



Prediction of COVID-19 New Cases Using Multiple Linear Regression Model Based on May to June 2020 Data in Ethiopia

**Alemayehu Siffir Argawu^{a*}, Gizachew Gobebo^a, Ketema Bedane^a,
Temesgen Senbeto^a, Reta Lemessa^a and Agassa Galdassa^a**

^a *College of Natural and Computational Sciences, Ambo University, Ambo, P. O. Box 19, Ethiopia.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2021/v33i51A33468

Editor(s):

(1) Dr. Dharmesh Chandra Sharma, G. R. Medical College & J. A. Hospital, India.

(2) Dr. Asmaa Fathi Moustafa Hamouda, Jazan University, Saudi Arabia.

Reviewers:

(1) Dzulsuhaimi Daud, Universiti Teknologi Mara (Perak Branch, Tapah Campus), Malaysia.

(2) Sónia Ferreira, University of Aveiro, Portugal.

(3) Jesús G. Rodríguez Diego, Autonomous Metropolitan University- Xochimilco Unit. México, México.

Complete Peer review History: <https://www.sdiarticle4.com/review-history/76078>

Original Research Article

Received 30 August 2021

Accepted 30 October 2021

Published 20 November 2021

ABSTRACT

The aims of this study was to predict COVID-19 new cases using multiple linear regression model based on May to June 2020 data in Ethiopia. The COVID-19 cases data was collected from the Ethiopia Ministry of Health Organization Facebook page. Pearson's correlation analysis and linear regression model were used in the study. And, the COVID-19 new cases was positively correlated with the number of days, daily laboratory tests, new cases of males, new cases of females, new cases from Addis Ababa city, and new cases from foreign natives. In the multiple linear regression model, COVID-19 new cases was significantly predicted by the number of days at 5%, the number of daily laboratory tests at 10%, and the number of new cases from Addis Ababa city at 1% levels of significance. Then, the researchers recommended that Ethiopian Government, Ministry of Health, and Addis Ababa city administrative should give more awareness and protections for societies, and they should open again more COVID-19 laboratory testing centers. And, this study will help the government and doctors in preparing their plans for the next times.

Keywords: COVID-19 new cases; laboratory tests; number of days; multiple linear regression model; Ethiopia.

1. INTRODUCTION

Coronavirus disease (COVID-19) is an infectious disease that is caused by severe acute respiratory syndrome known as coronavirus. The COVID-19 was first identified on 31 December 2019 in the city of Wuhan, which is the capital of Hubei Province in China. Some of the common signs of COVID-19 include fever, shortness of breath, and dry coughs. Other uncommon symptoms include muscle pain, mild diarrhea, abdominal pain, sputum production, loss of smell, and sore throat [1-3]. On 11 March 2020, the WHO announced that it was a global pandemic [4].

As the Worldometer coronavirus updates information reported on 15th of June 2020, we have 8,028,253 total COVID-19 cases, and 4,148,128 totals recovered with 51.7% recovered rate as globally. It was distributed from highest to lowest ranks of the new cases by the World Regions as follows: North America has led by 2,480,701 total new cases (1st), Europe has 2,398,779 total new cases (2nd), Asia has 1,616,962 total new cases (3rd), South America has 1,425,696 total new cases (4th), Africa has 244,578 total new cases (5th), and the last was Oceania by 8,931 cases. In the report, the male and female cases were 71% and 29%, respectively [5]. And, Ethiopia was ranked as the 2nd, 15th, 16th, and 23rd on the table by 176 new cases, by 3,521 total COVID-19 cases, and by 620 total recovered (17.6%) as compared from African countries on this date as the Worldometer coronavirus updates information showed [5].

