Supplemental Materials: Consistent Comparison of Symptom-based Methods for COVID-19 Infection Detection

A. Materials and Methods

A.1. Estimation Detection Methods

Detection methods were grouped into three types. More precisely, the rules-based methods give weights to the symptoms using some criteria (physician, from some organization, etc.) to identify positive cases. In addition, machine-learning-based techniques build tree-based classifiers to detect infected people from datasets containing information on symptoms. Finally, the regression-based approaches build prediction models using logistic regression techniques.

A.1.1. Rule-based methods

The rule-based methods used in this works are:

- Smith

  This work builds a clinical prediction rule to identify COVID-19 active cases in symptomatic individuals [10]. To this end, this method implemented a multivariable logistic regression analysis to identify the independent predictors of COVID-19 active cases. Specifically, the Smith method selects a few symptoms associated with positive cases and assigns them different coefficients: loss of smell/loss of taste (2), fever and cough (1), and chest pain (-1). The chest pain variable has a negative score because this symptom being likely caused by another virus. The dataset used was obtained from a standardized clinical questionnaire that was administered to patients before applying the RT-PCR test. Moreover, the performance of the Smith method was tested using a dataset with 120 SARS-CoV-2-positive cases and 120 SARS-CoV-2-negative cases for training, and 40 cases for validation of the classification model.

- Centers for Disease Control and Prevention (CDC)

  In August 2020, the Centers for Disease Control and Prevention approved the COVID-like illness (CDC) metric [20]. This metric declares a COVID-19 positive case if an individual presents at least two of the following symptoms: fever, chills, rigors, myalgia (muscle aches and pain), headache, sore throat. This metric also identifies an active case if the individual has at least one of the following symptoms: cough, shortness of breath, difficulty breathing, loss of smell, or loss of taste. Notice that the UMD-CTIS survey does not register myalgia and rigors. Hence, we estimate the CDC metric without those symptoms.

- WHO

  Under the context of continuous monitoring of the COVID-19 pandemic, the World Health Organization (WHO) published a COVID-like illness[1] metric that declares a potential active case when individuals report the following symptoms: fever, cough, and fatigue [16]. This metric was used in [17] to estimate COVID-19 active cases in various coutries such as Spain, Peru, Ecuador, UK, Greece, and India considering the information extracted from the UMD-CTIS.

- Akinbami

  The Akinbami method uses a combination of three symptoms to declare a COVID-19 positive case [7]. To this end, the study built three classification models depending on the combination of the symptoms: 1) **Akinbami_1** which uses fever, shortness of breath, and chills, 2) **Akinbami_2** which uses fever, shortness of breath, and anosmia/ageusia, and 3) **Akinbami_3** which uses fever, shortness of breath, and headache. The dataset was provided by a serologic survey collected in Detroit and New York from May 17 to July 2, 2020. The extracted dataset contains 40,938 tested individuals of which 6,631 are positive.

- Salomon

  This method defines a metric also referred to as COVID-like illness (CLI)[2] different from those specified in [20] and [17]. Specifically, this metric identifies a positive case if the participant reports fever, cough, or shortness of breath/difficulty breathing [21, 18, 17]. This method was evaluated on datasets obtained from the CMU-CTIS survey that was collected in the United States from April 2020 to April 2021 by the Delphi Group at the Carnegie Mellon University (CMU) and the CTIS in partnership with Facebook. Like the UMD-CTIS project, the CMU-CTIS survey also obtained information about individual features (such as age groups, gender, testing, and vaccination) and COVID-19 symptoms.

A.1.2. Machine Learning methods

The machine learning techniques usedn in this work are:

- Astley

  This study focused on building COVID-19 diagnostic models based on machine learning techniques [15]. More precisely, this approach selected the Light Gradient Boosting Machine (LightGBM) engine to build the COVID-19 diagnostic models considering two individual features (age groups and gender) and twelve symptoms. Furthermore, this method extracted the datasets from the UMD-CTIS collected in 114 countries from April to December 2020. Note that the training set used to build the classification models for each country is a subset of rows derived from questionnaires reporting a laboratory test.

- Zoabi

  The Zoabi method considers eight features: gender, age ($\geq$ 60), cough, fever, sore throat, shortness of

[1] Aka UMD CLI WHO in [17]

[2] Aka CLI in [21, 18] and UMD CLI in [17]

breath, headache, and known contact with an individual COVID-19 confirmed positive case. It builds a classification model based on a machine learning approach [5]. In essence, this method builds a gradient-boosting classification model with decision-tree base-learners. This approach trained and evaluated the classification model from data released by the Israeli Ministry of Health. This dataset contains information on individuals with RT-PCR tests. Specifically, the training set consists of 51,831 individuals of whom 4,769 are confirmed positive cases. On the other hand, the test set consists of 47,401 individuals of whom 3,624 are positive cases. In this case, it is worth noting that UMD-CTIS data ranges of ages does not have a boundary at 60. The boundary is either at 55 or 65. Hence, we have created 2 different models, one for each range of age labeled **Zoabi_55** and **Zoabi_65**, to go around this difference in the data.

A.1.3. Regression-based methods

The regression-based methods used in this work are:

- Menni

  This method performs logistic regressions to build the optimal COVID-19 classification model for a set of individual features such as age, gender, loss of smell and taste, cough, fatigue, and loss of appetite [9]. The building and evaluation of the classification model used a dataset extracted from a symptom tracker based on a smartphone app launched in the United Kingdom and the United States in March 2020. Specifically, this method was evaluated over responses from 2, 618, 862 participants voluntarily recording their symptoms. In this study, the best classification model according to the Akaike information criterion (AIC) was described as

$$\begin{aligned} x &= -1.32 - (0.01 \times \text{age}) + (0.44 \times \text{gender}) \\ &\quad +(1.75 \times \text{loss of smell and taste}) \\ &\quad +(0.31 \times \text{cough}) + (0.49 \times \text{fatigue}) \\ &\quad +(0.39 \times \text{skipped meals}), \end{aligned}$$

  where symptoms represent binary variables. More precisely, every feature coded as 1 reports the presence of the symptom, while each variable coded as 0 indicates its absence. The gender variable also stands for a binary variable, the one-value indicates a male, and the zero-value represents a female. Afterward, this method identified a COVID-19 active case whether $\frac{e^x}{1+e^x} \geq 0.5$. It is worth noting that UMD-CTIS data did not register the skipped meal variable. Therefore, we modified the Menni method by computing the x score with the skipped meals variable fixed to zero. This approach is labeled as **Menni_1**. Furthermore, we followed the procedure reported in [9] to build the logistic regression model from individual features available in our dataset (**Menni_2**). In other words, we built a logistic regression model that considers the features: age, gender, loss of smell and taste, cough, and fatigue.

