

1 **Article type:** Original Article

2 **Development and external validation of a diagnostic multivariable prediction model**
3 **for a prompt identification of cases at high risk for SARS-COV-2 infection among**
4 **patients admitted to the emergency department.**

5 Nicola UGHI^{1*}, Antonella ADINOLFI¹, Michel CHEVALLARD¹, Laura BELLOLI¹, Michele
6 SENATORE², Alessandro TOSCANO¹, Andrea BELLONE³, Cristina GIANNATTASIO^{4,5}, Paolo
7 TARSIA⁶, Massimo PUOTI⁷, Francesco SCAGLIONE^{2,8}, Fabrizio COLOMBO⁹, Michaela
8 BERTUZZI¹⁰, Giuseppe BETTONI³, Davide FERRAZZI³, Alessandro MALOBERTI^{4,5}, Armanda
9 DICUONZO¹¹, Francesca DEL GAUDIO¹¹, Claudio ROSSETTI¹², Oscar Massimiliano EPIS¹; On
10 behalf of the Niguarda COVID group.

11

12 ¹Division of Rheumatology, Multispecialist Medical Department, ASST Grande Ospedale
13 Metropolitano Niguarda, Milan, Italy

14 ²Division of chemical-Clinical and Microbiological Analyses, Department of Laboratory Medicine,
15 ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy

16 ³Division of Emergency Medicine and Emergency Room, Department of Emergencies and
17 Admissions (DEA), ASST Grande Ospedale Metropolitano Niguarda, Milan, Italy

18 ⁴Division of Cardiology 4, Cardio-Thoraco-Vascular Department, ASST Grande Ospedale
19 Metropolitano Niguarda, Milan, Italy

20 ⁵School of Medicine and Surgery, Milano-Bicocca University, Milan, Italy

21 ⁶Division of Pneumology, Multispecialist Medical Department, ASST Grande Ospedale
22 Metropolitano Niguarda, Milan, Italy

23 ⁷Division of Infectious Disease, Multispecialist Medical Department, ASST Grande Ospedale
24 Metropolitano Niguarda, Milan, Italy

25 ⁸Department of Oncology and Hemato-Oncology, Università degli Studi di Milano, Milan, Italy

26 ⁹Division of Internal Medicine 1, Multispecialist Medical Department, ASST Grande Ospedale
27 Metropolitano Niguarda, Milan, Italy

28 ¹⁰Division of Quality and Clinical Risk, Continuous Quality Improvement, ASST Grande Ospedale
29 Metropolitano Niguarda, Milan, Italy

30 ¹¹Functional Department for Higher Education, Research, and Development, ASST Grande
31 Ospedale Metropolitano Niguarda, Milan, Italy

32 ¹²Functional Department for Higher Education, Research, and Development, Interhospital
33 Functional Department of Nuclear Medicine, ASST Grande Ospedale Metropolitano Niguarda,
34 Milan, Italy

35

36

37

38 ***Corresponding Author:**

39 Nicola Ughi, Division of Rheumatology, ASST Grande Ospedale Metropolitano Niguarda, Milan,
40 Italy, p.zza Ospedale Maggiore 3, 20162, Milano (MI), Italy, Tel.: (+39) 02.6444.2168, Fax (+39)
41 02.6444.3277, email: nicola.ughi@ospedaleniguarda.it.

42

43 **Target Journal:** Internal and Emergency Medicine

44 ***Authors' contributions***

45 Oscar Epis, Nicola Ughi, Antonella Adinolfi, Michel Chevallard, Michele Senatore and Alessandro
46 Toscano contributed to the study conception, study design, material preparation and data
47 collection. The analyses were performed by Nicola Ughi. Data interpretation: all authors.

48 The first draft of the manuscript was written by Nicola Ughi and all authors commented on previous
49 versions of the manuscript. All authors read and approved the final manuscript.

50 **Abstract**

51 **Background:** An urgent need exists for an early detection of cases with a high-risk of SARS-CoV-
52 2 infection, particularly in high-flow and -risk settings, such as emergency departments (EDs). The
53 aim of this work is to develop and validate a predictive model for the evaluation of SARS-CoV-2
54 infection risk, with the rationale of using this tool to manage ED patients.

55 **Methods:** A retrospective study was performed by cross-sectionally reviewing the electronic
56 case records of patients admitted to Niguarda Hospital or referred to its ED in the period 15 March
57 to 24 April 2020.

58 Derivation sample was composed of non-random inpatients hospitalized on 24 April and admitted
59 before 22 April 2020. Validation sample was composed of consecutive patients who visited the ED
60 between 15 and 25 March 2020. The association between the dichotomic outcome and each
61 predictor was explored by univariate analysis with logistic regression models.

62 **Results:** A total of 113 patients in the derivation sample and 419 in the validation sample were
63 analyzed. History of fever, elder age and low oxygen saturation showed to be significant predictors
64 of SARS-CoV-2 infection. The neutrophil count improves the discriminative ability of the model,
65 even if its calibration and usefulness in terms of diagnosis is unclear.

66 **Conclusion:** The discriminatory ability of the identified models makes the overall performance
67 suboptimal; their implementation to calculate the individual risk of infection should not be used
68 without additional investigations. However, they could be useful to evaluate the spatial allocation of
69 patients while awaiting the result of the nasopharyngeal swab.

