Assessing the capacity of symptom scores to predict COVID-19 positivity in Nigeria: a national derivation and validation cohort study


ABSTRACT

Objectives This study aimed to develop and validate a symptom prediction tool for COVID-19 test positivity in Nigeria.

Design Predictive modelling study.

Setting All Nigeria States and the Federal Capital Territory.

Participants A cohort of 43221 individuals within the national COVID-19 surveillance dataset from 27 February to 27 August 2020. Complete dataset was randomly split into two equal halves: derivation and validation datasets. Using the derivation dataset (n=21 477), backward multivariable logistic regression approach was used to identify symptoms positively associated with COVID-19 positivity (by real-time PCR) in children (<17 years), adults (18–64 years) and elderly (>65 years) patients separately.

Outcome measures Weighted statistical and clinical scores based on beta regression coefficients and clinicians’ judgements, respectively. Using the validation dataset (n=21 744), area under the receiver operating characteristic curve (AUROC) values were used to assess the predictive capacity of individual symptoms, unweighted score and the two weighted scores.

Results Overall, 27.6% of children (4415/15 988), 34.6% of adults (9154/26 441) and 40.0% of elderly (317/792) that had been tested were positive for COVID-19. Best individual symptom predictor of COVID-19 positivity was loss of smell in children (AUROC 0.56, 95% CI 0.55 to 0.56), either fever or cough in adults (AUROC 0.57, 95% CI 0.56 to 0.58) and difficulty in breathing in the elderly (AUROC 0.53, 95% CI 0.48 to 0.58) patients. In children, adults and the elderly patients, all scoring approaches showed similar predictive performance.

Conclusions The predictive capacity of various symptom scores for COVID-19 positivity was poor overall. However, the findings could serve as an advocacy tool for more investments in resources for capacity strengthening of molecular testing for COVID-19 in Nigeria.

INTRODUCTION

The index case of COVID-19 in Nigeria was recorded on 27 February 2020, in Ogun State; one of the contacts of the index case was later diagnosed with the disease in the same state. As of epidemiological week 44 (26 October–1 November), Nigeria had recorded 62 964 confirmed cases, of whom 1146 died and 58 790 recovered, resulting in a case fatality rate of 1.8%. The rapid increase in...
COVID-19 cases in Nigeria caused a strain on the healthcare systems, particularly in the area of molecular testing using real-time PCR (RT-PCR). Deficiencies in molecular laboratory capacity can hinder the rapid identification of persons with COVID-19 infection and initiation of appropriate treatment and public health measures.\(^5\) For example, as of 9 April 2020, only 5000 tests for SARS-CoV-2 were performed across Nigeria by eight laboratories across six states.\(^3\) Concerted efforts have since been made to expand molecular testing capacity across the country, with 675183 tests for SARS-CoV-2 conducted and the Nigeria Centre for Disease Control (NCDC) COVID-19 laboratory networks expanded to 69 functioning laboratories as of 6 November 2020.\(^1\) However, with a population of over 200 million and community transmission of COVID-19, the testing rate for Nigeria to date is still well below the Africa Centre for Disease Control-set target of 1% for population. Generally, laboratory capacity to test for SARS-CoV-2 remains a challenge in many sub-Saharan African countries, especially in the early phase of the COVID-19 pandemic.\(^3\)

Although PCR testing is considered the gold standard for COVID-19 diagnosis, limited access and centralised laboratory systems often contribute to delay.\(^6\) This has contributed to making diagnostic results unavailable for urgent decision-making by frontline healthcare workers, with consequent effect on healthcare delivery and increased risk of nosocomial infection.\(^7\) Between 27 February and 6 June 2020, for example, the median turnaround time for laboratory diagnosis of COVID-19 in Nigeria was 2 (IQR 1–4) days; 9% of the 12289 confirmed COVID-19 cases during this period was recorded among healthcare workers.\(^8\) Thus, to mitigate the impact of COVID-19 on the healthcare system, while also improving community surveillance and care for patients, rapid and accurate diagnosis are crucial.\(^9\) NCDC and its technical partners are currently conducting a study to validate rapid diagnostic test kits to determine their accuracy and clinical utility in Nigeria.

One approach to supplement laboratory testing is, however, the use of diagnostic prediction models that use clinical variables (including standard blood counts when available) to estimate the likelihood of individuals being COVID-19 positive or experiencing poor clinical outcome.\(^9\) A systematic review of developed and/or validated prediction models found the most frequently reported predictors of COVID-19 diagnosis to be influenza-like symptoms (eg, chills, fatigue) and neutrophil count, with C-statistic estimates ≥0.90 (an excellent discriminative performance).\(^9\) However, to date, these prediction models have been judged to be at a high risk of bias, largely because they were fitted on data that were not representative of the target population and have suffered from poor reporting on intended use of the models or their calibration performance. While the reported performances of models to date are encouraging, their adoption in clinical practice has not been recommended.\(^9\) In addition, there has been a dearth of evidence from sub-Saharan Africa as the majority of available models use data from China and European countries, thus limiting their generalisability to settings such as Nigeria with different COVID-19 epidemiology, healthcare systems, socioeconomic conditions and health-seeking behaviours.

The majority of confirmed COVID-19 cases in Nigeria are asymptomatic at the point of diagnosis;\(^8\) for the confirmed cases who present with symptoms, the most common symptoms tend to be non-specific (eg, fever, cough and difficulty in breathing).\(^8\)\(^10\)\(^11\) Using data from 27 February to 6 June 2020, presentation with cough, fever, loss of smell and loss of taste was found to be positively and independently associated with COVID-19 positivity in Nigeria.\(^12\) But the dearth of evidence on the predictive capacity of clinical symptoms to predict COVID-19 in Nigeria (and indeed in sub-Saharan Africa) is the rationale for this study. Thus, we aimed to develop and validate the predictive capacity of clinical signs and symptoms with regard to testing positive for COVID-19 and to investigate whether there are any gains in the predictive capacity of statistically and clinically derived weighted combined symptom scores, as compared with an unweighted combined score in Nigeria.

**METHODS**

**Study design and data source**

This is a predictive modelling study using a retrospective cohort of persons enrolled in the Surveillance, Outbreak Response Management and Analysis System (SORMAS) database from 27 February to 27 August 2020. A detailed description of SORMAS in the COVID-19 context in Nigeria is available elsewhere.\(^8\) Briefly, SORMAS is an open-source real-time electronic health surveillance and laboratory database, which has been in use in Nigeria since 2017. SORMAS is hosted and coordinated at the NCDC’s headquarter in Abuja.