- Roland

  This study performs logistic regression analysis to build a classification model based on five symptoms: loss of taste and smell, body aches, fever or chills, shortness of breath, and sore throat [11]. This method uses a dataset extracted from an anonymous electronic survey publicized with ads on social networks (Facebook, Twitter, Reddit, and Nextdoor) from March 31 to April 10, 2020. Specifically, the Roland method was evaluated from a dataset provided by 620 participants of whom 339 reported COVID test outcomes. This work built a stepwise logistic model whose training set was obtained by randomly extracting 25% of the rows belonging to the COVID-tested individuals. The remaining rows were used to examine the performance of the classification model.

- Mika

  This method is similar to the Roland method. In other words, the Mika method fits a logistic regression model with the following symptoms: fever $> 38°C$, cough, loss of taste and smell, and gastro-intestinal (GI) [13]. The data set consisted of 3114 participants of which 778 were tested positive. The UMD-CTIS survey does not have a question on GI symptoms. Therefore, we use the answer for the presence of nausea instead, as it is the closest related symptom.

- Shoer

  This research builds two models based on logistic regression analysis to estimate the probability of individuals testing positive for SARS-CoV-2 [14]. In particular, this study obtained the datasets from two surveys launched in Israel in 2020. On one hand, the online survey registered various individual features such as age, gender, prior medical conditions, and self-reported symptoms. On the other hand, the shortened survey captured the information by means of an interactive voice response (IVR) platform. Specifically, the IVR version collected information on variables such as age group, prior medical conditions, general feeling, and a shortened list of symptoms. To generate the first model, an integrated dataset is constructed from both the features collected by the online version and the reduced set of attributes acquired by the IVR version. The second model used the information provided by the online version only. The UMD-CTIS survey does not have questions on prior medical conditions and general feeling, and therefore, we do not include them in the model.

- Bhattacharya

Table SM1
F_1 score and its 95% confidence interval for the selected countries for 2020, in %.

Method	Brazil	Canada	Israel	Japan	Turkey	South Africa
Menni_1	65.56 (65.48 - 65.64)	54.33 (53.66 - 54.99)	59.76 (59.16 - 60.36)	46.33 (45.33 - 47.33)	63.93 (63.68 - 64.17)	61.39 (61.07 - 61.70)
Menni_2	71.13 (71.01 - 71.24)	49.33(48.77 - 49.88)	57.50 (57.04 - 57.97)	39.91 (39.27 - 40.54)	67.41 (67.21 - 67.60)	66.36 (66.10 - 66.62)
Roland	69.38 (69.30 - 69.46)	51.44 (50.86 - 52.02)	61.93 (61.46 - 62.41)	40.68 (39.98 - 41.39)	67.06 (66.87 - 67.26)	67.32 (67.05 - 67.58)
Smith	71.11 (71.05 - 71.18)	53.43 (52.85 - 54.01)	62.47 (61.98 - 62.97)	45.12 (44.42 - 45.82)	67.30 (67.11 - 67.49)	62.06 (61.80 - 62.32)
Zoabi_55	70.71 (70.65 - 70.77)	32.96 (32.37 - 33.54)	47.76 (47.32 - 48.20)	29.95 (29.29 - 30.60)	57.86 (57.69 - 58.03)	59.05 (58.80 - 59.31)
Zoabi_65	70.73 (70.67 - 70.79)	32.86 (32.28 - 33.44)	47.79 (47.36 - 48.23)	29.91 (29.27 - 30.55)	57.72 (57.55 - 57.88)	59.00 (58.74 - 59.25)
CDC	73.42 (73.36 - 73.48)	23.43 (23.14 - 23.72)	45.84 (45.46 - 46.21)	27.38 (27.00 - 27.75)	62.60 (62.42 - 62.78)	62.13 (61.88 - 62.39)
Shoer	70.45 (70.39 - 70.52)	50.95 (50.37 - 51.54)	62.41 (61.93 - 62.89)	44.57 (43.86 - 45.28)	67.49 (67.30 - 67.69)	66.76 (66.52 - 67.00)
Bhattacharya	69.77 (69.70 - 69.83)	51.90 (51.31 - 52.50)	62.78 (62.30 - 63.26)	39.41 (38.84 - 39.97)	67.67 (67.48 - 67.87)	66.81 (66.52 - 67.10)
WHO	23.92 (23.83 - 24.01)	24.08 (23.45 - 24.70)	24.69 (24.15 - 25.24)	27.29 (26.52 - 28.06)	25.14 (24.90 - 25.38)	30.97 (30.59 - 31.35)
Perez	59.47 (59.39 - 59.55)	45.20 (44.56 - 45.83)	52.27 (51.71 - 52.82)	32.93 (32.23 - 33.64)	58.12 (57.89 - 58.35)	61.00 (60.70 - 61.30)
Mika	69.43 (69.37 - 69.49)	51.43 (50.86 - 52.01)	62.16 (61.68 - 62.63)	45.29 (44.65 - 45.94)	67.08 (66.89 - 67.28)	66.40 (66.13 - 66.68)
Akinbami_1	12.85 (12.77 - 12.94)	11.33 (10.72 - 11.93)	10.22 (9.82 - 10.62)	13.38 (12.58 - 14.18)	11.48 (11.26 - 11.70)	17.70 (17.34 - 18.07)
Akinbami_2	14.69 (14.60 - 14.78)	9.41 (8.89 - 9.92)	9.59 (9.16 - 10.01)	13.16 (12.35 - 13.98)	10.81 (10.60 - 11.03)	17.14 (16.80 - 17.49)
Akinbami_3	27.84 (27.73 - 27.94)	20.23 (19.66 - 20.81)	21.67 (21.14- 22.19)	18.98 (18.22 - 19.73)	26.31 (26.05 - 26.56)	28.93 (28.57 - 29.29)
Salomon	30.97 (30.87 - 31.07)	25.52 (24.84 - 26.20)	27.12 (26.58 - 27.66)	30.64 (29.93 - 31.35)	28.36 (28.10 - 28.61)	39.35 (38.98 - 39.72)
Astley	73.72 (73.65 - 73.78)	48.29 (47.58 - 49.00)	62.47 (61.98 - 62.97)	44.13 (43.32 - 44.93)	67.45 (67.24 - 67.65)	66.85 (66.61 - 67.09)

