed 16 April 2023).
- 604 17. Department of Health and Human Services. USA government. What is long COVID? Long
605 COVID terms and definitions development explained. 2023. Available from
606 <https://www.covid.gov/longcovid/definitions> (accessed 12 January 2023).
- 607 18. Digitale Checklisten. Grip op ieder proces. 2023. Available from: <https://tempobonaire.com/nl>
608 (accessed 16 January 2023).
- 609 19. Tempo Flexible Manpower. Let's shape the future of staffing. [place unknown], [publisher
610 unknown]. 2023. Available from: <https://digitalechecklisten.nl/> (accessed 16 January 2023).
- 611 20. The Central Committee on Research Involving Human Subjects. Home page. [place
612 unknown], [publisher unknown]. 2023. Available from: <https://english.ccmo.nl/> (accessed 12
613 January 2023).
- 614 21. Mutubuki EN, van der Maaden T, Leung KY, et al. Prevalence and determinants of persistent
615 symptoms after infection with SARS-CoV-2: protocol for an observational cohort study
616 (LongCOVID-study). *BMJ Open* 2022;12:e062439; [https://doi.org/10.1136/bmjopen-2022-](https://doi.org/10.1136/bmjopen-2022-062439)
617 [062439](https://doi.org/10.1136/bmjopen-2022-062439)

- 618 22. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze
619 and compare ROC curves. BMC Bioinformatics 2011;12:77; [https://doi.org/10.1186/1471-](https://doi.org/10.1186/1471-2105-12-77)
620 [2105-12-77](https://doi.org/10.1186/1471-2105-12-77)
- 621 23. Harrell FE. Regression modeling strategies: With Applications to Linear Models, Logistic and
622 Ordinal Regression, and Survival Analysis. New York, USA: Springer 2015; 11-105.
- 623 24. Brightling CE, Evans RA. Long COVID: which symptoms can be attributed to SARS-CoV-2
624 infection? Lancet 2022;400:411-413; [https://doi.org/10.1016/S0140-6736\(22\)01385-X](https://doi.org/10.1016/S0140-6736(22)01385-X)
- 625 25. World Health Organization. Post covid-19 condition (long covid). [place unknown], [publisher
626 unknown]. 2023. Available from [https://www.who.int/europe/news-room/fact-](https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition)
627 [sheets/item/post-covid-19-condition](https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition) (accessed 12 January 2023).
- 628 26. Ballering AV, van Zon SKR, Olde Hartman TC, et al. Lifelines Corona Research Initiative.
629 Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort
630 study. Lancet 2022;400:452-461; [https://doi.org/10.1016/S0140-6736\(22\)01214-4](https://doi.org/10.1016/S0140-6736(22)01214-4)
- 631 27. Goërtz YMJ, Van Herck M, Delbressine JM, et al. Persistent symptoms 3 months after a
632 SARS-CoV-2 infection: the post-COVID-19 syndrome? ERJ Open Res 2020;6:00542-2020;
633 <https://doi.org/10.1183/23120541.00542-2020>
- 634 28. Greenhalgh T, Sivan M, Delaney B, et al. Long covid-an update for primary care. BMJ
635 2022;378:e072117; <https://doi.org/10.1136/bmj-2022-072117>

636

637

638

639

640

641

642

643

644

645

646

647 **STATEMENTS & DECLARATIONS**

648

649 **Funding**

650 "This work was funded by the RIVM department CIb their RAC programme budget: Research budget
651 (grant number 0113/2021, August 5th 2021)".

652 **Competing Interests**

653 The authors have no conflict of interest to declare that are relevant to the content of this article.

654 **Author Contributions**

655 LJ secured funding for the performance of the study; LJ, DB, GM, IT and EG contributed to the study
656 conception and design; LJ, GM and IT established data collection procedures; IT was responsible for
657 data registry; DB and EG defined the statistical analysis plan; DB, TD, LJ, EG and GM contributed to
658 the synthesis of key findings and interpretation of the results; DB performed the primary literature
659 search, carried out descriptive and univariate analyses, and took co-lead in drafting and writing of the
660 manuscript prior to publication. TD carried out descriptive analyses, quantitative analyses, and took
661 co-lead in drafting the manuscript prior to publication; EG, LJ, IT and GM critically reviewed the
662 manuscript.

663 **Ethics approval**

664 The study protocol was sent to the Central Committee on Research Involving Human Subjects
665 (CCMO Netherlands), who confirmed on 8 October 2021 that the study did not require ethical
666 approval due to its observational design.

667 **Consent to participate**

668 Informed consent was obtained from all individual participants included in the study.

669 **Consent to publish**

670 Not applicable.

671

672 Supplementary Information for:

673

674 **The prevalence of post COVID-19 condition (PCC) and a simple risk scoring tool for PCC**
675 **screening on Bonaire, Caribbean Netherlands: a retrospective cohort study.**

676 Danytza SF Berry^{1*}, Thomas Dalhuisen^{1*}, Giramin Marchena², Ivo Tiemessen³, Eveline Geubbels¹
677 Loes Jaspers²

678 *Joint first authorship.