**Study participants**

The study population comprised children (≤17 years), adults (18–64 years) and the elderly (≥65 years) who were tested for SARS-CoV-2 infection. Eligibility for RT-PCR test during the study period was based on the NCDC COVID-19 standard case definition for suspected cases (online supplemental table 1).\(^13\) However, provisions were made to test persons concerned about COVID-19 infection on presentation to designated testing centres regardless of whether they met the NCDC case definitions or not.

**Data collection and management**

Sociodemographic (eg, age, sex, occupation, education and geopolitical zone of residence) and clinical presentation (signs and symptoms in the 14 days prior to testing) were collected from all persons tested. The collection and transportation of respiratory samples (oropharyngeal and nasopharyngeal swabs) for laboratory analysis were facilitated by trained healthcare workers according to the

NCDC guidelines. Testing for SARS-CoV-2 by RT-PCR was performed according with the WHO guidelines. All the collected data across the country, including laboratory diagnostic outcomes, were submitted by trained healthcare workers to NCDC via the SORMAS application installed on tablets or laptops. Patients who tested positive for COVID-19 were managed either in a government health facility or at home, depending on the illness severity or availability of bed space at health facilities, while observing the NCDC case management guidelines.

Since this was an analysis of a secondary dataset, there was no formal sample size calculation for this study; however, the study met the standard sample size requirement of 10 outcome events per degree of freedom in prediction models (eg, 10 binary variables in the model require 100 COVID-19 positive cases). We used the train-test split-sample method to randomly split the complete dataset into two (ie, training and testing datasets), assigning half of the records to either the training (derivation) or testing (validation) datasets based on a random number generated within the statistical software. However, we made four major assumptions to define the study eligibility criteria. First, we assumed that a missing variable was indicative of absence (ie, record not present), supported by findings ($p<0.001$) from a $\chi^2$ test of association between two ad hoc variables ‘any-missing’ and ‘any-absent’. Second, we decided that eligible study participants needed to have complete records for both age and sex since these are crucial demographic variables. Third, given the high proportion of asymptomatic COVID-19 cases in Nigeria, the study participants needed to have had at least one symptom positively recorded (ie, symptomatic). Lastly, analyses were performed separately for children, adults and the elderly, given the evidence supporting the age dependence of COVID-19 symptom presentation, with children more likely to be asymptomatic than adults.

Figure 1 shows a flowchart showing the processes for selecting the study population as stated in the eligibility criteria (definitions of suspected and probable cases are available in the NCDC case definition). All symptoms were coded as ‘1’ if present and ‘0’ if not. Sex was classified as a binary variable, with ‘1’ and ‘0’ indicative of male and female gender, respectively.

**Outcome and predictor variables**

The outcome was COVID-19 positivity, defined as either presence or absence of SARS-CoV-2 by RT-PCR confirmation (yes/no). Clinical prediction variables were informed by evidence from our previous study using the same dataset, although it covered a shorter period (27 February to 5 June 2020). The clinical variables were collected using a combination of self-reports by COVID-19 suspected cases and/or their caretakers as well as by objective assessment by a healthcare worker. The

---

**Figure 1** Flowchart showing the processes for selecting records from the Surveillance, Outbreak Response Management and Analysis System.
three categories of their respective signs and symptoms (each coded yes/no) with their respective examples include (1) ‘obvious or as observed by a healthcare worker’ (chills, sweat, cough, breathing difficulty, rapid breathing and runny nose); (2) ‘elicitable or information volunteered or given on inquiry by patients/relatives’ (abdominal pain, diarrhoea, gastrointestinal tract symptoms, chest pain, fatigue/weakness, headache, musculoskeletal pain, sore throat, loss of taste and loss of smell); (3) and ‘measurable as assessed by a healthcare worker’ (fever, defined as an axillary temperature ≥37.5°C).

Statistical analyses

The data were set up and analysed using Stata V.13.0 (Stata Corporation, College Station, Texas, USA).

Model development

Using the derivation dataset, the association between individual symptoms and COVID-19 positivity was investigated, in turn, using univariable logistic regression. Analyses were stratified by children, adults, and the elderly. The results are presented as unadjusted ORs and 95% CIs. To determine which of the symptoms were independently associated with COVID-19 positivity, we first conducted a test for multicollinearity by the various age groups (children, adults and the elderly) using the variance inflation factor (VIF) and tolerance values of each statistically significant symptom identified from the univariable logistic regression model (see online supplemental table 2a, b and c for results). This was to ensure that symptoms were independent of each other or not collinear and therefore suitable for combining into a single score. Both VIF and tolerance values measure how much the regression coefficient for a symptom is determined by the other symptoms in the model; low VIF and high tolerance values suggest the absence of multicollinearity and vice versa.18 Thereafter, we used the backward multivariable logistic regression approach to select all the statistically significant symptoms for the model. As before, analyses were conducted for children, adults and the elderly separately.

We then created three different symptom scores. First, an unweighted combined score was calculated for each person (separately for children, adults and the elderly) by assigning ‘1’ point for each symptom before summation. The unweighted combined scores were included in the model as a continuous variable. The combined scores were further recoded as binary categorical variables to represent clinical scores ≥1 (yes/no), ≥2 (yes/no), ≥3 (yes/no) and ≥4 (yes/no). Additionally, weighted combined scores were developed using two approaches. First, we used the beta regression coefficients obtained from the multivariable logistic regression model that was run previously. Each symptom was given the value of ‘1’ if present and ‘0’ if otherwise. Each symptom was then multiplied by its weight, obtained by multiplying the respective beta regression coefficients (or log of the ORs) by 10. Essentially, the beta regression coefficients measure the relative prognostic strength of each symptom when they are included simultaneously in a multivariable regression model so that the bigger the value of a symptom, the more its weight. Second, we engaged six clinicians (including infectious disease consultants) who have been managing patients with COVID-19 in treatment centres and at home in the Federal Capital Territory, Gombe, Delta and Kaduna. Here, the clinicians were provided with a list of symptoms identified to be statistically significant from the univariable logistic models, with the freedom to add any missing sign or symptom in the model as needed. Each clinician was then asked to independently assign a weight of 1–5 for each symptom based on experiences from managing patients with COVID-19 (see online supplemental table 3 for detail of weighted scores assigned by each clinician for the identified symptoms). Similar to the statistical weighting approach, each symptom was multiplied by the average of the combined weights assigned by all the clinicians. Predictive capacity of the various score thresholds was also compared to see which combination of symptoms was more predictive of COVID-19 positivity.