The Bhattacharya method proposes a clinical symptom-based score [12] given by

$$\begin{aligned} \text{score} \;=\; & 41.7 \times \text{Fever}(> 100^{\circ}\text{F}) + (13.5 \times \text{Cough}) \\ & +(15.8 \times \text{Headache}) + (10 \times \text{Myalgia}) \\ & +(94.7 \times \text{Loss of smell}). \end{aligned} \tag{SM1}$$

If the score is greater than 41.7, then the individual is declared a COVID-19 positive case. The method was examined on responses registered in a clinical screening applied to individuals with suspicion of having COVID-19. The number of participants in this study was 378 of which 125 individuals reported a positive COVID test result.

- Perez

  This method builds a classifier based on logistic regression that considers the following symptoms: anosmia (loss of smell), ageusia (loss of taste), shortness of breath, digestive symptoms, fever, tiredness, sore throat absence, headache, and cough [6]. Then, the Perez method defined different risk scores and assigned them to four symptoms: severe tiredness (1), absence of sore throat (1), fever (2), and anosmia/ageusia (5). This approach declares an individual COVID-19 positive case whether the number of symptoms present is at least 4 and the cumulative score is at least 3. This study considers the data of the seroepidemiological study performed in Spain from April to June 2020. More precisely, more than 61000 participants nationwide completed a questionnaire on symptoms along with SARS-CoV-2 antibodies assays [28]. Notice that the number of positive cases are 2669, out of which 781 (approx 30%) are asymptomatic. The UMD-CTIS survey does not have a question on digestive symptoms. Furthermore, we consider that *severe tiredness* is equivalent to *fatigue*, and *shortness of breath* is equivalent to *difficulty in breathing*.

B. Results

Table SM2
Sensitivity and its 95% confidence interval for the selected countries for 2020, in %

Method	Brazil	Canada	Israel	Japan	Turkey	South Africa
Menni_1	53.11 (53.02 - 53.21)	52.53 (51.78 - 53.28)	48.66 (48.02 - 49.30)	39.15 (38.09 - 40.21)	53.60 (53.30 - 53.90)	50.47 (50.13 - 50.81)
Menni_2	63.94 (63.72 - 64.16)	58.46 (57.75 - 59.16)	56.81 (56.05 - 57.58)	58.57 (57.32 - 59.81)	63.70 (63.43 - 63.98)	65.12 (64.82 - 65.43)
Roland	59.21 (59.08 - 59.33)	62.28 (61.55 - 63.00)	58.96 (58.36 - 59.56)	57.59 (56.31 - 58.87)	60.82 (60.55 - 61.09)	64.96 (64.65 - 65.27)
Smith	72.10 (72.02 - 72.17)	56.26 (55.54 - 56.98)	62.24 (61.65 - 62.84)	50.75 (49.86 - 51.65)	61.52 (61.05 - 61.98)	63.97 (62.01 - 65.93)
Zoabi_55	90.43 (90.18 - 90.67)	35.24 (34.00 - 36.47)	57.15 (56.26 - 58.04)	37.36 (35.86 - 38.86)	79.44 (78.45 - 80.43)	65.39 (64.72 - 66.07)
Zoabi_65	90.68 (90.50 - 90.87)	35.28 (33.87 - 36.69)	58.32 (57.55 - 59.09)	37.82 (36.26 - 39.38)	79.49 (78.34 - 80.63)	65.04 (64.50 - 65.57)
CDC	88.09 (88.03 - 88.16)	88.44 (88.03 - 88.85)	85.84 (85.43 - 86.24)	86.99 (86.36 - 87.63)	89.11 (88.96 - 89.25)	88.57 (88.35 - 88.79)
Shoer	61.40 (61.30 - 61.50)	61.75 (61.05 - 62.46)	58.22 (57.67 - 58.76)	58.24 (57.27 - 59.20)	62.13 (61.86 - 62.41)	64.24 (63.93 - 64.56)
Bhattacharya	60.08 (60.00 - 60.16)	61.97 (61.24 - 62.70)	58.45 (57.91 - 58.99)	60.72 (59.92 - 61.53)	61.40 (61.13 - 61.67)	63.50 (63.18 - 63.82)
WHO	13.88 (13.82 - 13.94)	16.63 (16.15 - 17.11)	15.26 (14.88 - 15.65)	23.66 (22.98 - 24.35)	14.91 (14.75 - 15.08)	19.32 (19.04 - 19.59)
Perez	45.71 (45.62 - 45.80)	46.14 (45.40 - 46.87)	42.38 (41.81 - 42.95)	39.22 (38.37 - 40.07)	46.18 (45.92 - 46.44)	50.64 (50.31 - 50.98)
Mika	59.17 (59.09 - 59.25)	62.25 (61.52 - 62.98)	58.85 (58.30 - 59.39)	55.67 (54.80 - 56.54)	61.08 (60.82 - 61.35)	61.18 (60.85 - 61.51)
Akinbami_1	6.94 (6.89 - 6.99)	6.51 (6.15 - 6.87)	5.53 (5.30 - 5.76)	7.70 (7.22 - 8.18)	6.18 (6.05 - 6.30)	9.93 (9.71 - 10.16)
Akinbami_2	7.99 (7.94 - 8.05)	5.07 (4.78 - 5.35)	5.11 (4.88 - 5.35)	7.21 (6.74 - 7.68)	5.77 (5.65 - 5.89)	9.51 (9.31 - 9.72)
Akinbami_3	16.88 (16.81 - 16.95)	15.28 (14.80 - 15.76)	13.56 (13.20 - 13.92)	15.49 (14.83 - 16.14)	16.31 (16.13 - 16.49)	17.95 (17.70 - 18.20)
Salomon	18.98 (18.91 - 19.05)	18.70 (18.14 - 19.26)	17.40 (16.99 - 17.81)	31.32 (30.58 - 32.05)	17.38 (17.20 - 17.56)	27.36 (27.05 - 27.68)
Astley	69.34 (69.24 - 69.44)	38.82 (38.04 - 39.59)	52.58 (52.04 - 53.12)	34.51 (33.71 - 35.31)	60.87 (60.58 - 61.16)	60.34 (60.01 - 60.67)