This report also showed that Ethiopia was listed in 27th place by the capacity of COVID-19 laboratory tests. It was 1,629 tests per 1,000,000 populations. This is bad news for Ethiopia. Currently, the Ethiopian population is near 115 million. This is the fact that Ethiopia has a very low proportion of COVID-19 laboratory tests compared with other countries' tests. This report indicated that Ethiopia needs increasing efforts and strategies to increase daily laboratory tests. Otherwise, Ethiopia will be the next "African USA" by beating COVID-19.

And, as the researchers observed COVID-19 new cases were alarmingly increased in the study period (12th of May to 10th June, 2020) as the Ethiopia Ministry of Health Organization Facebook page report shown. Then, this study

aimed to predict COVID-19 new cases using multiple linear regression model based on May to June 2020 data in Ethiopia.

2. MATERIALS AND METHODS

2.1 Source of Data and Study Period

The COVID-19 new case report data were collected from the Ethiopia ministry of health organization Facebook page. The study period of data was from 12th of May to 10th of June 2020 (for the last 30 days) since complete information was available on this study period only but not on the previous dates on the Facebook page. Total number of COVID-19 new cases, date of record, number of new recoveries, number of new cases from Addis Ababa city and some regions, number of males and females, maximum and minimum ages of the patients were collected and included in the study.

2.2 Pearson's Correlation Coefficient

Correlation is a statistical method used to assess a possible linear association between two continuous variables. It is simple both to calculate and to interpret. Pearson's correlation coefficient is denoted as r for a sample statistic. For a correlation between variables x and y , the formula for calculating the sample Pearson's correlation coefficient is given below [6].

$$r = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{[\sum_{i=1}^n (x_i - \bar{x})^2][\sum_{i=1}^n (y_i - \bar{y})^2]}} \quad [1]$$

Where: x_i and y_i are values for variables x and y for the i^{th} individual.

2.3 Linear Regression Model

The regression model has many variants such as linear regression, polynomial regression, and others [7]. In this study, the fitted simple and multiple linear regression models were used to determine the most predictor variables for COVID-19 new cases from 12th May to 10th June 2020 in Ethiopia.

The fitted simple linear regression model equation is written as follow:

$$\hat{Y} = b_0 + b_1 X_1 \quad [2]$$

Where: \hat{Y} an estimated COVID-19 new cases. And X_1 is an independent variable with its corresponding estimated coefficients (b_1). And, b_0 is the intercept coefficient in the model.

The fitted multiple linear regression model equation is given as follow:

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \quad [3]$$

Where, \hat{Y} an estimated COVID-19 new cases. And X_1 , X_2 , and X_3 are independent variables with their corresponding estimated coefficients (b_1 , b_2 , and b_3). And, b_0 is the intercept coefficient in the model.

2.4 Polynomial Regression Model

The fitted quadratic and cubic regression models were used to estimate the parameters of independent variables as X , X^2 , and X^3 . All the estimated parameters (b_1 , b_2 , and b_3) shown the change of Y when the independent variable changed from x to $x+1$ [7]. The fitted quadratic regression model equation is given as follow:

$$\hat{Y} = b_0 + b_1X + b_2X^2 \quad [4]$$

The fitted cubic regression model equation is given as follow:

$$\hat{Y} = b_0 + b_1X + b_2X^2 + b_3X^3 \quad [5]$$

Where: all terms are stated above.

3. RESULTS

3.1 Frequency Statistics of COVID-19 Cases

The total numbers of COVID-19 laboratory tests were 36,583 and 121,938 from 14th of March to 11th of May and 12th of May to 10th of June, respectively. This indicates that the number of total laboratory tests was increased by 3.33 times. The newly reported COVID-19 laboratory tests peaked at 6,092 on June 07, 2020. The total numbers of COVID-19 cases were 249 and 2,257 from 14th of March to 11th of May and from 12th of May to 10th of June, respectively. This indicates that the number of total cases was

increased by 9.1 times. The new cases peaked at 190 cases on June 09, 2020. The numbers of total recoveries were 107 and 294 from 14th March to 11th of May and from 12th May to 10th June, respectively. This indicates that the number of total recoveries was increased by 2.75 times (Table 1).