679 Corresponding author: Loes Jaspers, loesjaspers@gmail.com

680

681 **Affiliations:**

682 ¹ National Institute for Public Health and the Environment, Centre for Infectious Disease Control,
683 Bilthoven, The Netherlands

684 ² Public Health Department Bonaire, Public Entity of Bonaire, Caribbean Netherlands

685 ³ Mobilito, Bonaire, Caribbean Netherlands

686

687 Supplementary content:

688

689 **Supplementary Table 1 (S1):** Variables used for model building

690

691 **Supplementary Table 2 (S2):** Coefficients from sparse and rich multivariate logistic regression models
692 assessing the risk of PCC in full patient – and non-hospitalized patient samples

693

694 **Supplementary Table 3 (S3):** Risk score formulas to calculate the risk of PCC in COVID-19 patients

695

696 **Supplementary Table 4 (S4):** Coefficients from intermediate multivariate logistic regression models assessing
697 the risk of PCC in full patient (3a)– and non-hospitalized patient (3b) samples

698

699 **Supplementary Table 5 (S5):** Sensitivity analyses for risk score formulas to calculate the risk of PCC in
700 COVID-19 patients

701

702 **Supplementary Table 6 (S6):** Prevalence of symptoms (n, %) by subgroup in the pre-pandemic, acute, and
703 post-acute phase

704

705 **Supplementary Table 7 (S7):** The prevalence of post-acute symptoms across different baseline characteristics
706 of PCC patients (n=160)

707

708 **Supplementary Table 8 (S8):** Characteristics of SARS-CoV-2 positive cases (n= 390) included in risk scoring
709 tool analyses

710 **Supplementary Figure 1 (SF1):** Epiurve and inclusion period for SARS-CoV-2 positive cases on Bonaire,
711 Caribbean Netherlands

712

713 **Supplementary Figure 2 (SF2):** Internal calibration plots for predicting the risk of PCC
714 PCC in COVID-19 patients

715

716 **Supplementary Figure 3 (SF3):** Receiver Operating Characteristic (ROC) curves for PCC risk scoring tools

717

718

719

720

721 **Supplementary Table 1 (S1):** Variables used for model building

Variable	Type	Description
Age	Categorical	Age in years. Categories are: 0-19 years, 20-39 years, 40-59 years, and 60+ years old
Sex	Binary	Male, female.
Education level	Categorical	Categories are: Low level education = no education, primary level education, lower vocational & lower secondary education, moderate level education = higher secondary education, high level education = university level & post-secondary education.
BMI	Categorical	BMI (kg/m ²), categorized into underweight (<18), normal (≥18-<25), overweight (≥25-<30), obese (≥30-<35) and morbidly obese (≥35).
Smoking	Binary	Binary variable indicating whether or not a patient smoked cigarettes before the start of the pandemic.
Alcohol use	Binary	Binary variable indicating whether or not a patient used alcohol before the start of the pandemic.
Presence of one or more comorbidities	Binary	Binary variable indicating whether a person suffers from any of the following prior to the start of the pandemic: self-reported post-acute respiratory problems (including asthma, chronic obstructive pulmonary disorder, tuberculosis), post-acute cardiovascular disease, diabetes, severe kidney disease (including dialysis or kidney transplantation), HIV, severe liver disease, being immunocompromised, cancer, depression or anxiety, stroke, rheumatism, arthritis, or memory loss due to a neurological condition or dementia.
Number of symptoms in the acute phase	Numerical	Number of symptoms a patient experienced during the acute phase of their SARS-CoV-2 infection. Symptoms include: chest pain, concentration problems, cough, fatigue, headache, heart palpitations, loss of appetite, reduced muscle strength, loss of sense of smell, loss of sense of taste, muscle ache, reduced physical endurance, shortness of breath, sleeping problems.
GP visit in the acute phase	Binary	Binary variable indicating whether a patient visited or was visited by a GP during the acute phase of their SARS-CoV-2 infection.
Hospitalized during the acute phase	Binary	Binary variable indicating whether a patient was hospitalized during the acute phase of their SARS-CoV-2 infection.

722

723 **Supplementary Table 2 (S2):**

724 Coefficients from sparse and rich multivariate logistic regression models assessing the risk of PCC in full patient – and non-hospitalized patient
725 samples.

Model	Age (ref cat: 0-19 years old)			Sex (ref: male)	Presence of one or more comorbidities (ref: no)	GP visit acute phase (ref: no)	Hospitalized during acute phase (ref: no)	Alcohol use (ref: no)	Smoking (ref: no)	BMI (ref: normal)				Education (ref: low)		726 Number of symptoms acute phase
	20-39	40-60	60+	Female	Yes	Yes	Yes	Yes	Yes	Under weight	Over weight	Obese	Morbidly obese	Moderate	High	
Sparse 1a	1.17*	1.27**	1.02	0.80***	0.47	0.66***	1.91***	-	-	-	-	-	-	-	-	-
Sparse 1b	1.12*	1.16*	0.75	0.72***	0.43	0.90	-	-	-	-	-	-	-	-	-	-
Rich 2a	0.30	0.59	0.50	0.46	0.33	0.31	1.17*	0.39	-0.18	0.49	0.33	0.35	0.90**	0.28	0.65	0.19***
Rich 2b	0.23	0.46	0.14	0.37	0.23	0.59	-	0.37	-0.06	0.44	0.23	0.37	0.90**	0.31	0.63	0.19***

727 **Note:** Reference categories are shown in each column. Significant coefficients are shown in bold; the Asterix refers to three categories: *p<0.05,
728 **p<0.01, ***p<0.001. Models 1a and 2a use the full patient sample (n=390) and Models 1b and 2b use the non-hospitalized patient sample
729 (n=362).