Table SM3
Specificity and its 95% confidence interval for the selected countries for 2020, in %

Method	Brazil	Canada	Israel	Japan	Turkey	South Africa
Menni_1	89.77 (89.71 - 89.84)	95.87 (95.79 - 95.96)	96.37 (96.26 - 96.48)	96.59 (96.48 - 96.70)	91.44 (91.31 - 91.57)	92.17 (91.99 - 92.34)
Menni_2	81.41 (81.23 - 81.58)	91.83 (91.71 - 91.94)	89.23 (88.83 - 89.63)	82.83 (82.32 - 83.34)	84.09 (83.91 - 84.27)	82.23 (82.00 - 82.45)
Roland	86.54 (86.44 - 86.64)	91.68 (91.56 - 91.81)	91.71 (91.45 - 91.97)	83.91 (83.16 - 84.66)	87.08 (86.90 - 87.25)	84.01 (83.81 - 84.21)
Smith	64.18 (64.09 - 64.27)	94.37 (94.28 - 94.47)	90.02 (89.68 - 90.36)	90.56 (90.39 - 90.72)	86.63 (86.22 - 87.04)	75.98 (73.43 - 78.53)
Zoabi_55	23.21 (22.64 - 23.79)	91.56 (90.65 - 92.48)	78.31 (77.69 - 78.93)	85.44 (84.14 - 86.73)	40.05 (38.43 - 41.67)	67.92 (66.89 - 68.95)
Zoabi_65	22.73 (22.31 - 23.15)	91.46 (90.45 - 92.47)	77.39 (76.93 - 77.85)	85.05 (83.71 - 86.38)	39.53 (37.65 - 41.41)	68.31 (67.53 - 69.08)
CDC	39.32 (39.21 - 39.42)	40.87 (40.68 - 41.06)	49.90 (49.59 - 50.22)	43.06 (42.75 - 43.37)	39.57 (39.38 - 39.77)	44.13 (43.84 - 44.41)
Shoer	84.86 (84.78 - 84.95)	91.60 (91.46 - 91.7)5	92.53 (92.35 - 92.71)	86.83 (86.39 - 87.27)	86.18 (86.02 - 86.34)	83.90 (83.67 - 84.14)
Bhattacharya	85.81 (85.74 - 85.88)	92.01 (91.88 - 92.13)	92.67 (92.52 - 92.83)	81.11 (80.89 - 81.33)	87.43 (87.28 - 87.58)	84.95 (84.73 - 85.16)
WHO	97.51 (97.48 - 97.54)	97.79 (97.74 - 97.85)	97.86 (97.78 - 97.94)	93.69 (93.55 - 93.84)	97.68 (97.62 - 97.73)	96.96 (96.88 - 97.05)
Perez	90.64 (90.58 - 90.70)	93.98 (93.88 - 94.08)	94.79 (94.66 - 94.92)	87.46 (87.26 - 87.66)	91.96 (91.84 - 92.08)	91.13 (90.96 - 91.30)
Mika	86.75 (86.68 - 86.82)	91.69 (91.57 - 91.82)	91.96 (91.81 - 92.11)	88.55 (88.37 - 88.74)	86.78 (86.64 - 86.93)	86.83 (86.57 - 87.10)
Akinbami_1	98.80 (98.78 - 98.82)	99.18 (99.14 - 99.21)	99.33 (99.29 - 99.38)	99.17 (99.11 - 99.22)	99.13 (99.09 - 99.16)	98.76 (98.69 - 98.82)
Akinbami_2	99.02 (99.00 - 99.04)	99.78 (99.75 - 99.80)	99.66 (99.62 - 99.69)	99.80 (99.77 - 99.83)	99.44 (99.42 - 99.47)	99.24 (99.19 - 99.29)
Akinbami_3	94.85 (94.81 - 94.89)	96.30 (96.21 - 96.40)	96.97 (96.87 - 97.06)	94.02 (93.88 - 94.17)	95.15 (95.06 - 95.23)	96.50 (96.41 - 96.59)
Salomon	95.80 (95.76 - 95.84)	97.14 (97.08 - 97.20)	97.17 (97.08 - 97.26)	90.72 (90.56 - 90.88)	96.74 (96.68 - 96.80)	93.36 (93.20 - 93.51)
Astley	77.93 (77.81 - 78.05)	97.75 (97.68 - 97.82)	95.86 (95.75 - 95.98)	97.26 (97.15 - 97.38)	87.67 (87.51 - 87.83)	88.50 (88.29 - 88.72)

Table SM4
Precision and its 95% confidence interval for the selected countries for 2020, in %