3.2 COVID-19 New Cases by Regions and Genders

From the total number of 2,257 COVID-19 new cases, the majorities (64%) were males and 36% of them were females. This indicated that the male group was infected less in Ethiopia as compared with world male cases was 71%. Addis Ababa city has covered the majority (74%) of the Pandemic. The Somali region (7%) has taken the 2nd highest coverage of the virus. Oromia and Amhara regions have equally shared the COVID-19 new cases (each 6%). However, Tigray and other regions (such as SNNP, Afar, Harare, etc.) have taken only 3% of COVID-19 new cases by each. The foreign natives had shared only 1% of the distribution (Table 2).

3.3 Descriptive Statistics of COVID-19 Cases

The average value of COVID-19 conducted laboratory tests was 4,065 per day with its min (1,775) and max (6,187) in the given duration. The average value of COVID-19 new cases was 75 per day, with minimum and maximum values of 2 and 190, respectively. Addis Ababa (ADDIS ABABA) city had recorded the highest COVID-19 new cases (56) per day in the given duration. In this duration, the maximum and minimum new cases in Addis Ababa city were 153 and 0 with mean value of 56 new cases, respectively. The city covered more than 70% of total cases in the country. The average values of the minimum and maximum ages of COVID-19 new cases were 9.4 years and 71 years with their smallest and largest ages of 1 month and 115 years, respectively (Table 3).

Table 1. Percentage and rates of COVID-19 cases from 14th of March to 11th of May and from 12th May to 10th June in Ethiopia, 2020

| Variable | Duration | |
|------------------------|--|---|
| | 14 th March to 11 th May | 12 th May to 10 th June |
| Total laboratory tests | 36,583 | 121, 938 |
| COVID-19 total cases | 249 | 2, 257 |
| Total recoveries | 107 | 294 |

Table 2. Prevalence of COVID-19 new cases by gender and region from 12th May to 10th June in Ethiopia, 2020

| Prevalence by Category | # of days | Total | % |
|-------------------------|-----------|-------|-----|
| Gender | | 2,257 | 100 |
| Male | 30 | 1,441 | 64 |
| Female | 30 | 816 | 36 |
| Region | | 2,237 | 99 |
| Addis Ababa City | 30 | 1,671 | 74 |
| Oromia Region | 30 | 130 | 6 |
| Amhara Region | 30 | 134 | 6 |
| Tigray Region | 30 | 78 | 3 |
| Somali Region | 30 | 157 | 7 |
| Other Regions* | 30 | 67 | 3 |
| Other (Foreign Natives) | 30 | 20 | 1 |

*Other regions are SNNP, Afar, Harare, Dire Daw administrative city, Gambela, and Benishagul-gumuz

Table 3. Descriptive statistics of the COVID-19 cases from 12th May to 10th June in Ethiopia, 2020

| Variables | Min. | Max. | Mean | S.D |
|---------------------------------|------|-------|-------|-------|
| Daily laboratory tests | 1775 | 6,187 | 4,065 | 1,094 |
| COVID-19 new cases | 2 | 190 | 75 | 57 |
| New cases from Addis Ababa city | 0 | 153 | 56 | 48 |
| Minimum age of new cases | 0.1 | 24 | 9.4 | 6.9 |
| Maximum age of new cases | 33 | 115 | 71 | 18 |

3.4 Correlation Analysis for COVID-19 New Cases

The correlation analysis showed that there were significant positive correlations between COVID-19 new cases and the number of days, daily laboratory tests, new cases of males, new cases of females, new cases from ADDIS ABABA city, and new cases from foreign natives (Table 4).