Model validation

The different models were then applied to the validation dataset. Separately for children, adults and the elderly, predictive capacity was assessed in terms of sensitivity, specificity, area under the receiver operating characteristic (ROC) curve (AUROC) value, positive predictive value (PPV) and negative predictive value (NPV)19; and ROC curves were plotted for the comparison of clinical prediction scores.

Where applicable, this study is reported according to the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) statement.20

External calibration belt and test for model goodness of fit

Using the validation dataset, external calibration test and belt were used to evaluate the goodness of fit of our models, an approach that examines the relationship between estimated probabilities and observed outcome rates.21 Tests and belts often return concordant outputs: non-significant tests are often associated with the belt encompassing the 45 degree lines (good fit) and significant tests with the belt deviating from the bisector (poor fit).

Patient and public involvement

Being an analysis of deidentified secondary dataset, it was not possible to involve patients or the public in the design, or conduct, or reporting or dissemination plans of this study. However, clinicians who have been managing patients with COVID-19 both in health facilities and at home in Nigeria were actively engaged both in planning the study and in developing the clinical scores.
RESULTS

Description of study population
The baseline sociodemographic and clinical characteristics of the study participants in the combined datasets are presented in table 1. Overall, 15,988 children met the study eligibility criteria, 4,415 (27.6%) of whom tested positive for COVID-19 and 62.2% were male. A total of 26,441 adults met the study eligibility criteria, 9,154 (34.6%) of whom tested positive for COVID-19, 64.0% (16,819/26,441) were male, and 11.3% (2,977/26,441) were healthcare workers. Seven hundred and ninety-two elderly patients met the study eligibility criteria, 317 (40.0%) of whom tested positive for COVID-19 and 62.3% were male. Similar characteristics were observed in the derivation (online supplemental table 4) and validation (online supplemental table 5) cohorts for all age groups.

Performance of individual symptoms for predicting COVID-19 positivity
In the unadjusted model, presentation with cough, runny nose, fatigue, loss of taste, loss of smell and fever was significantly associated with higher odds of COVID-19 positivity in children (left panel of table 2). In the adjusted model, however, presentation with the following symptoms remained significantly (p<0.001) associated with higher odds of COVID-19 positivity: cough (adjusted OR (aOR) 1.32, 95% CI 1.19 to 1.47), runny nose (aOR 1.48, 95% CI 1.32 to 1.66), fatigue (aOR 1.53, 95% CI 1.21 to 1.93), loss of taste (aOR 2.26, 95% CI 1.63 to 3.12), loss of smell (aOR 4.87, 95% CI 3.66 to 6.49) or fever (aOR 1.47, 95% CI 1.33 to 1.63). Regarding predictive performance, loss of smell recorded the highest AUROC value (0.56, 95% CI 0.55 to 0.56). Specifically, cough recorded the highest sensitivity (47.7%, 95% CI 45.6% to 49.7%) and NPV (74.7%, 95% CI 73.4% to 75.9%), while loss of smell recorded the highest specificity (98.1%, 95% CI 97.7% to 98.4%) and PPV (72.9%, 95% CI 68.3% to 77.2%) (see details in online supplemental table 6).

The presentation of symptoms significantly associated with increased COVID-19 positivity was higher in adults than in children (middle panel of table 2). In the adjusted model, however, it was only the presentation with cough (aOR 1.77, 95% CI 1.64 to 1.91), runny nose (aOR 1.18, 95% CI 1.08 to 1.30), chest pain (aOR 1.50, 95% CI 1.21 to 1.85), fatigue (aOR 1.39, 95% CI 1.17 to 1.64), headache (aOR 1.24, 95% CI 1.11 to 1.39), loss of taste (aOR 2.33, 95% CI 1.83 to 2.96), loss of smell (aOR 4.18, 95% CI 3.30 to 5.28) or fever (aOR 1.58, 95% CI 1.47 to 1.71) that remained independently associated with COVID-19 positivity. Regarding predictive performance, presentation with either fever or cough appeared to be the best predictor of COVID-19 positivity (AUROC 0.57, 95% CI 0.56 to 0.58). Additionally, presenting with cough or fever recorded the highest sensitivity at 54.3% (52.9% to 55.8%) and NPV at 70.8% (69.7% to 71.8%); however, presenting with loss of smell recorded the highest specificity (98.5%; 95% CI 98.2% to 98.8%) and PPV (79.4%, 95% CI 76.0% to 82.5%) (see details in online supplemental table 7).

Only three symptoms in the elderly patients were significantly associated with COVID-19 positivity both in the unadjusted and adjusted models, with cough (aOR 1.59, 95% CI 1.04 to 2.43; p=0.033), difficulty in breathing (aOR 1.74, 95% CI 1.12 to 2.72; p=0.015) or loss of smell (aOR 7.15, 95% CI 1.44 to 35.44; p=0.016) being independently associated with COVID-19 positivity in the latter model (right panel of table 2). Regarding predictive performance, presenting with difficulty in breathing appeared to be the best predictor of COVID-19 positivity (AUROC 0.53, 0.48 to 0.58). However, presenting with cough recorded the highest sensitivity (61.4%, 95% CI 53.3% to 69.0%) and NPV (62.1%, 95% CI 54.1% to 69.6%), while loss of smell recorded the highest specificity (99.1%, 95% CI 97.0% to 99.9%) and PPV (60.0%, 95% CI 14.7% to 94.7%) (see details in online supplemental table 8).

Performance of unweighted, statistically and clinically weighted scores for predicting COVID-19 positivity
In adults, the statistically derived weighted combined score appeared to be slightly better in predicting COVID-19 positivity (AUROC 0.65), when compared with both the clinically derived weighted combined and unweighted scores with AUROC 0.63 each (table 3). A similar pattern was recorded in children and elderly populations. For individual symptom thresholds, presenting with ≥2 symptoms on the unweighted score appeared to be a better predictor of COVID-19 positivity (AUROC 0.61, 95% CI 0.60 to 0.61) than the other symptom thresholds in adults. For the statistically weighted score, presenting with ≥4 symptoms appeared to be a better predictor of COVID-19 positivity (AUROC 0.59, 95% CI 0.58 to 0.59) than the other symptom thresholds in adults. For the clinically derived weighted score, presenting with either ≥3 or ≥4 symptoms was a better predictor of COVID-19 positivity (AUROC 0.58, 95% CI 0.58 to 0.59) in adults. The detailed results including sensitivity, specificity, PPV and NPV values for each age group are available in online supplemental tables 9-11. Figure 2 shows the ROC curves comparing the predictive performance of unweighted, statistically and clinically derived weighted combined scores in children (A), adult (B) and the elderly (C) populations.