730 **Supplementary Table 3 (S3):**

731 Risk score formulas to calculate the risk of PCC in COVID-19 patients.

732 **1a.** Sparse risk score formula based on full patient sample:

$$\begin{aligned}
 733 \quad & \textbf{Risk Score} = \textbf{Age Score} + \mathbf{8}(\textbf{If sex} = \textbf{female}) + \mathbf{5}(\textbf{If comorbidities} \\
 734 \quad & \quad = \textbf{yes}) + \mathbf{7}(\textbf{If GP visit in acute phase} = \textbf{yes}) + \mathbf{19}(\textbf{If hospitalized} \\
 735 \quad & \quad = \textbf{yes})
 \end{aligned}$$

736 **Score table:**

Age Category	Age Score
0-19	0
20 – 39	12
40 – 60	13
60 +	10

737

738 **1b.** Sparse risk score formula based on non-hospitalized sample :

$$\begin{aligned}
 739 \quad & \textbf{Risk score} = \textbf{Age Score} + \mathbf{7}(\textbf{If sex} = \textbf{female}) + \mathbf{4}(\textbf{If comorbidities} \\
 740 \quad & \quad = \textbf{yes}) + \mathbf{9}(\textbf{If GP visit in acute phase} = \textbf{yes})
 \end{aligned}$$

741 **Score table:**

Age Category	Age Score
0-19	0
20 – 39	11
40 – 60	12
60 +	8

742

743 **2a.** Rich risk score formula based on the full patient sample:

$$\begin{aligned}
 744 \quad & \textbf{Risk Score} = \textbf{Age Score} + \textbf{BMI Score} + \textbf{Education Score} + \mathbf{2} \\
 745 \quad & \quad * (\textbf{Number of symptoms in the acute phase}) + \mathbf{4}(\textbf{If sex} \\
 746 \quad & \quad = \textbf{female}) + \mathbf{2}(\textbf{If comorbidities} \\
 747 \quad & \quad = \textbf{yes}) + \mathbf{3}(\textbf{If GP visit in acute phase} = \textbf{yes}) + \mathbf{12}(\textbf{If hospitalized} \\
 748 \quad & \quad = \textbf{yes}) + \mathbf{4}(\textbf{If alcohol consumption} = \textbf{yes}) - \mathbf{2}(\textbf{If smoker} = \textbf{yes})
 \end{aligned}$$

749 **Score table:**

Age Category	Age Score
0-19	0
20 – 39	3
40 – 60	6
60 +	5

BMI Category	BMI Score
Normal	0
Underweight	5
Overweight	3
Obese	4
Morbidly Obese	9
Education Category	Education Score
Low	0
Moderate	3
High	7

750

751 **2b.** Rich risk score formula based on non-hospitalized sample:

$$\begin{aligned}
 752 \quad \text{Risk score} &= \text{Age Score} + \text{BMI Score} + \text{Education Score} + 2 \\
 753 \quad & * (\text{Number of symptoms in the acute phase}) + 4(\text{If sex} \\
 754 \quad & = \text{female}) + 2(\text{If comorbidities} \\
 755 \quad & = \text{yes}) + 6(\text{If GP visit in acute phase} \\
 756 \quad & = \text{yes}) + 4(\text{If alcohol consumption} = \text{yes}) - 1(\text{If smoker} = \text{yes})
 \end{aligned}$$

757 **Score table:**

Age Category	Age Score
0-19	0
20 – 39	3
40 – 60	6
60 +	5
BMI Category	BMI Score
Normal	0
Underweight	5
Overweight	3
Obese	4
Morbidly Obese	9
Education Category	Education Score
Low	0
Moderate	3
High	7

758

759 **Supplementary Table 4 (S4):**

760 Coefficients from intermediate multivariate logistic regression models assessing the risk of
 761 PCC in full patient (3a)– and non-hospitalized patient (3b) samples.
 762

Model	Age (ref :0-19)			Sex (ref: Male)	Comorbidities (ref;no)	Hospitalized (ref: no)	GP visit acute phase (ref: no)	Alcohol use (ref: no)	Smoking (ref: no)	BMI (ref: normal)			
	20-39	40-60	60+	Female	Yes	Yes	Yes	Yes	Yes	Under weight	Over weight	Obese	Morbidly obese
3a	0.86	1.00*	0.84	0.74**	0.36	1.92***	0.77	0.31	-0.16	0.35	0.17	0.30	0.80*
3b	0.90	1.06*	1.04	0.61*	0.49	-	0.81	0.22	-0.08	0.29	0.27	0.45	0.89**

763 **Note:** Reference categories are shown in each column. Significant coefficients are shown in
 764 bold; the Asterix refers to three categories: *p<0.05, **p<0.01, ***p<0.001.

Supplementary Table 5 (S5):

Sensitivity analyses for risk score formulas to calculate the risk of PCC in COVID-19 patients.