3.5 Regression Model for COVID-19 New Cases

The linear regression model had the highest F-value (120.7) and the smallest MSE value (637.4) as compared with quadratic and cubic models. And, the number of days was a significant predictor for new cases in the linear regression model (p-value of 0.000). But, this variable and its two expressions were not predictors in the quadratic and cubic regression model. And, this fitted linear regression model was much better than the quadratic and cubic models. However, the three models have similar R square values like 81% and 82% variations of COVID-19 new cases was explained by the models (Table 5 and Fig. 1).

The estimated linear regression equation was given as:-

$\widehat{\text{New Cases}} = -378.2 + 5.851 * \text{number of the day}$. This implied that the new cases will be increased to 585 after 100 days.

Daily laboratory test was also significant predictor for new cases in the linear regression model (p-value of 0.000). The fitted linear regression model has the highest F-value (19.5) and but not the smallest MSE value (1993.3) as compared with quadratic and cubic models. Thus, the fitted linear regression model was much better than the quadratic and cubic models. However, the cubic regression model has a better R square value as 46% variations of COVID-19 new cases was explained by the model. And, the linear regression model explained 41% of the variations (Table 5 and Fig. 2).

The fitted linear regression equation is given below.

$\widehat{\text{New Cases}} = -61 + 0.034 * \text{daily laboratory tests}$. This indicated that the new cases will be raised to 3,400 if 100,000 laboratory tests were conducted daily.

Similarly, new case from Addis Ababa city was predicted significantly in the new cases in the linear regression model (p-value of 0.000) with

$R^2=93\%$ of the new cases variations were explained by cases from Addis Ababa city (Table 5 and Fig. 3).

Its estimated linear regression equation is defined as follow.

$\widehat{\text{New Cases}} = 10.8 + 1.2 * \text{new cases from AA city}$. This suggested that the country's new cases will be increased to 12,000

if 10,000 new cases were found in ADDIS ABABA city.

3.6 Multiple Linear Regression Model for COVID-19 New Cases

In this model, COVID-19 new cases were predicted significantly by the number of days, daily laboratory tests, and new cases from Addis Ababa city at 5%, 10%, and 1% levels of significance, respectively (Fig. 4).