External calibration belt and test for model goodness-of-fit
The calibration belt in the produced plots and tests for the unweighted, statistically and clinically weighted scores in children, adults and elderly are presented in a supplemental figure. With the exception of unweighted and clinically weighted scores in children, all other scores showed good calibration.
Table 1  Baseline sociodemographic and clinical characteristics of all the study participants in relation to COVID-19 infection (combined dataset)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Children (≤17 years)</th>
<th>Adults (18-64 years)</th>
<th>Elderly (≥65 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCR-confirmed cases</td>
<td>PCR positive and</td>
<td>PCR-confirmed cases</td>
</tr>
<tr>
<td></td>
<td>(n=4415 (%))</td>
<td>negative cases</td>
<td>(n=9154 (%))</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=26441 (%))</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=317 (%))</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=792 (%))</td>
</tr>
<tr>
<td>Sociodemographic features</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1773 (40.16)</td>
<td>6051 (37.85)</td>
<td>9622 (36.39)</td>
</tr>
<tr>
<td>Male</td>
<td>2642 (59.84)</td>
<td>9937 (62.15)*</td>
<td>16819 (63.61)*</td>
</tr>
<tr>
<td>Geopolitical zone‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South-west</td>
<td>1302 (29.49)</td>
<td>3539 (22.14)</td>
<td>6074 (13.72)</td>
</tr>
<tr>
<td>South-south</td>
<td>1435 (32.50)</td>
<td>5689 (35.58)</td>
<td>10911 (41.27)</td>
</tr>
<tr>
<td>South-east</td>
<td>90 (2.04)</td>
<td>278 (1.74)</td>
<td>273 (2.98)</td>
</tr>
<tr>
<td>North-central</td>
<td>816 (18.48)</td>
<td>3056 (19.11)</td>
<td>4061 (15.15)</td>
</tr>
<tr>
<td>North-west</td>
<td>658 (14.90)</td>
<td>3118 (19.50)</td>
<td>4005 (15.15)</td>
</tr>
<tr>
<td>North-east</td>
<td>114 (2.58)</td>
<td>308 (1.93)*</td>
<td>610 (2.31)*</td>
</tr>
<tr>
<td>Setting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>230 (5.21)</td>
<td>887 (5.55)</td>
<td>1562 (5.91)</td>
</tr>
<tr>
<td>Urban</td>
<td>2048 (46.39)</td>
<td>6943 (43.43)</td>
<td>13088 (49.50)</td>
</tr>
<tr>
<td>Missing</td>
<td>2137 (48.40)</td>
<td>8158 (53.78)*</td>
<td>11791 (44.59)NS†</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>58 (1.31)</td>
<td>217 (1.36)</td>
<td>368 (1.39)</td>
</tr>
<tr>
<td>Nursery</td>
<td>30 (0.68)</td>
<td>143 (0.89)</td>
<td>7 (0.03)</td>
</tr>
<tr>
<td>Primary</td>
<td>130 (2.94)</td>
<td>520 (3.25)</td>
<td>322 (1.22)</td>
</tr>
<tr>
<td>Secondary</td>
<td>385 (8.72)</td>
<td>1379 (8.63)</td>
<td>2262 (8.55)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>1410 (31.94)</td>
<td>4066 (25.43)</td>
<td>8145 (30.80)</td>
</tr>
<tr>
<td>Other</td>
<td>179 (4.05)</td>
<td>1065 (6.66)</td>
<td>688 (2.60)</td>
</tr>
<tr>
<td>Missing</td>
<td>2223 (50.35)</td>
<td>8598 (53.78)*</td>
<td>14649 (55.40)*</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student/pupil</td>
<td>599 (13.57)</td>
<td>2809 (17.57)</td>
<td>1516 (5.73)</td>
</tr>
<tr>
<td>Child/housewife</td>
<td>133 (3.01)</td>
<td>489 (3.06)</td>
<td>500 (1.89)</td>
</tr>
<tr>
<td>Business/trading</td>
<td>217 (4.92)</td>
<td>765 (4.78)</td>
<td>2043 (7.73)</td>
</tr>
<tr>
<td>Transporter</td>
<td>11 (0.25)</td>
<td>80 (0.50)</td>
<td>174 (0.66)</td>
</tr>
<tr>
<td>Healthcare worker</td>
<td>601 (13.61)</td>
<td>1953 (12.22)</td>
<td>2977 (11.26)</td>
</tr>
<tr>
<td>Laboratory</td>
<td>15 (0.34)</td>
<td>40 (0.25)</td>
<td>63 (0.24)</td>
</tr>
<tr>
<td>Farmer</td>
<td>36 (0.82)</td>
<td>192 (1.20)</td>
<td>589 (2.23)</td>
</tr>
<tr>
<td>Animal-related worker</td>
<td>8 (0.18)</td>
<td>28 (0.18)</td>
<td>80 (0.30)</td>
</tr>
<tr>
<td>Religious/traditional leader</td>
<td>5 (0.11)</td>
<td>24 (0.15)</td>
<td>129 (0.49)</td>
</tr>
<tr>
<td>Other</td>
<td>1633 (36.99)</td>
<td>5126 (32.06)</td>
<td>11408 (43.15)</td>
</tr>
<tr>
<td>Missing</td>
<td>1157 (26.21)</td>
<td>4482 (28.03)*</td>
<td>6962 (26.33)*</td>
</tr>
<tr>
<td>Obvious (visible to healthcare workers on sight)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chills/sweat</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 1 Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Children (≤17 years)</th>
<th>Adults (18–64 years)</th>
<th>Elderly (≥65 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCR-confirmed cases (n=4415 (%))</td>
<td>PCR positive and negative cases (n=15988 (%))</td>
<td>PCR-confirmed cases (n=9154 (%))</td>
</tr>
<tr>
<td>No</td>
<td>4355 (98.64)</td>
<td>15764 (98.60)</td>
<td>9009 (98.42)</td>
</tr>
<tr>
<td>Yes</td>
<td>60 (1.36)</td>
<td>224 (1.40)†</td>
<td>145 (1.58)</td>
</tr>
<tr>
<td>Cough</td>
<td>No</td>
<td>2345 (53.11)</td>
<td>9393 (58.75)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>2070 (46.89)</td>
<td>6595 (41.25)*</td>
</tr>
<tr>
<td>Breathing difficulty</td>
<td>No</td>
<td>3936 (89.15)</td>
<td>14283 (89.34)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>479 (10.85)</td>
<td>1705 (10.66)NS†</td>
</tr>
<tr>
<td>Rapid breathing</td>
<td>No</td>
<td>4345 (98.41)</td>
<td>15783 (98.72)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>70 (1.59)</td>
<td>205 (1.28)§</td>
</tr>
<tr>
<td>Runny nose</td>
<td>No</td>
<td>3057 (69.24)</td>
<td>11835 (74.02)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>1358 (30.76)</td>
<td>4153 (25.98)*</td>
</tr>
<tr>
<td>Elicitable (can be found out by asking questions of patients/relatives)</td>
<td>Abdominal pain/diarrhoea</td>
<td>No 4142 (93.82)</td>
<td>14939 (93.44)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>273 (6.18)</td>
<td>1049 (6.56)NS†</td>
</tr>
<tr>
<td>Gastrointestinal symptoms</td>
<td>No</td>
<td>3972 (89.97)</td>
<td>14375 (89.91)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>443 (10.03)</td>
<td>1613 (10.09)NS†</td>
</tr>
<tr>
<td>Chest pain</td>
<td>No</td>
<td>4286 (97.08)</td>
<td>15577 (97.43)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>129 (2.92)</td>
<td>411 (2.57)NS†</td>
</tr>
<tr>
<td>Fatigue</td>
<td>No</td>
<td>4143 (93.84)</td>
<td>15229 (95.25)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>272 (6.16)</td>
<td>759 (4.75)*</td>
</tr>
<tr>
<td>Headache</td>
<td>No</td>
<td>3753 (85.01)</td>
<td>13712 (85.76)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>662 (14.99)</td>
<td>2276 (14.24)NS†</td>
</tr>
<tr>
<td>Musculoskeletal pain</td>
<td>No</td>
<td>4299 (97.37)</td>
<td>15595 (97.54)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>116 (2.63)</td>
<td>393 (2.46)NS</td>
</tr>
<tr>
<td>Sore throat</td>
<td>No</td>
<td>3306 (74.88)</td>
<td>11968 (74.86)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>1109 (25.12)</td>
<td>4020 (25.14)NS†</td>
</tr>
<tr>
<td>Loss of taste</td>
<td>No</td>
<td>3925 (88.90)</td>
<td>15283 (95.59)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>490 (11.10)</td>
<td>705 (4.41)*</td>
</tr>
<tr>
<td>Loss of smell</td>
<td>No</td>
<td>3815 (86.41)</td>
<td>15160 (94.82)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>600 (13.59)</td>
<td>828 (5.18)§</td>
</tr>
</tbody>
</table>