70 **Keywords:**

71 Diagnostic prediction model; COVID-19; SARS-CoV-2 infection; emergency department.

72 **Declarations**

73 **Funding**

74 There was no explicit funding for the development of this work.

75 ***Conflicts of interest/Competing interests***

76 None of the Authors declared conflict of interests.

77 ***Availability of data and material (data transparency)***

78 Data may be made available upon reasonable request.

79 ***Ethics approval***

80 The study was conducted within the protocol approved the ethics committee Milano Area 3
81 (register number 249-13052020).

82 ***Consent to participate***

83 All the participants signed an informed consent form

84 ***Consent for publication***

85 Not required as this manuscript doesn't include details, images or videos related to the participants

86 ***Acknowledgments***

87 Editorial assistance was provided by Simonetta Papa, PhD, and Aashni Shah (Polistudium SRL,
88 Milan, Italy). Graphical assistance was provided by Massimiliano Pianta (Polistudium SRL, Milan,
89 Italy).

90

91 **Key Messages box:**

92 What is already known on this topic

93 1 year after the onset of the coronavirus disease 2019 (COVID-19) pandemic, the trend of its
94 spread has not shown a substantial global reduction. An urgent need exists for efficient early
95 detection of cases with a high risk of SARS-CoV-2 infection and a number of diagnostic prediction

96 models have been developed, but a few models were externally validated in high-flow and -risk
97 settings, such as emergency departments (EDs).

98 What this study adds

99 This study develops and validate predictive models for the evaluation of SARS-CoV-2 infection
100 risk, with the rationale of using these tools to promptly manage patients who are afferent to the ED,
101 allocating them accordingly to the risk of infection while awaiting swab result. History of fever, older
102 age and low oxygen saturation showed to be significant predictors of the presence of SARS-CoV-2
103 infection. The use of laboratory findings, such as neutrophil count, showed to improve the
104 discriminative ability of the model, even if its calibration and usefulness in terms of diagnosis is
105 unclear.

106 INTRODUCTION

107 1 year after the onset of the coronavirus disease 2019 (COVID-19) pandemic, the trend of its
108 spread has not shown a substantial global reduction, with 238,031 new cases registered on 11
109 March 2021, 117,573,007 total confirmed cases and 2,610,925 total deaths [1] and recurring new
110 peaks of contagion are still observed locally in various geographical areas, including Italy (22,409
111 new cases registered on 11 March 2021, 3,123,368 total confirmed cases, 100,811 total deaths
112 [2]).

113 Pending the effects of the vaccine, test and tracing methods still represent the most effective tools
114 for containing the SARS-CoV-2 infection and preventing the spread of COVID-19 [3].

115 Rapid identification of SARS-CoV-2 infection is important for a prompt treatment and
116 implementation of isolation procedures, particularly in high-flow and high-risk settings, such as
117 emergency departments (EDs).

118 The standard reference test is represented by the search for the SARS-CoV-2 genome by reverse
119 transcriptase-polymerase chain reaction (RT-PCR) of nasopharyngeal swab material, but although
120 the technologies and execution times have been made more efficient, RT-PCR testing is time-
121 consuming and its execution time (up to 4 hours) is often not compatible with urgent/emergency
122 situations, consequently shorter decision times for the allocation of patients are still needed [4].

123 In order to reduce the identification time of SARS-CoV-2 cases, numerous models have been
124 developed for the diagnosis of infection and disease, but with numerous concerns in terms of risk
125 of bias. Consequently, the search for an improved predictive diagnostic tool is still open [5].

126 The aim of this work is to develop and validate a predictive model for the evaluation of SARS-CoV-
127 2 infection risk, with the rationale of using this tool to promptly manage patients who are afferent to
128 the ED, allocating them accordingly to the risk of infection while awaiting swab result.

129

130 PATIENTS AND METHODS

131 Study design and participants

132 A retrospective study was performed by cross-sectionally reviewing the electronic case records
133 of patients who were admitted to Niguarda Hospital or referred to its ED after the first Italian
134 COVID-19 case, between 15 March and 24 April 2020.

135 A non-random sample of inpatients from low-intensity general and specialist medical units who
136 were hospitalized on 24 April 2020 and admitted before 22 April 2020, was used to develop the
137 prediction model (derivation sample) and a sample of consecutive patients who visited the ED
138 between 15 and 25 March 2020 was used to perform the historical external validation (validation
139 sample).

140 Adult patients (≥ 18 years old) were eligible if at least one nasopharyngeal swab to search for the
141 SARS-CoV-2 genome was performed, independently of the presence of the respiratory disease.

142 The exclusion criteria were the following: current pregnancy, admission to pediatrics, obstetrics,
143 surgery, psychiatric departments, or the high-intensity units, including critical care, at the time of
144 inclusion.

145 All inpatients were tested for SARS-CoV-2 infection before the admission, while patients visiting
146 ED were tested in the judgment of the attending physician. All the swab tests were processed at
147 Niguarda Hospital.

148 The study is conformed to Helsinki's Declaration and was approved by the ethics committee Milano
149 Area 3 (register number 249-13052020). An informed consent was provided by the enrolled
150 participants.

151 **Outcome and predictor measurements**

152 The SARS-CoV-2 infection as the outcome was defined if at least one out of three SARS-Cov-2
153 genes tested on at least one nasopharyngeal swab was detected by RT-PCR.