3a. Intermediate risk score formula based on full patient sample:

AUC: 0.708 (95CI%, 0.655 – 0.7602)

$$\begin{aligned}
 \text{Risk Score} &= \text{Age Score} + \text{BMI Score} + 8(\text{If sex} = \text{female}) \\
 &+ 3(\text{If comorbidities} = \text{yes}) + 8(\text{If GP visit in acute phase} = \text{yes}) \\
 &+ 19(\text{If hospitalized} = \text{yes}) + 3(\text{If alcohol consumptions} \\
 &= \text{yes}) - 2(\text{If smoker} = \text{yes})
 \end{aligned}$$

Age Category	Age Score
0-19	0
20 – 39	9
40 – 60	10
60 +	8
BMI Category	BMI Score
Normal	0
Underweight	4
Overweight	2
Obese	3
Morbidly Obese	8

3b. Intermediate risk score formula based on non-hospitalized patient sample:

AUC: 0.6742 (95%CI 0.6149 – 0.7291)

$$\begin{aligned}
 \text{Risk Score} &= \text{Age Score} + \text{BMI Score} + 6(\text{If sex} = \text{female}) + 5(\text{If comorbidities} \\
 &= \text{yes}) + 8(\text{If GP visit in acute phase} \\
 &= \text{yes}) + 2(\text{If alcohol consumption} = \text{yes}) - 1(\text{If smoker} = \text{yes})
 \end{aligned}$$

Age Category	Age Score
0-19	0
20 – 39	9
40 – 60	11
60 +	10
BMI Category	BMI Score
Normal	0
Underweight	3
Overweight	3
Obese	4
Morbidly Obese	9

Supplementary Table 6 (S6): Prevalence of symptoms (n, %) by subgroup in the pre-pandemic, acute, and post-acute phase.

	Pre-pandemic phase, n (%)				Acute phase, n (%)		Post-acute phase, n (%)
	PCC patients (n=160)	Non-PCC cases (n=254)	COVID-19 positive cohort (n=414)	Community controls (n=75)	PCC patients (n=160)	Non-PCC cases (n=254)	PCC patients (n=160)
Chest pain	12 (7.5)	12 (4.7)	24 (5.8)	3 (4.0)	81 (50.6)	59 (23.2)	52 (32.5)
Concentration problems ^B	10 (6.3)	13 (5.1)	23 (5.6)	2 (2.7)	80 (50)	46 (18.1)	65 (40.6)
Cough ^B	13 (8.1)	31 (12.2)	44 (10.6)	5 (6.7)	120 (75.0)	129 (50.8)	78 (48.8)
Fatigue ^B	15 (9.4)	20 (7.9)	35 (8.5)	7 (9.3)	136 (85.0)	121 (47.6)	107 (66.9)
Headache ^B	33 (20.6)	52 (20.5)	85 (20.5)	9 (12.0)	124 (77.5)	154 (60.6)	79 (49.4)
Heart palpitations ^B	13 (8.1)	9 (3.5)	22 (5.3)	2 (2.7)	56 (35.0)	35 (13.8)	37 (23.1)
Loss of appetite ^B	6 (3.8)	11 (4.3)	17 (4.1)	3 (4.0)	113 (70.6)	113 (44.5)	96 (59.4)
Reduced muscle strength ^B	8 (5.0)	4 (1.6)	12 (2.9)	5 (6.7)	122 (76.2)	98 (38.6)	81 (50.6)
Loss of sense of smell ^B	5 (3.1)	9 (3.5)	14 (3.4)	0 (0)	97 (60.6)	118 (46.5)	62 (38.8)
Loss of sense of taste ^B	3 (1.9)	9 (3.5)	12 (2.9)	0 (0)	105 (65.6)	126 (49.6)	61 (38.1)
Muscle ache ^B	17 (10.6)	14 (5.5)	31 (7.5)	8 (10.7)	114 (71.2)	102 (40.2)	72 (45.0)
Reduced physical endurance ^{A, B}	12 (7.5)	6 (2.4)	18 (4.3)	4 (5.3)	133 (83.1)	117 (46.1)	105 (65.6)
Shortness of breath ^{A, B}	20 (12.5)	15 (5.9)	35 (8.5)	3 (4.0)	107 (66.9)	86 (33.9)	95 (59.4)
Sleeping problems ^B	20 (12.5)	20 (7.9)	40 (9.7)	6 (8.0)	98 (61.3)	75 (29.5)	66 (41.2)

^A Statistically significant difference between PCC patients and non-PCC cases ($p < 0.05$) in the pre-pandemic phase, with Fisher's exact test.

^B Statistically significant difference between PCC patients and non-PCC cases ($p < 0.05$) in the acute phase, with Fisher's exact test.

^C Statistically significant difference between PCC patients and community controls ($p < 0.05$) in the pre-pandemic phase, with Fisher's exact test (does not occur).

^D Statistically significant difference between COVID-19 positive cases and community controls ($p < 0.05$) in the pre-pandemic phase, with Fisher's exact test (does not occur).

Supplementary Table 7 (S7): The prevalence of post-acute symptoms across different baseline characteristics of PCC patients (n=160).