Method	Brazil	Canada	Israel	Japan	Turkey	South Africa
Menni_1	85.64 (85.55 - 85.73)	56.41 (55.64 - 57.18)	77.58 (76.92 - 78.24)	57.27 (56.13 - 58.42)	79.22 (78.94 - 79.49)	78.42 (77.97 - 78.87)
Menni_2	80.17 (80.06 - 80.29)	42.76 (42.18 - 43.34)	58.58 (57.76 - 59.39)	30.51 (29.86 - 31.17)	71.59 (71.33 - 71.86)	67.69 (67.33 - 68.05)
Roland	83.80 (83.70 - 83.89)	43.89 (43.30 - 44.48)	65.41 (64.70 - 66.12)	32.00 (31.04 - 32.96)	74.77 (74.51 - 75.03)	69.90 (69.54 - 70.26)
Smith	70.17 (70.09 - 70.26)	50.99 (50.35 - 51.64)	62.45 (61.61 - 63.29)	40.75 (40.05 - 41.46)	74.46 (73.99 - 74.93)	63.42 (61.33 - 65.51)
Zoabi_55	58.07 (57.94 - 58.19)	32.48 (31.33 - 33.63)	41.25 (40.69 - 41.82)	26.07 (25.15 - 26.99)	45.70 (45.27 - 46.13)	54.01 (53.53 - 54.49)
Zoabi_65	57.98 (57.86 - 58.09)	32.55 (31.35 - 33.76)	40.62 (40.12 - 41.12)	25.84 (24.92 - 26.75)	45.57 (45.09 - 46.05)	54.09 (53.68 - 54.50)
CDC	62.94 (62.86 - 63.02)	13.51 (13.33 - 13.70)	31.28 (30.96 - 31.61)	16.26 (16.01 - 16.51)	48.25 (48.06 - 48.45)	47.87 (47.59 - 48.15)
Shoer	82.66 (82.58 - 82.75)	43.48 (42.84 - 44.11)	67.35 (66.74 - 67.96)	36.36 (35.54 - 37.17)	73.90 (73.65 - 74.14)	69.54 (69.17 - 69.91)
Bhattacharya	83.19 (83.11 - 83.27)	44.74 (44.13 - 45.36)	67.88 (67.31 - 68.45)	29.23 (28.73 - 29.74)	75.40 (75.16 - 75.64)	70.53 (70.15 - 70.91)
WHO	86.76 (86.60 - 86.92)	43.98 (43.01 - 44.95)	65.37 (64.37 - 66.36)	32.43 (31.45 - 33.41)	80.16 (79.76 - 80.55)	78.42 (77.85 - 79.00)
Perez	85.11 (85.01 - 85.20)	44.42 (43.73 - 45.12)	68.34 (67.67 - 69.01)	28.49 (27.80 - 29.17)	78.42 (78.14 - 78.70)	76.76 (76.36 - 77.16)
Mika	84.00 (83.92 - 84.08)	43.91 (43.32 - 44.50)	65.93 (65.38 - 66.48)	38.29 (37.65 - 38.94)	74.42 (74.19 - 74.64)	72.68 (72.25 - 73.11)
Akinbami_1	87.17 (86.95 - 87.39)	45.02 (43.20 - 46.84)	68.68 (66.88 - 70.48)	53.66 (51.23 - 56.09)	81.61 (80.88 - 82.33)	82.30 (81.48 - 83.13)
Akinbami_2	90.55 (90.37 - 90.73)	70.32 (68.01 - 72.64)	79.68 (77.85 - 81.52)	81.58 (79.19 - 83.97)	86.57 (85.93 - 87.22)	87.77 (87.04 - 88.50)
Akinbami_3	79.35 (79.20 - 79.51)	30.27 (29.42 - 31.13)	54.35 (53.29 - 55.41)	24.78 (23.78 - 25.77)	68.12 (67.66 - 68.58)	74.82 (74.21 - 75.43)
Salomon	84.15 (84.00 - 84.30)	40.49 (39.57 - 41.41)	61.96 (61.09 - 62.82)	30.12 (29.35 - 30.89)	77.02 (76.65 - 77.40)	70.22 (69.67 - 70.76)
Astley	78.69 (78.59 - 78.79)	64.33 (63.51 - 65.16)	77.06 (76.49 - 77.63)	61.82 (60.67 - 62.97)	75.65 (75.39 - 75.91)	75.02 (74.63 - 75.41)

Table SM5
F_1 score and its 95% confidence interval for the selected countries for 2021, in %

Method	Brazil	Canada	Israel	Japan	Turkey	South Africa
Menni_1	59.24 (59.18 - 59.31)	49.38 (49.02- 49.74)	57.31 (56.96 - 57.65)	49.24 (49.16 - 49.83)	59.65 (59.44 - 59.87)	58.28 (58.06 - 58.50)
Menni_2	66.54 (66.49 - 66.59)	39.82 (39.59 - 40.05)	53.46 (53.21 - 53.70)	42.60 (42.37 - 42.84)	62.71 (62.56 - 62.85)	66.50 (66.33 - 66.68)
Roland	65.76 (65.71 - 65.82)	46.28 (46.03 - 46.53)	57.16 (56.86 - 57.46)	42.82 (42.62 - 43.03)	64.13 (63.96 - 64.31)	64.41 (64.23 - 64.59)
Smith	63.37 (63.32 - 63.42)	50.28 (49.99 - 50.57)	58.00 (57.68 - 58.33)	51.48 (51.23 -51.74)	64.38 (64.21 - 64.55)	61.62 (61.45 - 61.80)
Zoabi_55	59.83 (59.79 - 59.88)	37.31 (37.01 - 37.60)	39.63 (39.28 - 39.98)	33.71 (33.45 - 33.98)	52.14 (51.88 - 52.40)	59.62 (59.47 - 59.77)
Zoabi_65	59.78 (59.74 - 59.83)	37.10 (36.81 - 37.39)	39.64 (39.29 - 39.99)	33.36 (33.11 - 33.62)	52.06 (51.80 - 52.31)	59.54 (59.38 - 59.69)
CDC	63.22 (63.17 - 63.26)	27.41 (27.28 - 27.55)	38.78 (38.59 - 38.97)	28.54 (28.40 - 28.68)	55.96 (55.81 - 56.11)	61.25 (61.10 - 61.39)
Shoer	65.81 (65.76 - 65.87)	41.10 (40.84 - 41.36)	53.67 (53.37 - 53.97)	45.42 (45.07 - 45.78)	64.18 (64.01 - 64.35)	64.97 (64.80 - 65.15)
Bhattacharya	64.16 (64.11 - 64.22)	49.22 (48.96 - 49.49)	58.76 (58.48 - 59.03)	45.82 (45.59 - 46.05)	64.61 (64.44 - 64.78)	63.40 (63.22 - 63.59)
WHO	23.62 (23.56 - 23.68)	26.01 (25.66 - 26.35)	27.92 (27.59 - 28.24)	34.05 (33.74 - 34.37)	27.72 (27.49 - 27.94)	32.78 (32.58 - 32.98)
Perez	54.85 (54.79 - 54.90)	44.70 (44.40 - 45.00)	51.27 (50.93 - 51.61)	39.72 (39.45 - 40.00)	56.03 (55.86 - 56.21)	59.17 (58.98 - 59.35)
Mika	65.33 (65.28 - 65.38)	46.76 (46.40 - 47.12)	57.50 (57.22 - 57.79)	52.41 (51.73 - 53.09)	64.13 (63.96 - 64.31)	63.98 (63.81 - 64.15)
Akinbami_1	12.02 (11.96 - 12.07)	11.43 (11.17 - 11.70)	10.60 (10.33 - 10.88)	11.11 (10.82 - 11.39)	13.86 (13.69 - 14.03)	15.86 (15.66 - 16.06)
Akinbami_2	12.02 (12.05 - 12.16)	8.03 (7.79 - 8.27)	11.48 (11.20 - 11.75)	9.10 (8.83 - 9.31)	11.80 (11.64 - 11.96)	13.61 (13.44 - 13.79)
Akinbami_3	26.59 (26.00 - 26.11)	20.96 (20.64 - 21.27)	21.96 21.62 - 22.30)	19.90 (19.63 - 20.17)	26.35 (26.12 - 26.58)	28.08 (27.85 - 28.31)
Salomon	30.15 (30.11 - 30.24)	28.06 (27.70 - 28.43)	30.72 (30.39 - 31.05)	37.27 (36.97 - 37.57)	31.31 (31.09 - 31.53)	38.03 (37.83 - 38.23)
Astley	65.95 (65.90 - 66.01)	45.07 (44.74 - 45.40)	58.62 (58.29 - 58.94)	50.39 (50.08 - 50.70)	63.67 (63.50 - 63.85)	64.06 (63.88 - 64.24)