Continued
Strengths and limitations of this study

To the best of our knowledge, this is the first study to have developed and validated symptom prediction scores with a view to aiding prompt recognition of COVID-19 by frontline healthcare workers in the Nigerian context and possibly in sub-Saharan Africa at large. Despite the limited accuracy of the developed prediction tool, the findings are very important for a country with limited capacity for molecular diagnosis of COVID-19, as it provides the evidential basis for advocacy for more investments in molecular diagnostics by policy makers in Nigeria. We also adopted a transparent methodology and adhered to the TRIPOD reporting statement, hence minimising the vagueness often associated with the reporting of studies on the predictive performance of diagnostic models for COVID-19. The methodology taken to the derivation of weighted scores is also a strength of this study. In accordance with the preferred approach for building prediction models, the participatory approach taken to deriving the clinically weighted score can enhance the study relevance in the medical community. The use of beta regression coefficients as opposed to ORs in deriving the statistical weighted scores has the advantage of being less prone to bias by small to moderate sample size. A common limitation of many COVID-19 diagnostic models is bias due to overfitting of the models on data that are not representative of the target population. By using SORMAS database (hosts data from all over the country) for both derivation and validation in the present study, our findings are considerably generalisable to the COVID-19 situation in Nigeria and less prone to overestimation of COVID-19 risk among individuals tested.

This study, however, has some limitations that warrant discussion. First, being an analysis of secondary data based on practical recording of routine clinical assessments, the fundamental assumption is that the data recorded on practical recording of routine clinical assessments, the fundamental assumption is that the data recorded on clinical laboratory data could potentially minimise differences in comparisons, in which case observed differences are likely to be in the direction of the null hypothesis. Second, the approach of splitting the dataset for both derivation and validation of symptom scores may have lowered the precision of estimated effect (wider 95% CIs) and potentially underestimated prediction performance due to loss of power. Moreover, evidence supporting the 10 events per outcome rule of thumb has been found by van Smeden et al to be weak. Third, the study lacked detailed clinical laboratory data, such as record for albumin or albumin/globin, direct bilirubin values and red cell distribution width, which have been found to be significant variables in COVID-19 diagnostic models. Practically, however, the time and technical requirements for testing these laboratory data could limit their clinical utility.

Interpretation and implications of findings

Based on the systematic and critical review of diagnostic scores by Wynants et al, the performance of a diagnostic

<table>
<thead>
<tr>
<th>Variable</th>
<th>Children (≤17 years)</th>
<th>Adults (18–64 years)</th>
<th>Elderly (≥65 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fever</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>2350 (53.23)</td>
<td>9430 (58.98)</td>
<td>15434 (58.37)</td>
</tr>
<tr>
<td>Yes</td>
<td>2065 (46.77)</td>
<td>6558 (41.02)</td>
<td>11007 (41.63)</td>
</tr>
</tbody>
</table>

*In adults, all the symptom scores showed similar performance, with the statistically weighted score (AUROC 0.65) slightly showing better performance than the unweighted (AUROC 0.63) and clinically derived weighted (AUROC 0.64) scores. Similar results were found in children and elderly patients. Overall, none of the symptom scores had good enough discrimination to use in practice.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Children (≤17 years)</th>
<th>Adults (18–64 years)</th>
<th>Elderly (≥65 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR-confirmed cases (n=4415 (%)</td>
<td>PCR positive and negative cases (n=15998 (%)</td>
<td>PCR-confirmed cases (n=9154 (%)</td>
<td>PCR positive and negative cases (n=26441 (%)</td>
</tr>
</tbody>
</table>

| Fever    |                      |                      |                     |
| No       | 2350 (53.23)        | 9430 (58.98)         | 15434 (58.37)       |
| Yes      | 2065 (46.77)        | 6558 (41.02)         | 11007 (41.63)       |

*p<0.001

*p>0.05 or not significant (NS).