	Chest pain (n=52)	Concentration problems (n=65)	Cough (n=78)	Fatigue (n=107)	Headache (n=79)	Heart palpitations (n=37)	Loss of appetite (n=95)	Reduced muscle strength (n=81)	Loss of sense of smell (n=62)	Loss of sense of taste (n=61)	Muscle ache (n=72)	Reduced physical endurance (n=105)	Shortness of breath (n=95)	Sleeping problems (n=66)
Age group (years), n (%)														
0 – 19 (n=6)	1 (17)	1 (17)	5 (83)	4 (67)	5 (83)	0 (0)	2 (33)	2 (33)	2 (33)	0 (0)	1 (17)	2 (33)	1 (17)	0 (0)
20 – 39 (n=61)	23 (38)	22 (36)	27 (44)	40 (66)	30 (49)	13 (21)	35 (57)	29 (48)	21 (34)	24 (39)	25 (41)	39 (64)	40 (66)	21 (34)
40-59 (n=64)	12 (19)	28 (44)	26 (41)	40 (63)	28 (44)	11 (17)	36 (56)	31 (48)	23 (36)	22 (34)	25 (39)	41 (64)	34 (53)	27 (42)
60+ (n=29)	16 (55)	14 (48)	20 (69)	22 (76)	16 (55)	13 (45)	22 (76)	19 (66)	16 (55)	15 (52)	21 (72)	23 (79)	20 (69)	18 (62)
Sex, n (%)														
Female (n= 114)	42 (37)	50 (44)	57 (50)	79 (69)	60 (53)	30 (26)	70 (61)	57 (50)	44 (39)	46 (40)	52 (46)	74 (65)	71 (62)	53 (46)
Male (n=46)	10 (22)	15 (33)	21 (46)	28 (61)	19 (41)	7 (15)	25 (54)	24 (52)	18 (39)	15 (33)	20 (43)	31 (67)	24 (52)	13 (28)
BMI category (kg/m²), n (%)														
Underweight (n=2)	0 (0)	0 (0)	2 (100)	1 (50)	1 (50)	0 (0)	0 (0)	0 (0)	1 (50)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Normal (n=32)	6 (19)	14 (44)	12 (38)	20 (63)	16 (50)	4 (13)	13 (41)	18 (56)	14 (44)	13 (41)	12 (38)	22 (69)	14 (44)	11 (34)

Overweight (n=41)	14 (34)	14 (34)	23 (56)	25 (61)	16 (39)	7 (17)	23 (56)	18 (44)	16 (39)	14 (34)	17 (41)	25 (61)	23 (56)	16 (39)
Obese (n=36)	13 (36)	22 (61)	15 (42)	22 (61)	18 (50)	8 (22)	25 (69)	15 (42)	10 (28)	13 (36)	13 (36)	20 (56)	21 (58)	11 (31)
Morbidly obese (n=49)	19 (39)	26 (53)	26 (53)	39 (80)	28 (57)	18 (37)	34 (69)	30 (61)	21 (43)	21 (43)	30 (61)	38 (78)	38 (78)	21 (43)
Comorbidities, n (%)														
Yes, one or more (n=33)	13 (39)	14 (42)	23 (70)	24 (73)	11 (33)	9 (27)	19 (58)	20 (61)	16 (48)	14 (42)	18 (55)	26 (79)	25 (76)	16 (48)
No (n=127)	39 (31)	51 (40)	55 (43)	83 (65)	68 (54)	28 (22)	76 (60)	61 (48)	46 (36)	48 (38)	54 (43)	79 (62)	70 (55)	50 (39)
Smoking, n (%)														
Yes (n=18)	6 (33)	6 (33)	12 (67)	10 (56)	9 (50)	4 (22)	14 (78)	7 (39)	7 (39)	8 (44)	5 (28)	11 (61)	10 (56)	6 (33)
No (n=142)	46 (32)	59 (42)	66 (46)	97 (68)	70 (49)	33 (23)	81 (57)	74 (52)	55 (39)	53 (37)	67 (47)	94 (66)	85 (60)	60 (42)
Alcohol use, n (%)														
Yes (n=109)	33 (30)	44 (40)	52 (48)	69 (63)	50 (46)	20 (18)	65 (60)	52 (48)	42 (39)	41 (38)	44 (40)	71 (65)	63 (58)	42 (39)
No (n=51)	19 (37)	21 (41)	26 (51)	38 (75)	29 (57)	17 (33)	30 (59)	29 (57)	20 (39)	20 (39)	28 (55)	34 (67)	32 (63)	24 (47)
Education level, n (%)														
Low (n=119)	40 (34)	47 (39)	61 (51)	82 (69)	62 (52)	31 (26)	75 (63)	61 (51)	48 (40)	47 (39)	60 (50)	81 (68)	69 (58)	52 (44)
Moderate (n=23)	7 (30)	11 (48)	10 (43)	16 (70)	10 (43)	3 (13)	11 (48)	13 (57)	7 (30)	8 (35)	9 (39)	16 (70)	15 (65)	9 (39)
High (n=17)	5 (29)	7 (41)	7 (41)	9 (53)	7 (41)	3 (18)	8 (47)	7 (41)	7 (41)	6 (35)	3 (18)	7 (41)	11 (65)	5 (29)
Hospitalized during acute phase, n (%)														
Yes (n=24)	6 (25)	15 (63)	12 (50)	16 (67)	9 (38)	6 (25)	17 (71)	17 (71)	7 (29)	7 (29)	13 (54)	21 (88)	19 (79)	10 (42)
No (n=136)	46 (34)	50 (37)	66 (49)	91 (66)	70 (51)	31 (23)	78 (57)	64 (47)	55 (40)	54 (39)	59 (43)	84 (61)	76 (56)	56 (41)
GP visit during acute phase, n (%)														