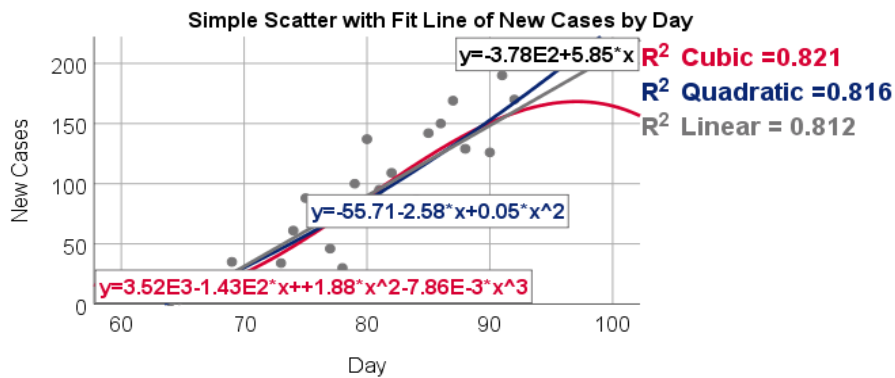


Fig. 1. Simple line graph of COVID-19 new cases by the number of days

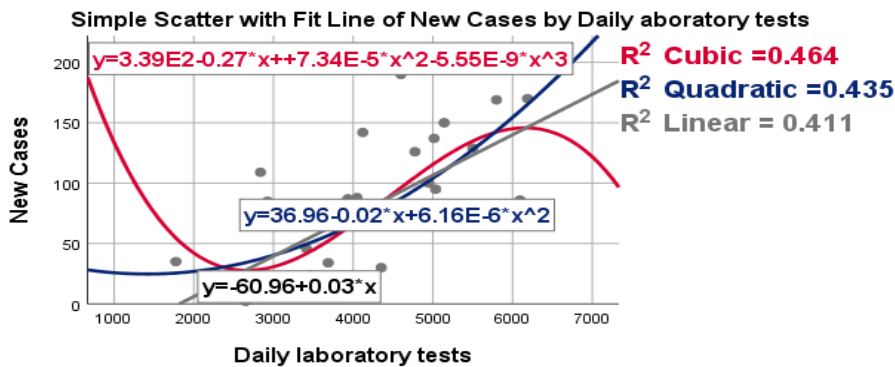


Fig. 2. Simple line graph of COVID-19 new cases by daily laboratory tests

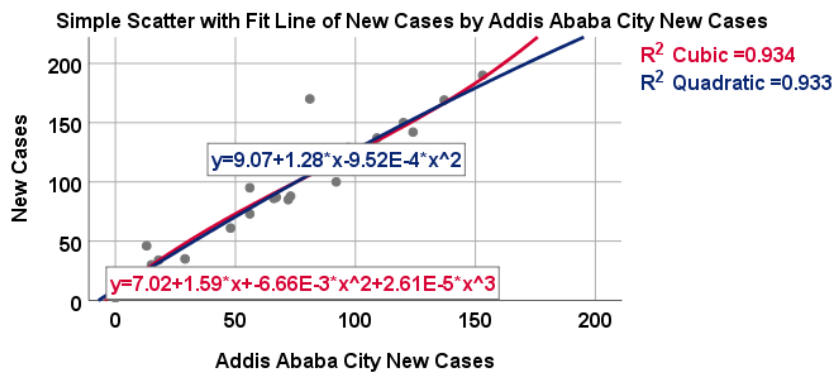


Fig. 3. Simple line graph of COVID-19 new cases by new cases from ADDIS ABABA city

Table 4. Correlation analysis between COVID-19 new cases and related variables

| | | | | | |
|--|--|---|--|-------------------------------------|---------------------------------------|
| COVID-19 New Cases (r & p values) | Number of days (0.901, 0.000*) | Laboratory tests (0.641, 0.000**) | New recoveries (0.389, 0.034**) | Males cases (0.985, 0.000**) | Females cases (0.964, 0.000**) |
| | New cases from ADDIS ABABA City (0.965, 0.000**) | New cases from Foreign Natives (0.416, 0.022) | | | |

Table 5. Regression models summaries and parameters estimates to predict COVID-19 New Cases from 12th May to 10th June in Ethiopia, 2020

| Covariates | Model | R² | F | MSE | Sig. | Constant | b1 | b2 | b3 |
|----------------------------|--------------|----------------------|----------|------------|-------------|-----------------|---------------|---------------|-----------|
| Day | Linear | 0.81 | 120.7 | 637.4 | 0.000 | -378.2(0.000)* | 5.851(0.000)* | | |
| | Quadratic | 0.82 | 59.8 | 646.3 | 0.000 | -55.7(0.894) | -2.577(0.813) | 0.054(0.440) | |
| | Cubic | 0.82 | 59.8 | 646.3 | 0.000 | -55.7(0.894) | -2.577(0.813) | 0.054(0.440) | 0 |
| Daily Lab Tests | Linear | 0.41 | 19.5 | 1993.3 | 0.000 | -61.0(0.000)* | 0.034(0.000)* | | |
| | Quadratic | 0.43 | 10.4 | 1983.8 | 0.000 | 37.0(0.707) | -0.017(0.722) | 0 | |
| | Cubic | 0.46 | 7.5 | 1952.2 | 0.001 | 338.8(0.220) | -0.273(0.223) | 0.0001(0.204) | 0 |
| Addis Ababa City New Cases | Linear | 0.93 | 382.6 | 230.8 | 0.000 | 10.8 (0.000)* | 1.2 (0.000)* | | |
| | Quadratic | 0.93 | 187.4 | 235.9 | 0.000 | 9.1(0.090) | 1.3 (0.000)* | -0.001(0.533) | |
| | Cubic | 0.93 | 122.3 | 241.2 | 0.000 | 7.02 (0.262) | 1.6(0.006)* | -0.007(0.469) | 0 |

* The regression coefficients are significant at 1%.