154 For the derivation sample, the inpatients' medical records were completed by interviews using *ad-*
155 *hoc* case report forms to collect data about age, gender, communal living situations, cohousing
156 and/or residence in care facilities, known contact with COVID-19 case; the presence (ever) of
157 fever, cough, dyspnea, pharyngodynia, hypo-/anosmia and/or hypo-/ageusia, nausea and/or vomit,
158 diarrhea; current history of syncope, seizure (if unprecedented), and/or conjunctivitis; systolic and
159 diastolic blood pressures, heart and respiratory rates per minute, body temperature ($^{\circ}\text{C}$), peripheral

160 oxygen saturation while breathing ambient air; the total leukocyte, neutrophil, lymphocyte, and
161 monocyte counts, and C-reactive protein levels (mg/dL, normal reference <0.5). The patient's
162 current medical history, symptoms and signs, vital signs, and laboratory exams at the time of the
163 ED admission were considered as predictors.

164 The data needed for the validation sample were extracted by three physicians who reviewed the
165 electronic medical records of the ED visits.

166 Data collection was unblinded to the outcome in both samples.

167 **Model development and validation methods**

168 The association between the dichotomic outcome and each predictor was explored by univariate
169 analysis with logistic regression models. The distribution of non-continuous predictors was checked
170 and when the absolute frequency was less than 5 per table, they were not included in the model
171 development due to potential separation and overfitting issues. Only main fixed effects were
172 considered for inclusion in the model while interactions terms as well as higher order polynomials
173 were not explored. A backwards stepwise algorithm based on the Akaike's information criterion
174 was used with bootstrap resampling ($B = 1000$) and the covariates whose statistical significance
175 was detected in more than 80% of the bootstrap samples were chosen. The candidate models
176 were compared in terms of performance by C-statistic, explained variation (McFadden's R^2 ; scaled)
177 Brier score, Hosmer-Lemeshow test (grouping the observed outcomes by decile of predictions),
178 and sensitivity and specificity (the Youden's J statistic was used to identify the cutpoint for their
179 maximization). The net reclassification improvement [6] was considered to further assess the effect
180 of adding new predictors and the Decision Curve Analysis [7] was performed to evaluate the
181 thresholds, as well as the quality of the predictions. Then, these models were internally validated
182 by bootstrapping ($B = 1000$) [8] for estimation of the optimism-corrected C-statistic and calibration
183 (slope and intercept). The original β coefficients of the candidate logistic models were used to
184 calculate the predicted probabilities in an historical sample to assess their external validation and
185 the above-mentioned performance measures were used. To explore the impact of overfitting due to

186 small samples, a sensitivity analyses was performed by using shrunk [9] and penalized [10]
187 estimations of the β coefficients and these predicted probabilities were compared.

188 Missing data were firstly managed by exploring the absence of remarkable differences between the
189 complete cases and the group with missing data. Then, after checking the arbitrary-pattern of
190 missing data and testing for the assumption of missing completely at random (Little's chi-squared
191 test) a complete case analysis was chosen both for the derivation and the validation samples.

192 The power to detect a statistically significant difference in the performance of the model at a
193 particular sample size was calculated with standard formulas based on the normal distribution. The
194 estimated size of the validation sample for a one-sample two-sided mean test was 351 to achieve
195 80% power if the mean difference between C-statistics 0.03 with standard deviation 0.20, and the
196 type I error 0.05 were assumed.

197 All the analyses were performed using Stata Statistical Software Release 15 (StataCorp. 2017,
198 College Station, TX: StataCorp LLC) and R (R Core Team 2018, R Foundation for Statistical
199 Computing, Vienna, Austria).

200 The TRIPOD checklist for transparent reporting of a multivariable prediction model for individual
201 prognosis or diagnosis was used as guidance for reporting the final version of this article [11]. The
202 PROBAST tool was considered to discuss the potential risk of bias and concerns regarding model
203 development and applicability [12].

204

205 **RESULTS**

206 **Study population**

207 A total of 576 patients were included in this study: 141 in the derivation sample and 435 in the
208 validation sample. Of them, 28 patients in the derivation sample and 16 patients in the validation
209 sample were excluded because of missing data, consequently 113 and 419 patients were
210 analyzed, respectively (Figure 1).

211 The baseline characteristics of included patients are reported in Table 1.

212 Regarding the case-mix, a significant higher proportion of positive outcome (71 out of 141, 50% vs
213 260 out of 435, 60%; $p=0.047$), male patients (73 out of 141, 52% vs 267 out of 435, 61%,
214 $p=0.044$) with fever (67 out of 139, 48% vs 300 out of 435, 69%, $p\leq 0.01$) were observed in the
215 validation sample compared to the derivation sample, while in the same sample age and neutrophil
216 counts were lower (respectively, mean [SD] age equal to 72 [14] vs 61 [16] years, $p < 0.01$, and
217 neutrophil mean [SD] count equal to 7.3 [5.3] vs 6.4 [4.4] 10^9 cells/L) (Table 1).

218 The differences in terms of outcome and predictors between the patients analysed and those
219 excluded were not statistically significant with the exception of a higher prevalence of cough in the
220 complete cases (39 out of 113, 34% vs 3 out of 26, 11%, $p=0.03$) of the derivation sample and a
221 lower frequency of hospitalization (363 out of 419, 87% vs 10 out of 16, 62%, $p\leq 0.01$) in the
222 validation set (Supplementary Table 1).

223 The average time between the patient's assessment and the test result of swab was less than 24
224 hours.

225

226 **Prediction models**

227 In the derivation sample, age, oxygen saturation at room air, fever, cough, and both total white
228 blood cell and neutrophil counts were significantly associated in the univariate analyses (Table 2).