Table SM6
Sensitivity and its 95% confidence interval for the selected countries for 2021, in %

Method	Brazil	Canada	Israel	Japan	Turquey	South Africa
Menni_1	45.56 (45.49 - 45.62)	41.96 (41.59 - 42.33)	47.46 (47.11 - 47.81)	38.87 (38.53 - 39.22)	48.34 (48.09 - 48.59)	45.44 (45.20 - 45.68)
Menni_2	61.38 (61.33 - 61.43)	70.30 (69.94 - 70.65)	67.92 (67.50 - 68.33)	67.74 (67.46 - 68.02)	59.35 (59.11 - 59.58)	67.11 (66.88 - 67.34)
Roland	59.02 (58.96 - 59.09)	56.82 (56.49 - 57.15)	61.26 (60.81 - 61.72)	66.10 (65.82 - 66.38)	59.02 (58.81 - 59.24)	60.86 (60.60 - 61.12)
Smith	69.00 (68.94 - 69.05)	48.06 (47.73 - 48.39)	62.43 (62.09 - 62.78)	50.07 (49.78 - 50.37)	60.80 (60.25 - 61.35)	74.04 (73.85 - 74.24)
Zoabi_55	66.13 (65.66 - 66.59)	38.77 (37.62 - 39.91)	48.89 (48.07 - 49.71)	37.39 (36.37 - 38.40)	62.51 (62.14 - 62.88)	69.70 (69.44 - 69.96)
Zoabi_65	66.66 (65.75 - 67.57)	39.34 (38.05 - 40.62)	48.95 (48.16 - 49.74)	40.18 (38.51 - 41.86)	62.07 (61.71 - 62.42)	69.61 (69.30 - 69.92)
CDC	42.56 (42.51 - 42.61)	88.39 (88.18 - 88.61)	85.05 (84.82 - 85.29)	85.90 (85.69 - 86.12)	87.19 (87.06 - 87.32)	88.09 (87.96 - 88.23)
Shoer	87.72 (87.68 - 87.76)	66.90 (66.53 - 67.27)	66.13 (65.70 - 66.57)	67.12 (66.67 - 67.57)	59.08 (58.85 - 59.30)	82.11 (81.92 - 82.31)
Bhattacharya	54.12 (54.06 - 54.18)	53.51 (53.17 - 53.85)	57.80 (57.44 - 58.16)	88.79 (88.73 - 88.85)	59.01 (58.79 - 59.22)	84.16 (84.03 - 84.29)
WHO	13.84 (13.80 - 13.88)	17.97 (17.71 - 18.24)	17.95 (17.71 - 18.19)	26.08 (25.81 - 26.36)	17.08 (16.92 - 17.24)	20.99 (20.84 - 21.15)
Perez	41.39 (41.33 - 41.44)	42.56 (42.24 - 42.89)	43.68 (43.35 - 44.02)	41.14 (40.82 - 41.46)	45.31 (45.11 - 45.51)	48.66 (48.45 - 48.87)
Mika	57.42 (57.36 - 57.48)	57.87 (57.35 - 58.40)	61.65 (61.30 - 62.00)	56.97 (56.54 - 57.41)	59.02 (58.81 - 59.24)	83.17 (83.05 - 83.28)
Akinbami_1	6.49 (6.46 - 6.52)	6.58 (6.42 - 6.75)	5.82 (5.66 - 5.98)	6.10 (5.93 - 6.27)	7.65 (7.55 - 7.75)	8.87 (8.75 - 8.99)
Akinbami_2	6.50 (6.47 - 6.53)	4.29 (4.16 - 4.43)	6.18 (6.02 - 6.34)	4.79 (4.66 - 4.93)	6.36 (6.27 - 6.45)	7.41 (7.31 - 7.51)
Akinbami_3	15.91 (15.87 - 15.95)	14.98 (14.74 - 15.23)	14.16 (13.92 - 14.41)	14.67 (14.46 - 14.89)	16.77 (16.61 - 16.94)	17.61 (17.44 - 17.77)
Salomon	18.76 (18.71 - 18.81)	20.50 (20.20 - 20.80)	20.68 (20.43 - 20.94)	32.02 (31.73 - 32.32)	20.11 (19.93 - 20.28)	26.68 (26.50 - 26.86)
Astley	56.60 (56.53 - 56.67)	34.44 (34.12 - 34.76)	48.80 (48.42 - 49.19)	38.99 (38.67 - 39.32)	55.19 (54.96 - 55.41)	89.09 (88.96 - 89.21)