The fitted MLR for COVID-19 New Cases,

$$\widehat{\text{COVID-19 New Cases}} = -99.1 + 1.352 * \text{number of days} + 0.00503 * \text{daily laboratory tests} + 0.88096 * \text{new cases from AA city.}$$

- 1) COVID-19 new cases are predicted to increase 135 when the number of days increases by 100 days while holding other variables constant.
- 2) COVID-19 new cases are predicted to increase 503 when the daily laboratory tests rise by 100,000 tests while holding other variables constant.
- 3) COVID-19 new cases are predicted to increase 881 when the new cases from Addis Ababa city increase by 10,000 tests while holding other variables constant.

Also, it is predicted to be 0 (negative cases not applicable) when the three variables are zero. In the model, R²=96% of the variation in COVID-19 new cases was explained by the model (predictors). In the hypothesis test (F-statistic=205.1 with DF=3 & 26, p-value=0.000), there was enough evidence to reject the null hypothesis that all the model's coefficients are 0. The residual standard error = 12.16 shows how far the observed total COVID-19 new cases (Y-values) are from the predicted total COVID-19 new cases (Ŷ) (Fig. 4, R-software output).

3.7 Multiple Linear Regression Assumptions

The multiple linear regression assumptions were tested correctly as shown on Fig. 5.

4. DISCUSSION

In the correlation analysis for COVID-19, new cases had significant and positive correlations with the number of days (r = 0.901), daily laboratory tests (r = 0.641), new recoveries (r =0.389), new cases from males (r =0.985), new cases from females (r = 0.964), new cases from ADDIS ABABA city (r = 0.965), and new cases from foreign natives (r = 0.416).

The simple linear regression model was a better fit for the data of COVID-19 new cases than quadratic and cubic regression models. In this fitted model, COVID-19 new cases were significantly predicted by the number of days (B= 5.85), daily laboratory tests (B= 0.034), and new cases from Addis Ababa city (B=1.2) at a 5% level of significance. A study from Indian found that the linear regression growth model was more specific to predict the number of affected cases of COVID-19 than the exponential growth model. These models are used for forecasting in long term intervals. And, another study from Indian showed that the linear model was the best fitting model for Region III from May 3rd to May 15th [8-10].

```
Call:
lm(formula = COVID19NewCases ~ NumberofDays + DailyLabor.Tests +
    NewCasesfromAACity, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-24.020  -6.188  -0.421   5.624  42.233

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -99.075139   35.424355  -2.797  0.00958 **
NumberofDays    1.351870    0.571886   2.364  0.02584 *
DailyLabor.Tests  0.005029    0.002882   1.745  0.09278 .
NewCasesfromAACity  0.880959    0.090683   9.715 3.86e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 12.16 on 26 degrees of freedom
Multiple R-squared:  0.9595,    Adjusted R-squared:  0.9548
F-statistic: 205.1 on 3 and 26 DF,  p-value: < 2.2e-16
```

Fig. 4. R Output of MLR for estimating the parameter to predict COVID-19 New Cases