*p<0.05 or not significant (NS).
Table 2  Associations and predictive performance of individual clinical characteristics of the study participants in relation to COVID-19 positivity (based on derivation datasets)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Children (n=7911)</th>
<th>Adults (n=13167)</th>
<th>Elderly (n=399)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted OR (95% CI)</td>
<td>Adjusted OR (95% CI)*</td>
<td>AUROC (95% CI)†</td>
</tr>
<tr>
<td>Obvious (visible to healthcare workers on sight)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chills/sweat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.15 (0.77 to 1.70)NS</td>
<td>1.14 (0.85 to 1.54)NS</td>
<td>0.30 (0.03 to 2.57)NS</td>
</tr>
<tr>
<td>Cough</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.54</td>
</tr>
<tr>
<td>Yes</td>
<td>1.39 (1.26 to 1.53)‡</td>
<td>1.32 (1.19 to 1.47)‡ (0.53 to 0.55)</td>
<td>1.85 (1.72 to 1.99)‡</td>
</tr>
<tr>
<td>Breathing difficulty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.02 (0.87 to 1.19)NS</td>
<td>1.23 (1.11 to 1.35)‡</td>
<td>1.10 (0.99 to 1.22)NS</td>
</tr>
<tr>
<td>Rapid breathing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.11 (0.73 to 1.68)NS</td>
<td>1.58 (1.15 to 2.16)§</td>
<td>1.24 (0.89 to 1.73)NS</td>
</tr>
<tr>
<td>Runny nose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.52</td>
</tr>
<tr>
<td>Yes</td>
<td>1.53 (1.37 to 1.70)‡</td>
<td>1.48 (1.32 to 1.66)‡ (0.51 to 0.53)</td>
<td>1.30 (1.20 to 1.42)‡</td>
</tr>
<tr>
<td>Elicitable (can be found out from patients’ responses or by asking questions of relatives)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal pain/ diarrhoea</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>0.87 (0.71 to 1.07)NS</td>
<td>1.12 (0.98 to 1.29)NS</td>
<td>0.62 (0.26 to 1.44)NS</td>
</tr>
<tr>
<td>GIT symptoms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>0.96 (0.82 to 1.14)NS</td>
<td>1.17 (1.04 to 1.32)‡</td>
<td>1.11 (0.98 to 1.26)NS</td>
</tr>
<tr>
<td>Chest pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.33 (0.98 to 1.79)NS</td>
<td>1.75 (1.44 to 2.14)‡</td>
<td>1.50 (1.21 to 1.85)‡ (0.50–0.51)</td>
</tr>
<tr>
<td>Fatigue/weakness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.51</td>
</tr>
<tr>
<td>Yes</td>
<td>1.29 (1.03 to 1.62)§</td>
<td>1.53 (1.21 to 1.93)‡ (0.51–0.52)</td>
<td>1.25 (1.07 to 1.46)§</td>
</tr>
<tr>
<td>Headache</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.52</td>
</tr>
<tr>
<td>Yes</td>
<td>1.12 (0.98 to 1.29)NS</td>
<td>1.44 (1.29 to 1.60)‡</td>
<td>1.24 (1.11 to 1.39)‡ (0.51 to 0.52)</td>
</tr>
<tr>
<td>Musculoskeletal pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.12 (0.98 to 1.29)NS</td>
<td>1.44 (1.29 to 1.60)‡</td>
<td>1.24 (1.11 to 1.39)‡ (0.51 to 0.52)</td>
</tr>
</tbody>
</table>

Continued
### Table 2 Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Children (n=7911)</th>
<th>Adults (n=13167)</th>
<th>Elderly (n=399)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted OR (95% CI)</td>
<td>Adjusted OR (95% CI)*</td>
<td>AUROC (95% CI)†</td>
</tr>
<tr>
<td>N0</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.09 (0.80 to 1.48) NS</td>
<td>1.04 (0.81 to 1.32) NS</td>
<td>1.80 (0.59 to 5.45) NS</td>
</tr>
<tr>
<td>Sore throat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>1.02 (0.91 to 1.14) NS</td>
<td>0.83 (0.76 to 0.90) §</td>
<td>0.82 (0.50 to 1.35) NS</td>
</tr>
<tr>
<td>Loss of taste</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.54</td>
</tr>
<tr>
<td>Yes</td>
<td>7.07 (5.59 to 8.95)</td>
<td>2.26 (1.63 to 3.12) †</td>
<td>(0.54 to 0.55)</td>
</tr>
<tr>
<td>Loss of smell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.56</td>
</tr>
<tr>
<td>Yes</td>
<td>7.85 (6.30 to 9.77) †</td>
<td>4.67 (3.66 to 6.49) †</td>
<td>(0.55 to 0.56)</td>
</tr>
<tr>
<td>Measurable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fever</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.00</td>
<td>1.00</td>
<td>0.54</td>
</tr>
<tr>
<td>Yes</td>
<td>1.43 (1.29 to 1.58) †</td>
<td>1.47 (1.33 to 1.63) †</td>
<td>(0.52 to 0.55)</td>
</tr>
</tbody>
</table>

Significant results in the adjusted model are in bold fonts.

*The fully adjusted model includes all statistically significant variables from the unadjusted model.
†Area under receiver operating characteristic curve; figures are for variables that were statistically significant in the adjusted model only; the best predictor AUROC value is highlighted in bold.
‡Wald’s p-values: <0.001.
§Wald’s p-values: <0.05.
NS, not significant (p>0.05).
model is influenced by its composition, with higher number of clinical and laboratory parameters in a model indicating better predictive performance. For instance, studies in China, Brazil, Italy, The Netherlands and France with several clinical and laboratory parameters recorded excellent discriminatory performance, although with substantial evidence of bias and limited clinical utility. Conversely, a prediction model (containing fewer number of symptoms, heart rate, systolic and diastolic blood pressure) developed by Sun et al in Singapore had a poor discriminatory capacity (C statistic: 0.65; 0.57–0.73). There is evidence to further suggest that the discriminatory accuracy of a prediction model, particularly its sensitivity, can be enhanced by including certain variables including a combination of loss of smell or taste and fever. In the absence of comparison study from a sub-Saharan African country, it is difficult to fully explain the variation in the findings from the present study and elsewhere. Thus, a follow-up study using both clinical and laboratory parameters in a Nigerian setting or in sub-Saharan Africa (with similar healthcare system and demographic structure) is recommended.