Yes (n=14)	6 (43)	7 (50)	10 (71)	11 (79)	6 (43)	5 (36)	9 (64)	7 (50)	8 (57)	9 (64)	5 (36)	11 (79)	10 (71)	11 (79)
No (n=146)	46 (32)	58 (40)	68 (47)	96 (66)	73 (50)	32 (22)	86 (59)	74 (51)	54 (37)	52 (36)	67 (46)	94 (64)	85 (58)	55 (38)

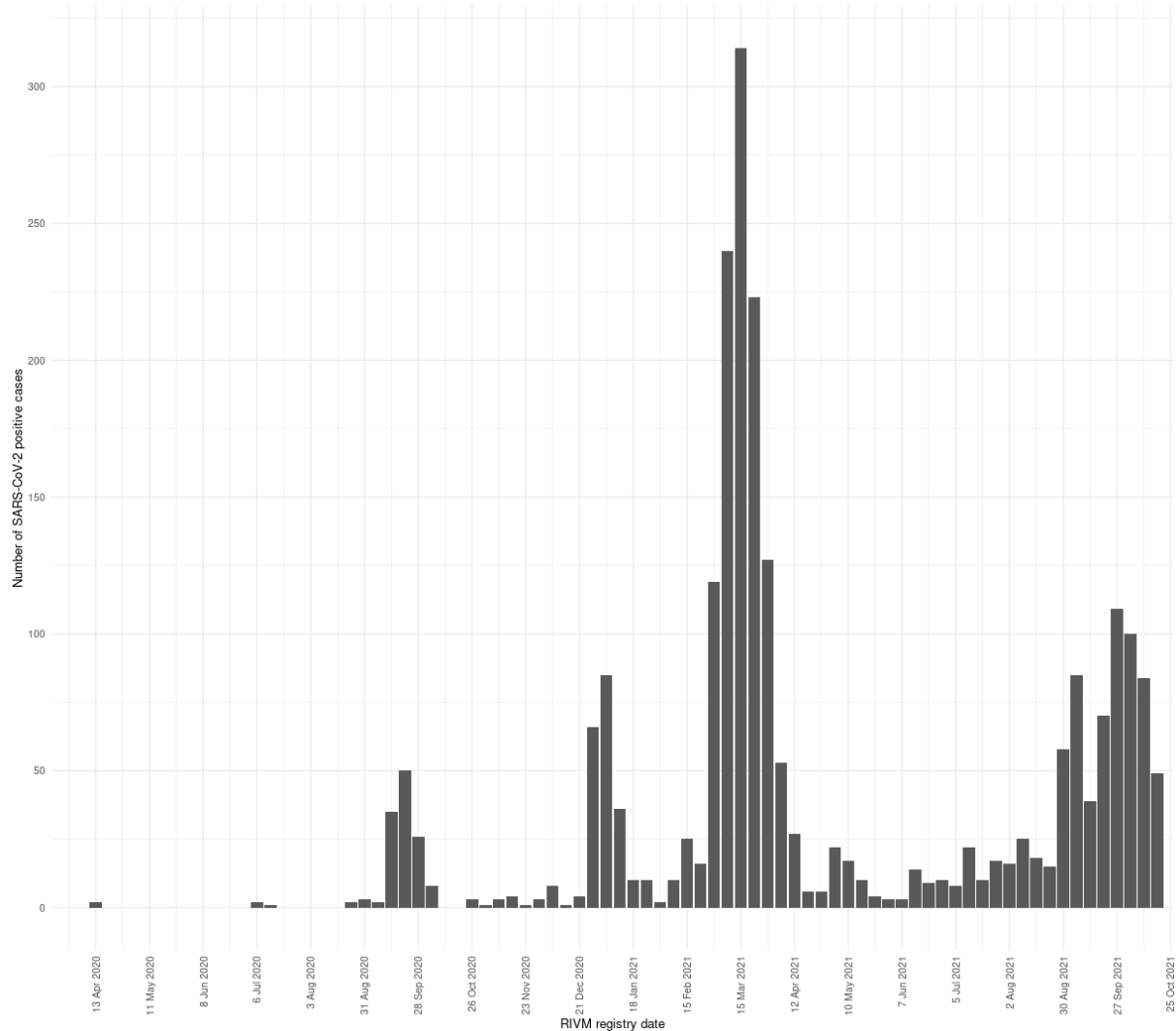
Supplementary Table 8 (S8): Characteristics of SARS-CoV-2 positive cases (n= 390) included in risk scoring tool analyses.