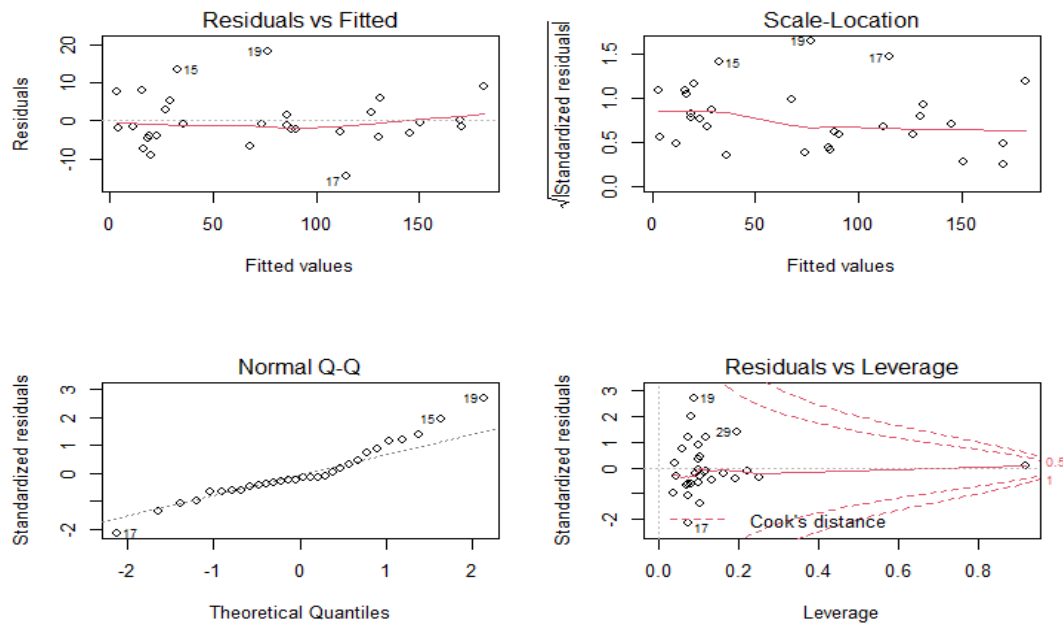


Fig. 5. R-Output of multiple linear regression assumptions

In the multiple linear regression model, COVID-19 new cases were predicted significantly by the number of days, daily laboratory tests, and new cases from Addis Ababa city at the 5%, 10%, and 1% levels of significance, respectively. Thus, COVID-19 new cases were predicted to increase 135, 503, and 881 when the number of days increased by 100 days, the daily laboratory tests increased by 100,000 tests, and the new cases from Addis Ababa city increased by 10,000 tests while holding other variables constant.

Odhiambo et al. [9] from Kenya showed that there was a correlation between COVID-19 new cases and contact persons made by the confirmed status as well as the number of flights from foreign countries to Kenya. The study used univariate analysis of the generalized linear model showed that contact persons in Kenya had 0.265 effects on COVID-19 cases in Kenya. In the multivariate analysis, the contact persons and flights to Kenya had 0.278 and 3,309 effects on COVID-19 cases in Kenya at 5% and 10% levels of significance, respectively. The researchers also used the compound Poisson regression model, which showed that as the COVID-19 day increased by 235, the COVID-19 new cases were projected to 83,418 new cases.

Ghosal et al. [11] from India used a linear regression analysis to predict the average week 5 and 6 death counts. Thus, our study agreed with this study on the correlation analysis but not on the linear regression analysis.

Mahnnty et al. [12] analyzed COVID-19 cases from India, Pakistan, Myanmar (Burma), Brazil, Italy and Germany till June 4, 2020 and predictions have been made for the number of positive cases for the next 28 days. In the study, Verhulst model fitting effect is better than Gompertz and SIR model with R-score 0.9973. The proposed model perform better as compare to other three existing models with R-score 0.9981. These above models can be adapted to forecast in long term intervals, based on the predictions for a short interval as of June 5, 2020 and June 30, 2020, active COVID-19 patients for India, Pakistan, Italy, Germany, Brazil and Myanmar predicted as (236,170, 88,998,234,066, 184,922, 645,057 and 235) and (486,357, 218,864, 240,545, 193, 727, 1,211,567 and 309).