229 In the selection of the predictors, only age, oxygen saturation at room air, fever, and the neutrophil
230 count (total white blood cell was excluded due to high correlation with the individual cell count)
231 showed a statistical significance in more than 80% of the 1000 bootstrap samples and they were
232 considered for model building.

233 Three multivariable models were consequently developed: an oversimplified one with fever and
234 oxygen saturation only (M1), and two models with the clinical predictors in absence (M2) and
235 presence (M3) of the results of the laboratory test, and their apparent performance was tested as
236 reported in Table 3.

237 The discrimination was >0.80 in M2 and M3, but the net reclassification improvement due to the
238 addition of the neutrophil count as predictor was not statistically significant. In the view of Decision

239 Curve Analysis, M2 was steadily superior to M1, and inferior to M3, but not for all the threshold
240 points. When these models were internally validated, the correction of the calibration slope for
241 optimism was <10% in M1 and M2 only.

242

243 **Validation results**

244 The external validation of the models and the sensitivity analyses of the regression coefficients by
245 uniform shrinking and penalization are detailed in Table 4 and Supplementary Table 2.

246 The C-statistic was significantly higher for M3 (0.819, 95%CI 0.775-0.862) compared to M2 (0.793,
247 95% 0.747-0.840, $p=0.014$), yet the difference was limited (2.6%) and statistically not significant in
248 comparison with M1 (0.817, 95% CI 0.774-0.861, $p=0.933$). However, only M2 showed to be
249 sufficiently well calibrated if shrunk or penalized coefficients were applied instead of the original
250 values. The calibration plots based on the original coefficients are shown in Figure 2. The
251 sensitivity was consistently >85% for all the models, but the specificity was >70% in M3 only when
252 the cut-off was statistically calculated to maximize both the performance measures (Supplementary
253 Table 2). In Figure 3, the curves show that each model is clinically useful for thresholds in the
254 range of 20-80% for the probability of the outcome. Moreover, each model showed to be superior
255 to the use of age alone as well as the oxygen saturation at room air, while the history of fever as
256 the only predictor seemed to be comparable.

257

258 **DISCUSSION**

259 Niguarda is one of the largest General Hospital in the North of Milan within a Metropolitan Area of
260 3,279,944 inhabitants (January 2020), and hosts all the medical and surgical disciplines for adults
261 and children, including a 24-hour ED with 96,588 visits and 32,612 in hospital admissions covering
262 every intensity of care in 2019.

263 In this hospital context, a prediction model for the prompt diagnosis of COVID-19 high risk was
264 developed in hospitalized patients and externally validated in the setting of the ER.

265 The history of fever, elder age and low oxygen saturation showed to be significant predictors of the
266 presence of SARS-CoV-2 infection in the derivation sample. The use of laboratory findings, such
267 as neutrophil count showed to improve the discriminative ability of the model, even if its calibration
268 and usefulness in terms of diagnosis is unclear.

269 Since the beginning of the pandemic, a remarkable number of prediction models have been
270 developed and critically appraised in a systematic review updated to July 2020 [5]. Out of 232
271 models, 118 were diagnostic models for detecting COVID-19. Of them, 75 were based on medical
272 imaging and 10 were used to diagnose disease severity. Only few diagnostic models were
273 externally validated, and all the appraised models were judged to be flawed by a high risk of bias
274 [5].

275 Age was included in most of the final models, while the inclusion of fever or varying laboratory
276 findings, including neutrophil count, were inconsistent and the oxygen saturation was considered
277 only in prognostic models (107 out of 232), despite its clinical importance in the early management
278 of disease [5].

279 Among the prediction models appraised, the diagnostic model by Jehi et al. was identified as
280 promising, even if requires external validation. It defines male, African–American, older patients,
281 and those with known COVID-19 exposure at higher risk of being positive for COVID-19 [13]. A
282 reduced risk was reported for patients who had influenza or pneumococcal polysaccharide vaccine
283 or who were on melatonin, paroxetine or carvedilol. Nevertheless, the inclusion of race as predictor
284 is subjected to a possible bias due to American healthcare inequality, and the COVID-19 exposure
285 is a data reported by patients [13].

286 The role of social variables within a COVID-19 predictive model was explored by Chew and
287 collaborators, who derived two models which include, among different risk factors, dormitory
288 residence in the first and contact with infective patients in the second model [14].

289 Considering the dormitory residence, as the institutions dedicated to migrants, authors conclude
290 that this represents an important predictor for COVID-19 infection, but this result is not confirmed if
291 we consider other types of institutions, such as retirement homes. This result is therefore not
292 generalizable to realities other than Singapore.

293 Considering contact with COVID-19 infective patients, authors highlight the role of traceability as
294 prognostic factor in the development of the disease. On the other hand, contact tracing is not easy
295 to execute, both in traditional and alternative ways, because of privacy and socio-cultural aspects
296 [15].

297 Both considerations could explain the lack of predictability of residency data in some specific
298 contexts and exposure data in the present study [14].

299 Of note, Italian studies are increasing accordingly to the last update of the systematic review which
300 counts a total of 23 (10%) Italian models, 21 more than the first version of April 2020.

301 The discriminative ability of our models is consistent with those whose performance was assessed
302 by external validation with C-statistics ranging from 0.73 to 0.91. Moreover, their calibration was
303 superior to that of the only study where the calibration slope was reported (0.77 to 0.95 vs 0.56,
304 respectively) [16].