### Table 3: Predictive performance of unweighted, statistically and clinically weighted score thresholds for predicting COVID-19 positivity in children, adults and elderly

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Score</th>
<th>Unweighted score</th>
<th>Statically weighted score</th>
<th>Clinically weighted score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Children (&lt;17 years), n=8077</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined score</td>
<td>0.6064</td>
<td>0.6177</td>
<td>0.5915</td>
<td></td>
</tr>
<tr>
<td>Symptom threshold</td>
<td>≥1</td>
<td>0.55 (0.54 to 0.56)</td>
<td>0.55 (0.54 to 0.56)</td>
<td>0.55 (0.54 to 0.56)</td>
</tr>
<tr>
<td></td>
<td>≥2</td>
<td>0.58 (0.57 to 0.59)</td>
<td>0.55 (0.54 to 0.56)</td>
<td>0.54 (0.53 to 0.55)</td>
</tr>
<tr>
<td></td>
<td>≥3</td>
<td>0.55 (0.54 to 0.56)</td>
<td>0.55 (0.54 to 0.57)</td>
<td>0.56 (0.55 to 0.57)</td>
</tr>
<tr>
<td></td>
<td>≥4</td>
<td>0.52 (0.52 to 0.53)</td>
<td>0.59 (0.58 to 0.60)</td>
<td>0.57 (0.56 to 0.58)</td>
</tr>
<tr>
<td><strong>Adults (17–64 years), n=13274</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined score</td>
<td>0.6333</td>
<td>0.6475</td>
<td>0.6389</td>
<td></td>
</tr>
<tr>
<td>Symptom threshold</td>
<td>≥1</td>
<td>0.56 (0.55 to 0.56)</td>
<td>0.56 (0.55 to 0.56)</td>
<td>0.56 (0.55 to 0.56)</td>
</tr>
<tr>
<td></td>
<td>≥2</td>
<td>0.61 (0.60 to 0.61)</td>
<td>0.58 (0.57 to 0.58)</td>
<td>0.56 (0.55 to 0.56)</td>
</tr>
<tr>
<td></td>
<td>≥3</td>
<td>0.56 (0.55 to 0.57)</td>
<td>0.58 (0.57 to 0.59)</td>
<td>0.58 (0.58 to 0.59)</td>
</tr>
<tr>
<td></td>
<td>≥4</td>
<td>0.53 (0.53 to 0.54)</td>
<td>0.59 (0.58 to 0.59)</td>
<td>0.58 (0.58 to 0.59)</td>
</tr>
<tr>
<td><strong>Elderly (≥65 years), n=393</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined score</td>
<td>0.5413</td>
<td>0.5453</td>
<td>0.5426</td>
<td></td>
</tr>
<tr>
<td>Symptom threshold</td>
<td>≥1</td>
<td>0.52 (0.48 to 0.57)</td>
<td>0.52 (0.48 to 0.57)</td>
<td>0.52 (0.48 to 0.57)</td>
</tr>
<tr>
<td></td>
<td>≥2</td>
<td>0.53 (0.49 to 0.57)</td>
<td>0.52 (0.48 to 0.57)</td>
<td>0.52 (0.48 to 0.57)</td>
</tr>
</tbody>
</table>

The best predictive AUROC values for each age group are highlighted in bold. AUROC, area under the receiver operating characteristic curve.
with a high proportion of asymptomatic COVID-19 cases and testing system that allows persons who are concerned about their COVID-19 risk to be tested. Thus, to minimise missed diagnoses and overburdening of the healthcare system, with attendant psychological effects on health personnel, there is a need for more economic investments on molecular testing across Nigeria.

Loss of smell recorded the highest specificity with regard to COVID-19 positivity for the three age groups: 98.1% in children, 98.5% in adults and 99.1% in the elderly. However, unlike the present study which explored the predictive capacity of loss of smell and taste separately, a combination of both symptoms has been shown to be more predictive. Thus, the potential use of both loss of smell and taste to differentiate COVID-19 from endemic febrile and respiratory illnesses in Nigeria, such as malaria and pneumonia, with overlapping symptoms warrant further study. Additionally, possibility of using both loss of smell and taste as early indicators of emerging COVID-19 wave or a surge in Nigeria would be useful in improving COVID-19 response, such as allocation of already limited...
testing resources, risk communication and aid decision-making concerning lockdowns and quarantines. The poor predictive capacity of cough or fever alone in the present study is congruent with that in a meta-analysis. Clinical validity (characterised by sensitivity, specificity and AUROC values) is an important criterion for assessing a clinical prediction tool as it is—the ability of the prediction tool to distinguish between who has an outcome (in this case SARS-CoV-2 infection) and who does not. The clinical validity of all our prediction scores was generally poor but appeared to be dependent on the number of symptoms. For instance, in our study, the unweighted and weighted (both statistical and clinical) predictive scores presenting with fewer number of symptoms were more sensitive compared with many symptoms in children and adults; it was, however, the opposite relative to specificity given ≥4 symptoms recorded higher specificity values than lower symptom thresholds. The poor sensitivity of many symptoms could potentially be attributable to a high proportion of false negatives, suggesting that some symptoms have limited validity for COVID-19 in children and adults. However, similarity in the predictive performance of various symptom thresholds on the two weighted scores in elderly suggests that weighting has less predictive value for this group of population. The high specificity of more symptoms could be indicative of low proportion of false positives, underlining the need to accurately assess symptoms. In practice, there is a trade-off between sensitivity and specificity such that when the consequences of having a false positive test is very serious, specificity is prioritised over sensitivity and vice versa. This is the case for the various symptom thresholds on the unweighted scale where specificity is higher than sensitivity. A higher specificity over sensitivity is of practical relevance when the political implication of refusing to test someone with suspected COVID-19 is considered, although higher sensitivity over specificity might be given preference in the early phases of a pandemic before surge capacity is reached.