Table SM7
Specificity and its 95% confidence interval for the selected countries for 2021, in %

Method	Brazil	Canada	Israel	Japan	Turquey	South Africa
Menni_1	94.67 (94.65 - 94.70)	96.88 (96.84 - 96.91)	96.98 (96.92 - 97.04)	98.12 (98.09 - 98.15)	93.73 (93.65 - 93.80)	93.61 (93.52 - 93.70)
Menni_2	84.24 (84.21 - 84.28)	79.32 (79.11 - 79.53)	84.77 (84.57 - 84.96)	83.62 (83.48 - 83.77)	85.95 (85.81 - 86.09)	79.86 (79.70 - 80.02)
Roland	86.03 (85.99 - 86.06)	89.97 (89.90 - 90.04)	90.61 (90.38 - 90.84)	84.46 (84.39 - 84.53)	88.24 (88.15 - 88.34)	83.69 (83.53 - 83.85)
Smith	67.04 (67.00 - 67.09)	95.14 (95.08 - 95.19)	90.65 (90.44 - 90.86)	95.15 (95.11 - 95.20)	86.82 (86.34 - 87.31)	61.61 (61.44 - 61.79)
Zoabi_55	62.52 (61.75 - 63.29)	92.17 (91.54 - 92.81)	85.42 (84.91 - 85.92)	90.86 (90.30 - 91.41)	70.58 (70.34 - 70.81)	62.74 (62.40 - 63.09)
Zoabi_65	61.52 (60.02 - 63.03)	91.75 (91.04 - 92.46)	85.40 (84.93 - 85.87)	89.05 (88.07 - 90.02)	70.91 (70.71 - 71.12)	62.68 (62.24 - 63.11)
CDC	85.62 (85.58 - 85.66)	48.56 (48.44 - 48.68)	55.38 (55.23 - 55.54)	54.73 (54.63 - 54.83)	42.38 (42.23 - 42.53)	42.44 (42.26 - 42.63)
Shoer	57.88 (57.81 - 57.94)	82.04 (81.80 - 82.28)	85.79 (85.51 - 86.07)	85.96 (85.63 - 86.29)	88.25 (88.13 - 88.38)	62.96 (62.68 - 63.23)
Bhattacharya	90.14 (90.12 - 90.17)	92.79 (92.73 - 92.85)	93.13 (93.04 - 93.22)	60.31 (60.00 - 60.61)	89.02 (88.93 - 89.12)	59.11 (58.88 - 59.35)
WHO	97.73 (97.72 - 97.74)	97.72 (97.68 - 97.75)	98.12 (98.08 - 98.17)	97.05 (97.01 - 97.08)	97.12 (97.07 - 97.17)	95.88 (95.81 - 95.94)
Perez	93.57 (93.54 -93.60)	94.60 (94.55 - 94.65)	95.30 (95.23 - 95.37)	92.82 (92.77 - 92.87)	92.41 (92.33 - 92.49)	90.86 (90.76 - 90.96)
Mika	87.47 (87.44 - 87.50)	89.78 (89.28 - 90.28)	90.67 (90.49 - 90.84)	93.26 (92.81 - 93.71)	88.24 (88.15 - 88.34)	60.61 (60.47 - 60.90)
Akinbami_1	98.97 (98.96 - 98.98)	99.04 (99.02 - 99.06)	99.31 (99.29 - 99.34)	99.60 (99.58 - 99.61)	98.73 (98.70 - 98.76)	98.24 (98.21 - 98.28)
Akinbami_2	99.38 (99.37 - 99.39)	99.70 (99.69 - 99.71)	99.75 (99.73 - 99.76)	99.90 (99.89 - 99.91)	99.33 (99.31 - 99.36)	99.16 (99.13 - 99.19)
Akinbami_3	95.80 (95.78 - 95.82)	96.84 (96.81 - 96.88)	97.41 (97.36 - 97.46)	96.43 (96.40 - 96.47)	95.13 (95.07 - 95.19)	95.50 (95.42 - 95.58)
Salomon	96.20 (96.18 - 96.22)	97.11 (97.07 - 97.15)	97.54 (97.48 - 97.59)	95.66 (95.62 - 95.70)	96.11 (96.06 - 96.17)	92.10 (92.01 - 92.20)
Astley	89.74 (89.70 - 89.78)	97.92 (97.88 - 97.96)	96.88 (96.82 - 96.94)	98.28 (98.25 - 98.31)	91.48 (91.40 - 91.57)	56.00 (55.76 - 56.24)

Table SM8
Precision and its 95% confidence interval for the selected countries for 2021, in %