305 With regards to the setting, the only study based on patients from ED reported a C-statistic
306 calculated by training test split (0.85) comparable to those of our models in the derivation sample,
307 but the number of predictors was higher (15 versus 2 to 4 variables) and its external validation was
308 not provided, and this may raise concerns about overfitting issues [17]. Moreover, another
309 prediction model for diagnosis in Italy was developed to assess the probability of community-
310 acquired pneumonia due to COVID-19 from multiple centres and its discriminative ability was
311 similar (C-statistic 0.84) [18]. However, this comparison may be flawed by the differences between
312 the purposes of the models as well as their settings (hospitalized patients vs ED admissions).

313 The implementation of our models is based on a limited number of predictors which are easy to be
314 measured as well as promptly available in everyday practice. Moreover, its temporal external
315 validation showed to be robust to the differences in the case-mix and particularly when the
316 prevalence of the outcome and the proportion of hospitalization were variable in real-life
317 consecutive sample of patients from the ED setting.

318 This study has several limitations. First, it was developed by using a case-control design and case
319 analysis and data collection were retrospective and unblinded. However, since hard outcome and
320 predictors like the detection of SARS-CoV-2 genome in the nasopharyngeal swab and age were

321 used, a limited impact of bias on predictions may be assumed. Secondly, predictors like communal
322 living and history of contact with a person diagnosed of SARS-CoV-2 infection were unmeasured,
323 even though their relevance for diagnostic predictions has been still to be determined. Finally, the
324 discriminatory ability of the models makes the overall performance still suboptimal and its
325 implementation to calculate the individual risk of infection should not be used without additional
326 investigations.

327 However, it could be considered to evaluate the spatial allocation of the patients while awaiting the
328 result of the nasopharyngeal swab, which is still the current reference standard for the diagnosis of
329 SARS-Cov-2 infection.

330 In conclusion, the development and validation of these models showed that prediction tools based
331 on clinical findings like history of fever, age, and oxygen saturation at presentation may be
332 accurate and robust to identify patients at high risk for a diagnosis of SARS-CoV-2 infection.

333 Future external validation studies should be considered to evaluate if such models will be also
334 robust to variable outcome prevalence, particularly in low proportions of infection, and to search for
335 additional predictors which could increase the model performance for individual risk prediction.

336

337

Table 1. Baseline characteristics of included patients and case-mix features			
Features at presentation	Development sample (n=141)	Validation sample (n=435)	p-value
Age (years), mean (SD) [range]	72 (14) [27–96]	61 (16) [19–93]	<0.01
Male	73 (51.8)	267 (61.4)	0.044
Nasopharyngeal swab SARS-Cov-2 positivity	71 (50.3)	260 (59.8)	0.047
Report of contact with a person diagnosed of SARS-Cov-2 infection	17/127 (13.4)	–	–
Communal living	9/135 (6.7)	–	–
Hospitalization	141 (100)	373 (85.7)	<0.01
Fever	67/139 (48.2)	300 (69.0)	<0.01
Dyspnea	42/139 (30.2)	–	–
Cough	42/139 (30.2)	–	–
Pharyngodynia	1/139 (0.7)	–	–
Hypo/anosmia	1/139 (0.7)	–	–
Dysgeusia	4/139 (2.9)	–	–
Conjunctivitis	0/139 (0)	–	–
Nausea and/or vomit	18/139 (12.9)	–	–
Diarrhea	11/139 (7.9)	–	–

Seizure	4/139 (2.9)	–	–
Syncope	14/139 (10.1)	–	–
Body temperature (°C)	36.6 (1.0) [35–39.8] [^]	–	–
Oxygen saturation (room air), %	94.4 (4.2) [78–100] [*]	93.6 (6.1) [40–100] [†]	0.148
Systolic blood pressure (mmHg)	140 (31) [70–230] [*]	–	–
Diastolic blood pressure (mmHg)	73 (14) [40–123] [*]	–	–
Heart rate per minute	91 (20) [45–70] [§]	–	–
Respiratory rate per minute	21 (6) [12–40] [#]	–	–
White blood cell count (10 ⁹ /L)	9.8 (6.7) [1.4–51.2]	–	–
Neutrophil count (10 ⁹ /L)	7.3 (5.3) [1.1–42.5] [°]	6.4 (4.4) [0.7–46.2] [‡]	0.046
Lymphocyte count (10 ⁹ /L)	1.6 (2.1) [0.1–16.8] [°]	–	–
Monocyte count (10 ⁹ /L)	0.8 (1.5) [0.04–16.5] [°]	–	–
C-reactive protein (mg/dL)	6.8 (7.9) [0–43.4]	–	–
Data presented as n (%) and mean (SD) [range].			
[^] 4/141 (2.8%) missing; [*] 2/141 (1.4%) missing; [§] 1/141 (0.7%) missing; [#] 24/141 (8.3%) missing; [°] 3/141 (2.1%) missing; [†] 2/435 (0.5%) missing; [‡] 15/435 (3.4%) missing.			
SD, standard deviation			