Given the rapid increase in community transmission of COVID-19 cases and deleterious impacts of instituting another lockdown (partial or complete), large-scale surveillance for capturing the epidemiological trend of COVID-19 in Nigeria is crucial. However, Nigeria has limited SARS-CoV-2 testing capacity with an average turnaround of 2 days, making syndromic surveillance (syndromic monitoring) a viable complementary surveillance system. As such, our findings would be relevant in informing the design of such a surveillance system, which has been demonstrated in Japan and in the USA to be useful in improving the understanding of COVID-19 epidemiology (often in real time), assessing the effectiveness of public health interventions and enhancing preparedness for the emergence of COVID-19 wave or a surge. For instance, an evaluation of a syndromic surveillance system in the USA found new taste/smell loss to be highly correlated with a range of COVID-19 outcomes, highlighting their usefulness in supporting the surveillance system as an early warning system for COVID-19 prevention and control. However, the feasibility (eg, considering selection bias and recall bias) and acceptability of a syndromic surveillance system first need to be ascertained given the large proportion of asymptomatic COVID-19 cases at diagnosis in Nigeria. PPVs across the various prediction thresholds, especially for the weighted scales, were generally low despite increasing proportionately with the thresholds. This could be attributable, in part, to the general mildness of the pandemic with resultant low incidence of mortality in Nigeria. For instance, 66% of the 12 289 confirmed COVID-19 cases in Nigeria between 27 February and 6 June 2020 were asymptomatic at diagnosis, with an overall cumulative incidence and case fatality rate of 5.6 per 100 000 population and 2.8%, respectively—these figures were substantially lower than those from European countries during the same period. As such, our predictive tools could perform differently during a more severe COVID-19 outbreak in Nigeria.

CONCLUSION

This study has investigated the possibility of using symptoms to predict COVID-19 positivity in Nigeria and found the predictive capacity of various symptom scores to be poor overall. However, the findings have the potential to serve as an advocacy tool for more investments in resources for capacity strengthening of molecular testing for COVID-19 in Nigeria, which is crucial for improving both clinical case management and surveillance.

Author affiliations
1Nigeria Centre for Disease Control, Abuja, Nigeria
2Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden
3Department of Epidemiology and Public Health, University of Nottingham, Nottingham, UK
4World Health Organization, Abuja, Nigeria
5University of Abuya Teaching Hospital, Gwagwalada, Nigeria
6Lagos State Ministry of Health, Ikeja, Nigeria
7Department of Epidemiology and Community Health, University of Ilorin Teaching Hospital, Ilorin, Nigeria
8Department of Internal Medicine, National Hospital, Abuja, Nigeria
9National Agency for the Control of AIDS, Abuja, Nigeria
10Federal Medical Centre Asaba, Asaba, Nigeria
11Federal Teaching Hospital Gombe, Gombe, Nigeria
12Kaduna State Infectious Disease Control Center Community Medicine, Kaduna, Nigeria
13Mainland Hospital Yaba, Lagos, Nigeria
14Ministry of Health, Lagos, Nigeria
15Lagos State Government Ministry of Health, Ikeja, Nigeria
16Department of Community Medicine, Ahmadu Bello University, Zaria, Nigeria
17Department of Community Medicine, Nnamdi Azikiwe University Teaching Hospital, Nnewi, Nigeria
18Department of Medical Education, Usmanu Danfodio University, Sokoto, Nigeria
19Department of Community Medicine, University College Hospital, Ibadan, Nigeria
20School of Public Health, University of the Western Cape, Bellville, South Africa

Twitter Kelly Osezele Elimian @KElimian, Chinye Lucia Ochu @chinye_ochu, Carina King @CarinaTKing and Ehimario Igumbor @Ehimario

Acknowledgements We acknowledge the contribution of all the COVID-19 treatment centres and laboratories within the NCDC Laboratory Networks of both epidemiological and laboratory data to SORMAS on a regular basis. We express
our profound gratitude to the NCDC SORMAS team for collating and maintaining a quality dataset utilised for this study. Last but not least, we thank Professor Jo Leonard-Bee of the University of Nottingham for her input on statistical analyses.

**Contributors**

KOE contributed to the conceptualisation, literature search, data management, data analysis, writing (original draft, review and editing) and formatting of the document. OA contributed to the conceptualisation, writing (review and editing), formatting of manuscript and funding acquisition. JG contributed to the conceptualisation, data management, data analysis, writing (original draft, review and editing) and formatting of the document. PM contributed to the writing (original draft, review and editing), literature search and formatting of the manuscript. CLO contributed to the writing (review and editing) and formatting of the manuscript. CQ contributed to the writing (review and editing) and formatting of the manuscript. CD-N contributed to the writing (review and editing) and formatting of the manuscript. OI contributed to the writing (review and editing) and formatting of the manuscript. CD-N contributed to the writing (review and editing) and formatting of the manuscript. JTA contributed to the writing (review and editing) and formatting of the manuscript. JFA contributed to the writing (review and editing) and formatting of the manuscript. AB contributed to the writing (review and editing) and formatting of the manuscript. BS contributed to the writing (review and editing) and formatting of the manuscript. OWA contributed to the writing (review and editing) and formatting of the manuscript. WN contributed to the writing (review and editing) and formatting of the manuscript. IA contributed to data acquisition and curation, and writing (review and editing) and formatting of the manuscript. BM contributed to the writing (review and editing) and formatting of the manuscript. JS contributed to the writing (review and editing) and formatting of the manuscript. AO contributed to the writing (review and editing) and formatting of the manuscript. CO contributed to the writing (review and editing) and formatting of the manuscript. BE contributed to the writing (review and editing) and formatting of the manuscript. WJ contributed to the writing (review and editing) and formatting of the manuscript. AB contributed to the writing (review and editing) and formatting of the manuscript. BM contributed to the writing (review and editing) and formatting of the manuscript. BS contributed to the writing (review and editing) and formatting of the manuscript. A Jinadu contributed to the writing (review and editing) and formatting of the manuscript. KF contributed to the writing (review and editing) and formatting of the manuscript. C Ike contributed to the writing (review and editing) and formatting of the manuscript. SB contributed to the writing (review and editing) and formatting of the manuscript. MCA contributed to the writing (review and editing) and formatting of the manuscript. NM contributed to supervision, writing (review), funding acquisition, project administration and resources. JO contributed to supervision, writing (review), funding acquisition, project administration and resources. EI contributed to supervision, writing (review), funding acquisition, project administration and resources. TA contributed to the writing (review and editing), literature search and formatting of the manuscript. C Chekwuzo contributed to conceptualisation, supervision, writing (review and editing), funding acquisition, project administration and resources. All authors invest significant contributions and approved the final draft.

**Funding**
The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** The Nigeria National Health Research Ethics Committee approved the use of the SORMAS database for this study (reference number: NHREC/01/01/2007-22/06/2020).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Kelly Osezie Elimian http://orcid.org/0000-0003-2729-1613
Chinwe Lucia Ochu http://orcid.org/0000-0002-0830-7332
Giulia Gaudenzi http://orcid.org/0000-0003-4922-6965
Oluwatosin Wuraola Akande http://orcid.org/0000-0001-6906-895X

**REFERENCES**