	PCC patients (n=156)	Non-PCC cases (n=234)	P-value
Age group (years), n (%)			
0-19 years	6 (34%)	31 (13%)	0.01^B
20 – 39 years	61 (39%)	90 (39%)	
40 – 59 years	63 (40%)	80 (34%)	
60 + years	26 (17%)	33 (14%)	
Sex, n (%)			
Male	45 (29%)	107 (46%)	<0.01^B
Female	111 (71%)	127 (54%)	
Comorbidities, n (%)			
Yes	30 (19%)	27 (12%)	0.04^B
No	126 (81%)	207 (89%)	
GP visit in acute phase, n (%)			
Yes	14 (9%)	8 (3%)	0.02^B
No	142 (91%)	226 (97%)	
Hospitalized for COVID-19 in acute phase, n (%)			
Yes	22 (14%)	6 (23%)	<0.01^B
No	134 (86%)	228 (97%)	
Median number of symptoms in acute phase (IQR)			
	10 (5)	5 (7)	<0.01^A
Pre-pandemic alcohol use			
Yes	106 (68%)	148 (63%)	0.81^B
No	50 (32%)	86 (37%)	
Pre-pandemic smoking			
Yes	18 (12%)	32 (14%)	0.63^B
No	138 (89%)	202 (86%)	
BMI (kg/m²) category			
Underweight	2 (1%)	4 (2%)	<0.01^B
Normal	32 (21%)	75 (32%)	
Overweight	40 (26%)	70 (30%)	
Obese	35 (22%)	48 (21%)	
Morbidly obese	47 (30%)	37 (16%)	
Education level			
Low	116 (74%)	194 (83%)	0.08^B
Moderate	24 (15%)	27 (12%)	
High	17 (11%)	13 (6%)	

^A Wilcoxon test. ^B Fisher's exact test.

SUPPLEMENTARY FIGURES

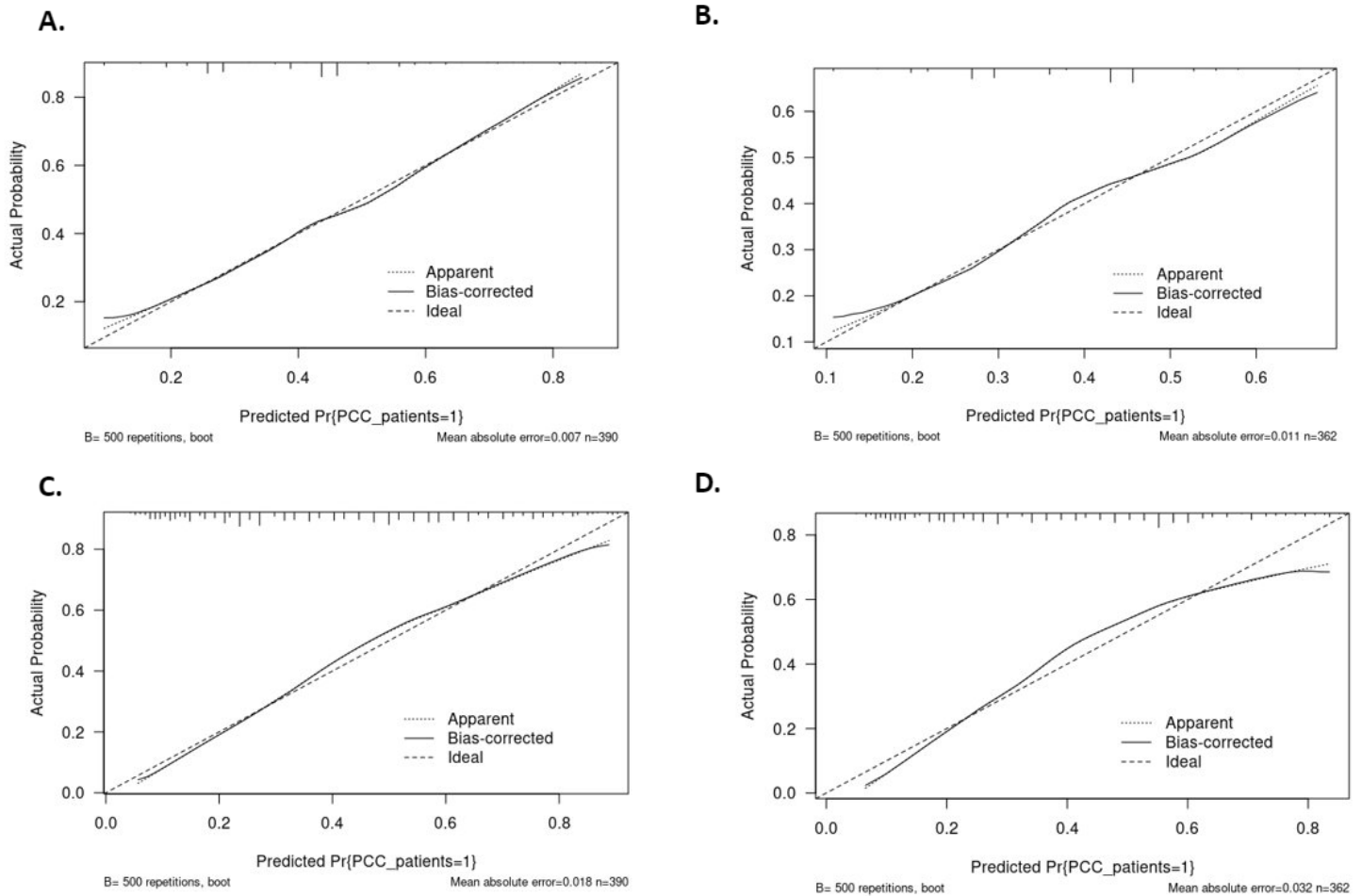
Supplementary Figure 1 (SF1): Epicurve and inclusion period for SARS-CoV-2 positive cases on Bonaire, Caribbean Netherlands.