338

339

340

Table 2. Predictors included in the selection process and results of their univariate analyses in the two study samples.		
	OR (95% CI)	
	Derivation sample (n=113)	Validation sample (n=419)
Age, 1-year increase	0.95 (0.93-0.99)	0.99 (0.97-1.00)
Male, female as reference	1.19 (0.57-2.50)	1.21 (0.81-1.81)
Fever, absence as reference	5.88 (2.61-13.23)	18.64 (10.87-31.98)
Dyspnoea, absence as reference	1.59 (0.72-3.55)	–
Cough, absence as reference	3.3 (1.45-7.52)	–
Nausea and/or vomit, absence as reference	0.71 (0.23-2.18)	–
Syncope, absence as reference	0.63 (0.17-2.36)	–
Body temperature (°C)	1.44 (0.95-2.13)	–
Oxygen saturation (room air), 1% increase	0.86 (0.78-0.96)	0.86 (0.81-0.90)
Systolic blood pressure, 10-mmHg increase	1.02 (0.91-1.14)	–
Diastolic blood pressure, 10-mmHg increase	1.07 (0.84-1.37)	–
Heart rate, 10-beat-per-minute increase	1.07 (0.89-1.30)	–
Respiratory rate, 1-breath-per-minute increase	1.01 (0.95-1.07)	–
White blood cell count, $1 \times 10^9/L$ increase	0.86 (0.79-0.95)	–

Neutrophil count, $1 \times 10^9/L$ increase	0.87 (0.78-0.96)	0.88 (0.83-0.93)
Lymphocyte count, $1 \times 10^9/L$ increase	0.61 (0.37-1.02)	–
Monocyte count, $1 \times 10^9/L$ increase	0.70 (0.39-1.26)	–
C-reactive protein, 1-mg/dL increase	0.96 (0.92-1.01)	–
OR, odds ratio; CI, confidence interval.		

341

342

343

Table 3. Performance of the three multivariable models in the development sample (n=113)			
	OR (95% CI)		
	M1	M2	M3
Age, 1-year increase	–	0.95 (0.92-0.98)	0.95 (0.92-0.99)
Fever, presence	5.63 (2.44-13.0)	4.89 (2.03-11.75)	4.80 (1.86-12.39)
Oxygen saturation (room air), 1% increase	0.87 (0.78-0.98)	0.83 (0.73-0.94)	0.78 (0.69-0.89)
Neutrophil count, 1×10 ⁹ cell/L increase	–	–	0.81 (0.71-0.92)
Intercept	1.65 E(05)	8.86 E(08)	8.13 E(11)
Performance measures			
C-statistic (95% CI)	0.7730 (0.6856-	0.8076 (0.7279-	0.8537 (0.7857-
• Optimism-corrected	0.8604)	0.8874)	0.9217)
• Optimism	0.7696	0.7959	0.8353
	-0.0034	-0.0117	-0.0184
Hosmer-Lemeshow test*, p-value	0.77	0.29	0.60
Calibration-in-the-large corrected for optimism	-0.0010	-0.0025	+0.0027
Slope corrected for optimism	0.9810	0.9431	0.8922
Brier score	0.1949	0.1781	0.1571
Scaled	0.2204	0.2876	0.3714
Akaike information criterion	136.001	128.292	117.077
Bayesian information criterion	144.183	139.2014	130.714
McFadden's R ²	0.1701	0.2321	0.3164
Cutpoints			

• Youden empirical	0.3818	0.4576	0.6398
• Sensitivity	0.82	0.79	0.70
• Specificity	0.63	0.70	0.88
Net Reclassification Improvement	-	-0.019 [§]	+0.053 [§]
Decision Curve Analysis, net reduction in intervention (%)			
Threshold point			
0.1	0	1.8	3.5
0.2	1.8	18.1	18.1
0.4	25.4	28.0	31.6
0.6	25.4	28.0	31.6
0.8	37.2	39.6	40.7
M, model; OR, odds ratio; CI, confidence interval; [§] p-value >0.05; *the observed outcomes were grouped by decile of predictions.			

344

345

Table 4. Intercept and coefficients of the three diagnostic models (original and values modified by uniform shrinking and penalization) and their discrimination performance in the validation sample (n=419).

MODEL			Age, 1- year increase	Fever, presenc e	Oxygen saturatio n (room air), 1% increase	Neutrophil count, 1*10 ⁹ cell/L increase	Intercept	DISCRIMINATION	C- statistic (95% CI)
			β_{original}	-	1.727861	-0.136564	-		12.013340
M1		β_{shrunk} (s=0.9249)	-	1.598142	-0.126312	-	11.111438	0.817 (0.774- 0.861)	
		$\beta_{\text{penalized}}$	-	1.688091	-0.128422	-	11.266010	0.819 (0.776- 0.862)	
		β_{original}	-0.050820	1.586681	-0.187662	-	20.602260	0.793 (0.747- 0.840)	
M2		β_{shrunk} (s=0.9175)	-0.046625	1.455731	-0.172174	-	18.901936	0.793 (0.747- 0.840)	
		$\beta_{\text{penalized}}$	-0.048163	1.531269	-0.175829	-	19.286960	0.794 (0.747- 0.841)	
		β_{original}	-0.050992	1.569582	-0.244274	-0.207961	27.423690	0.819 (0.775- 0.862)	

M3	β_{shrunk} (s=0.9193)	-0.046877	1.442926	-0.224563	-0.191180	25.210763	0.819 (0.775- 0.862)
	$\beta_{\text{penalized}}$	-0.047870	1.496003	-0.225975	-0.193822	25.403820	0.819 (0.776- 0.863)
β , coefficient of the logistic regression from the development sample; CI, confidence interval.							