Method	Brazil	Canada	Israel	Japan	Turquey	South Africa
Menni_1	84.66 (84.66 - 84.74)	59.99 (59.68 - 60.29)	72.38 (71.90 - 72.87)	67.16 (67.89 - 68.31)	77.92 (77.69 - 78.15)	81.23 (81.15 - 81.32)
Menni_2	72.65 (72.6 - 72.7)	27.78 (27.61 - 27.95)	44.12 (43.80 - 44.43)	31.07 (30.88 - 31.27)	66.50 (66.27 - 66.73)	65.90 (65.79 - 66.03)
Roland	74.24 (74.21 - 74.28)	39.04 (38.84 - 39.24)	53.71 (53.16 - 54.26)	31.67 (31.51 - 31.83)	70.23 (70.01 - 70.44)	68.40 (68.32 - 68.48)
Smith	58.59 (58.55 - 58.64)	52.72 (52.47 - 52.96)	54.26 (53.71 - 54.82)	52.97 (52.77 - 53.19)	68.70 (68.07 - 69.33)	52.77 (52.62 - 52.93)
Zoabi_55	54.70 (54.50 - 54.90)	37.09 (36.22 - 37.96)	33.59 (33.05 - 34.13)	31.67 (30.82 - 32.52)	44.75 (44.47 - 45.03)	52.11 (51.90 - 52.32)
Zoabi_65	54.44 (54.05 - 54.84)	36.51 (35.55 - 37.46)	33.55 (33.03 - 34.07)	30.17 (29.16 - 31.18)	44.85 (44.58 - 45.12)	52.04 (51.80 - 52.29)
CDC	50.14 (50.10 - 50.18)	16.22 (16.14 - 16.31)	25.12 (24.96 - 25.28)	17.11 (17.02 - 17.20)	41.21 (41.05 - 41.36)	46.95 (46.81 - 47.07)
Shoer	52.66 (52.61 - 52.72)	29.66 (29.46 - 29.86)	45.25 (44.80 - 45.71)	34.32 (34.04 - 34.62)	70.27 (70.02 - 70.52)	53.75 (53.60 - 53.91)
Bhattacharya	78.77 (78.75 - 78.83)	45.57 (45.37 - 45.78)	59.80 (59.44 - 60.15)	30.88 (30.68 - 31.08)	71.42 (71.21 - 71.62)	50.86 (50.67 - 51.05)
WHO	80.52 (80.48 - 80.56)	47.07 (46.56 - 47.45)	62.91 (62.25 - 63.57)	49.04 (48.70 - 49.37)	73.59 (73.25 - 73.93)	74.79 (74.61 - 74.84)
Perez	81.28 (81.25 - 81.31)	47.07 (46.79 - 47.33)	62.09 (61.64 - 62.55)	38.39 (38.17 - 38.64)	73.44 (73.20 - 73.67)	75.47 (75.36 - 75.55)
Mika	75.77 (75.74 - 75.80)	39.23 (38.96 - 39.49)	53.95 (53.50 - 54.41)	48.53 (47.67 - 49.37)	70.23 (70.01 - 70.44)	51.99 (51.81 - 52.17)
Akinbami_1	80.80 (80.48 - 81.13)	43.47 (42.94 - 43.87)	59.89 (58.75 - 61.02)	62.18 (61.69 - 62.77)	73.71 (73.20 - 74.22)	74.83 (74.47 - 75.20)
Akinbami_2	87.91 (87.60 - 88.23)	61.62 (61.14 - 62.10)	81.22 (80.28 - 82.16)	83.71 (83.45 - 83.98)	81.64 (81.09 - 82.18)	83.73 (83.26 - 84.20)
Akinbami_3	71.91 (71.89 - 71.93)	34.89 (34.42 - 35.25)	49.06 (48.39 - 49.73)	30.93 (30.55 - 31.25)	61.52 (61.15 - 61.90)	69.26 (69.09 - 69.58)
Salomon	77.08 (77.07 - 77.08)	44.45 (44.06 - 44.9)	59.81 (59.19 - 60.43)	44.58 (44.28 - 44.86)	70.84 (70.51 - 71.16)	66.19 (66.08 - 66.29)
Astley	79.01 (78.99 - 79.04)	65.49 (64.96 - 66.02)	73.47 (73.04 - 73.90)	71.21 (71.04 - 71.35)	75.27 (75.06 - 75.49)	50.01 (49.83 - 50.19)

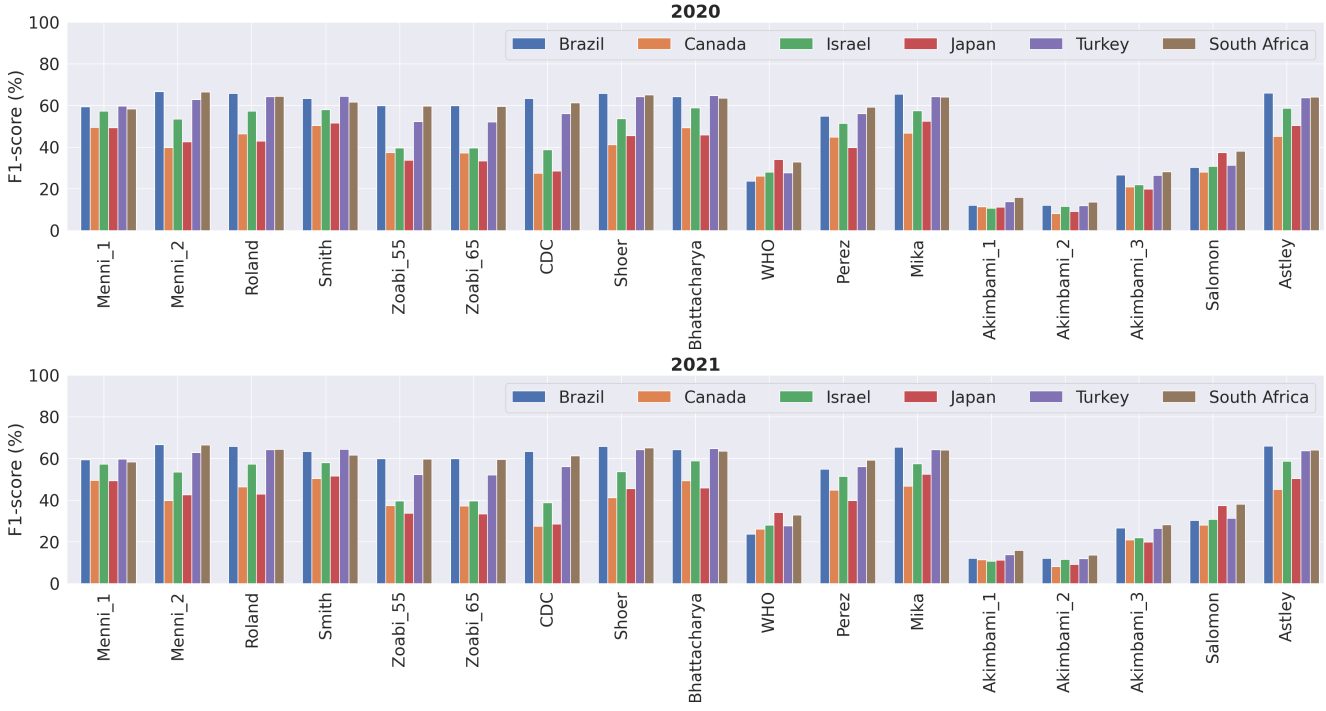

Figure SM1: F_1 scores in % obtained by each COVID-19 detection method across the six countries for 2020 and 2021.