Epidemiological curve of the SARS-CoV-2 outbreak on Bonaire, Caribbean Netherlands, displaying the weekly number of confirmed cases by 4-week intervals from the start of the pandemic until 27 October 2021. In this study, we included permanent residents of Bonaire, who had a SARS-CoV-2 positive test result between the start of the pandemic in 2020 until 1 October 2021.

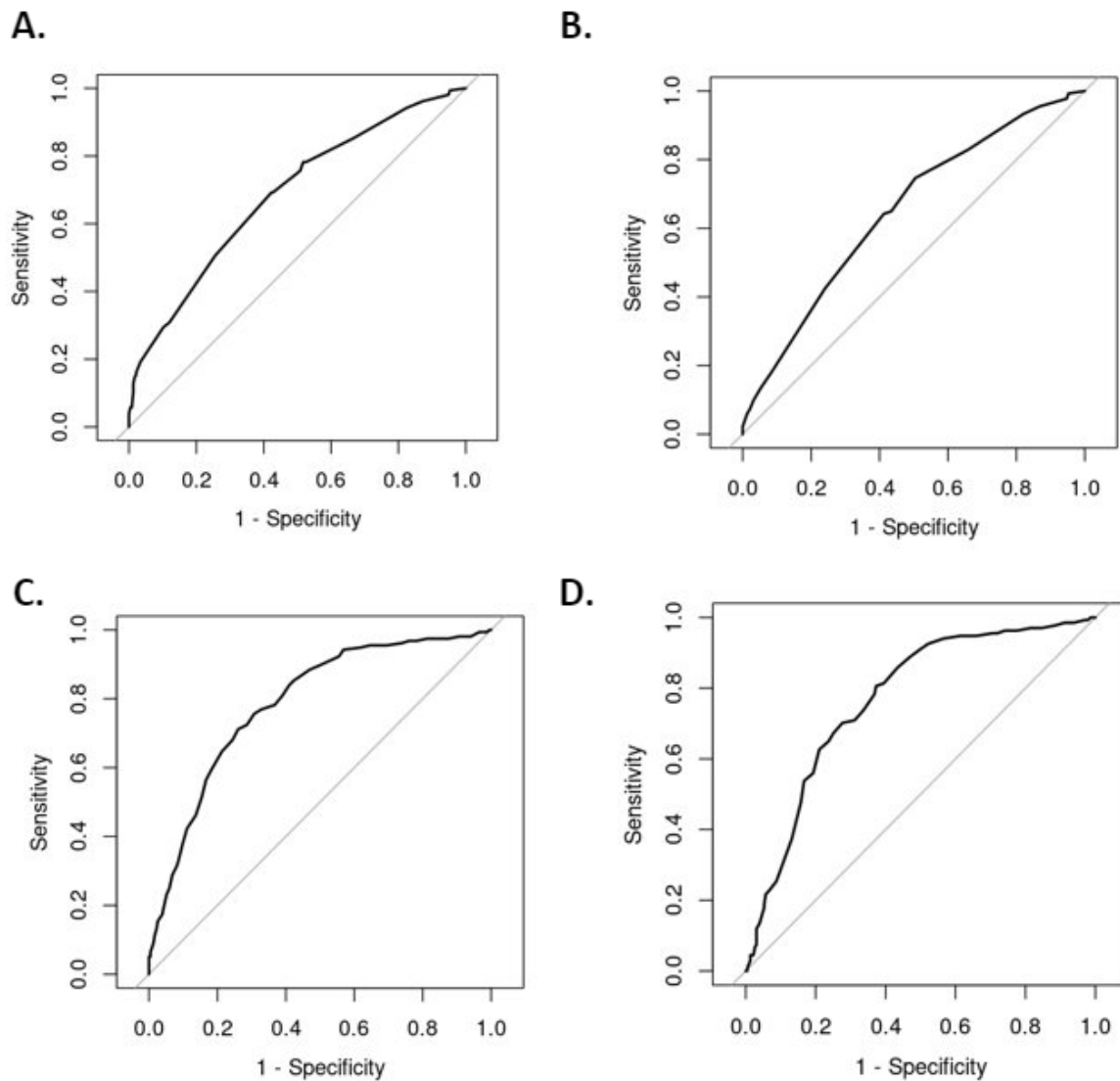
Source: Data from RIVM (downloaded on 20 January 2023)

Supplementary Figure 2 (SF2): Internal calibration plots for predicting the risk of PCC in COVID-19 patients.



Internal calibration plots for predicting the risk of PCC in COVID-19 patients. Dashed lines show ideal calibration, dotted line shows apparent calibration and the solid line shows bias-corrected calibration. (A). Sparse risk scoring tool based on full patient sample. (B) Sparse risk scoring tool based on non-hospitalized sample. (C) Rich risk scoring tool based on full patient sample. (D) Rich risk scoring tool based on non-hospitalized sample

Supplementary Figure 3 (SF3): Receiver Operating Characteristic (ROC) curves for PCC risk scoring tools.



Receiver Operating Characteristic (ROC) curves for PCC risk scoring tools (A). Sparse risk scoring tool based on full patient sample. (B) Sparse risk scoring tool based on non-hospitalized sample. (C) Rich risk scoring tool based on full patient sample. (D) Rich risk scoring tool based on non-hospitalized sample.