346

347

348 **Supplementary Material:**

349 Supplementary Table 1.

350 Supplementary Table 2.

351

352 **REFERENCES**

353 1. <https://covid19.who.int> access 11 March 2021

354 2. Ministry of Health/Civil Protection data, accessed 11 March 2021
355 <https://opendatadpc.maps.arcgis.com/>

356 3. Kevadiya BD, Machhi J, Herskovitz J, Oleynikov MD, Blomberg WR, Bajwa N, Soni D, Das
357 S, Hasan M, Patel M, Senan AM, Gorantla S, McMillan J, Edagwa B, Eisenberg R, Gurumurthy
358 CB, Reid SPM, Punyadeera C, Chang L, Gendelman HE (2021) Diagnostics for SARS-CoV-2
359 infections. *Nat Mater*. doi: 10.1038/s41563-020-00906-z.

360 4. Kurstjens S, van der Horst A, Herpers R, Geerits MWL, Kluiters-de Hingh YCM, Göttgens
361 EL, Blaauw MJT, Thelen MHM, Elisen MGLM, Kusters R (2020) Rapid identification of SARS-CoV-
362 2-infected patients at the emergency department using routine testing. *Clin Chem Lab Med*.
363 58(9):1587-1593. doi: 10.1515/cclm-2020-0593. PMID: 32598302.

364 5. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, Bonten MMJ, Dahly
365 DL, Damen JAA, Debray TPA, de Jong VMT, De Vos M, Dhiman P, Haller MC, Harhay MO,
366 Henckaerts L, Heus P, Kammer M, Kreuzberger N, Lohmann A, Luijken K, Ma J, Martin GP,
367 McLernon DJ, Andaur CL, Reitsma JB, Sergeant JC, Shi C, Skoetz N, Smits LJM, Snell KIE,
368 Sperrin M, Spijker R, Steyerberg EW, Takada T, Tzoulaki I, van Kuijk SMJ, van Bussel B, van
369 Royen FS, Verbakel JY, Wallisch C, Wilkinson J, Wolff R, Hooft L, Moons KGM, van Smeden M
370 (2020) Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and
371 critical appraisal. *BMJ*. 369:m1328. doi: 10.1136/bmj.m1328 [update 3].

- 372 6. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS (2008) Evaluating the added
373 predictive ability of a new marker: from area under the ROC curve to reclassification and beyond.
374 *Stat Med.* 27(2):157-72; discussion 207-12. doi: 10.1002/sim.2929.
- 375 7. Vickers AJ, Elkin EB (2006) Decision curve analysis: a novel method for evaluating
376 prediction models. *Med Decis Making.* 26(6):565-74. doi: 10.1177/0272989X06295361.
- 377 8. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD
378 (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression
379 analysis. *J Clin Epidemiol.* 54(8):774-81. doi: 10.1016/s0895-4356(01)00341-9.
- 380 9. Copas JB (1983) Regression, prediction and shrinkage. *J R Stat Soc Ser B* 45(3):311–354
- 381 10. Firth D (1993) Bias reduction of maximum likelihood estimates. *Biometrika* 80:27-38.
- 382 11. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable
383 prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.*
384 2015 Jan 7;350:g7594. doi: 10.1136/bmj.g7594. PMID: 25569120.
- 385 12. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB,
386 Kleijnen J, Mallett S; PROBAST Group (2019) PROBAST: a tool to assess the risk of bias and
387 applicability of prediction model studies. *Ann Intern Med.* 170(1):51-58. doi: 10.7326/M18-1376.
- 388 13. Jehi L, Ji X, Milinovich A, Erzurum S, Rubin BP, Gordon S, Young JB, Kattan MW (2020)
389 Individualizing risk prediction for positive coronavirus disease 2019 testing: results from 11,672
390 patients. *Chest.* 158(4):1364-1375. doi: 10.1016/j.chest.2020.05.580.
- 391 14. Chew WM, Loh CH, Jalali A, Fong GSE, Senthil Kumar L, Sim RHZ, Tan RP, Gill SR, Liang
392 TR, Koh JMK, Tay TR (2021) A risk prediction score to identify patients at low risk for COVID-19
393 infection. *Singapore Med J.* doi: 10.11622/smedj.2021019.
- 394 15. Montanari Vergallo G, Zaami S, Marinelli E (2021) The COVID-19 pandemic and contact
395 tracing technologies, between upholding the right to health and personal data protection. *Eur Rev*
396 *Med Pharmacol Sci.* 25(5):2449-2456. doi: 10.26355/eurrev_202103_25286.

- 397 16. Wang F, Hou H, Wang T, Luo Y, Tang G, Wu S, Zhou H, Sun Z (2020) Establishing a
398 model for predicting the outcome of COVID-19 based on combination of laboratory tests. *Travel*
399 *Med Infect Dis.* 36:101782. doi: 10.1016/j.tmaid.2020.101782.
- 400 17. de Moraes Batista AF, Miraglia JL, Rizzi Donato TH, Porto Chiavegatto Filho AD (2020)
401 COVID-19 diagnosis prediction in emergency care patients: a machine learning approach
402 (preprint). *Epidemiology.* <https://doi.org/10.1101/2020.04.04.20052092>
- 403 18. Sambataro G, Giuffrè M, Sambataro D, Palermo A, Vignigni G, Cesareo R, Crimi N, Torrissi
404 SE, Vancheri C, Malatino L, Colaci M, Del Papa N, Pignataro F, Roman-Pognuz E, Fabbiani M,
405 Montagnani F, Cassol C, Cavagna L, Zuccaro V, Zerbato V, Maurel C, Luzzati R, Di Bella S (2020)
406 The Model for Early COvid-19 Recognition (MECOR) Score: A Proof-of-Concept for a Simple and
407 Low-Cost Tool to Recognize a Possible Viral Etiology in Community-Acquired Pneumonia Patients
408 during COVID-19 Outbreak. *Diagnostics (Basel).* 10(9):619. doi: 10.3390/diagnostics10090619.

Figures

Development and external validation of a diagnostic multivariable prediction model for a prompt identification of cases at high risk for SARS-COV-2 infection among patients admitted to the emergency department.

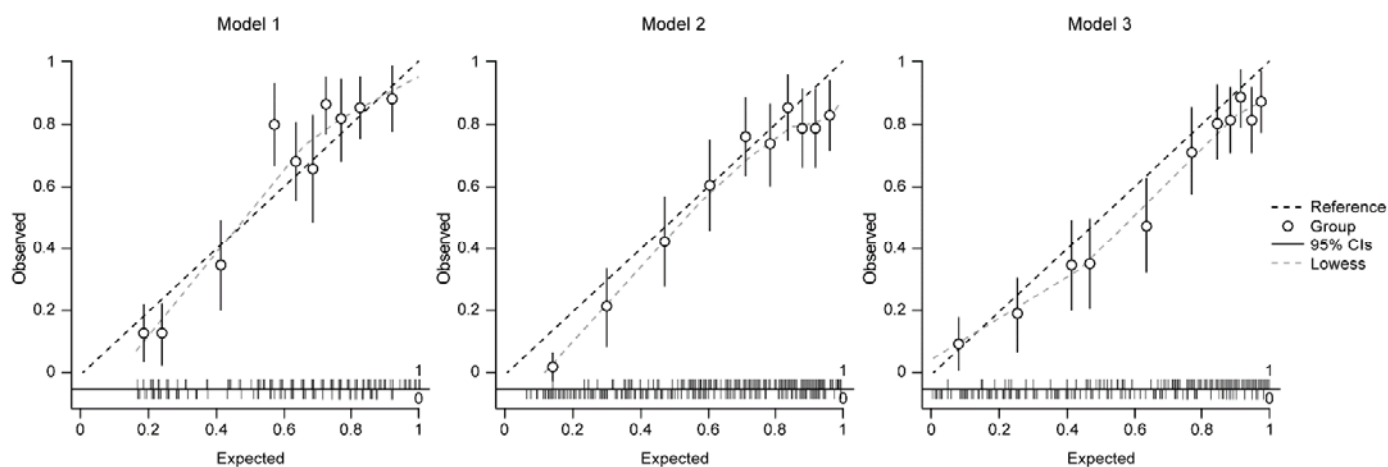


Figure 1. Calibration plots of the three models with the original coefficients in the external validation. CI: Confidence Interval.

Figures

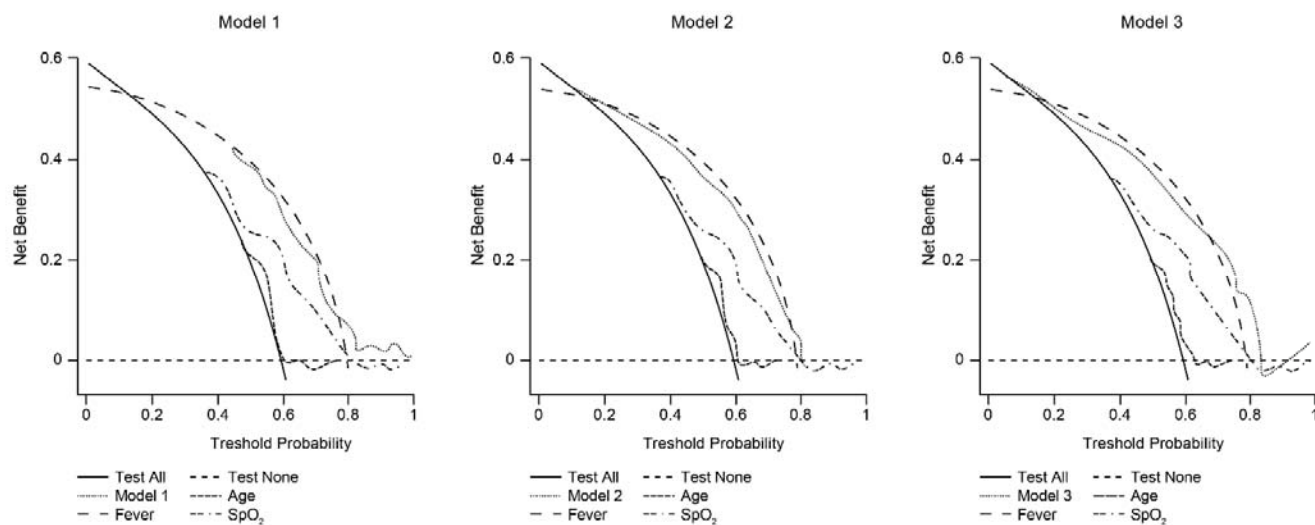


Figure 2. The Decision Curve Analysis of the three models applied into the validation sample in comparison with the use of the individual predictors across different probability thresholds for clinical decision. SpO₂: oxygen saturation.