Pandey et al. [13] from India analyzed COVID-19 cases from 30th January to 30th March 2020 and predictions were made for the number of cases for the next 2 weeks. SEIR model and Regression model were used for predictions based on the data collected from John Hopkins University repository. The performance of the models was evaluated using RMSLE and achieved 1.52 for SEIR model and 1.75 for the regression model. The RMSLE error rate between SEIR model and Regression model was found to be 2.01. Also, the value of R0 which is the spread of the disease was calculated to be 2.02. Expected cases may rise between 5000–6000 in the next two weeks of time.

Ayyoubzadeh et al. [14] from Iran used linear regression model predicted the incidence with root mean square error (RMSE) of 7.562 (SD 6.492). The most effective factors besides previous day incidence included the search frequency of hand washing, hand sanitizer, and antiseptic topics. The RMSE of the long short-term memory model was 27.187 (SD 20.705).

Rath et al. [15] used a comparison of linear regression and multiple linear regression model was performed where the score of the model tends to be 0.99 and 1.0 which indicates a strong prediction model to forecast the next coming days active cases. Using the multiple linear regression model as on July month, the forecast value of 52,290 active cases are predicted towards the next month of 15th August in India and 9,358 active cases in Odisha if situation continues like that way.

Teresa et al. [16] from Canada found a positive association but not statistically significant between cumulative incidence and ambient temperature (14.2 per 100,000 people; 95%CI: -0.60–29.0) using multiple linear regression models. The study showed that there was no a statistically significant association between total cases or effective reproductive number of COVID-19 and ambient temperature.

Researchers from South Korea Lee et al. [17] found newly confirmed COVID-19 patients have been decreasing since March 2020 while the traffic has been increasing. The study also showed that traffic was increasing indicates greater contact between people, which in turn increases the risk of further COVID-19 spread using non-linear regression and single linear regression models.

Another researcher Sansa N.A [18] from China found that there was a significant positive correlation between the COVID-19 confirmed cases and Recovered cases using a simple regression linear model from the period dated 20 January 2020 to 23 February 2020.

5. CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The total number of COVID-19 cases from May 12th to June 10th 2020 was increased by 9.1 times as compared with 14th of March to 11th May 2020 in Ethiopia. In the correlation analysis, the

COVID-19 new cases were significantly correlated with the number of days, daily laboratory tests, new recoveries, new cases of males, new cases of females, and new cases from Addis Ababa city. In the simple linear regression, variables such as the number of days, daily laboratory tests, new recoveries, new cases of males, new cases of females, new cases from Addis Ababa city, and new cases from foreign natives were significantly predicted COVID-19 new cases. But only the number of days, daily laboratory tests, and new cases from Addis Ababa city were significantly predicted COVID-19 new cases at 5%, 10%, and 1% levels of significance using the multiple linear regression model. And based the model prediction, COVID-19 new cases will be increased by 135, 503, and 881 new cases when the number of days, daily laboratory tests, and new cases from Addis Ababa city are increased by 100 days, 100,000 tests, and 10,000 cases, respectively. Then, if strong preventions and actions will not been taken in the country, the predicted values of COVID-19 new cases will be 590 after the 9th of August 2020.

5.2 Recommendation

The researchers recommended that Ethiopia government, Ministry of Health and Regional Governments (especially the Addis Ababa city administrative) should give more awareness and protections collaboratively for societies, and they should also open more COVID-19 laboratory testing health centers in different areas of the country to ensure that those health centers can test more persons as the number of days increases, and the number of new cases will be highly increased as predicted in this study. With these preventive and curative measures, the severity of COVID-19 will be limited when compared to other countries, such as the USA, South Africa, and Egypt, which are now leading in the number of new cases in the world and Africa. This research work will be extended after looking for the spread of the disease instantaneously by using a comparison of linear regression and time serious models. And, this study will help the government and doctors in preparing their plans for the next times. Based on the predictions for short-term interval, these models can be tuned for forecasting in long-term intervals.

CONSENT AND ETHICAL APPROVAL

It is not applicable.

ACKNOWLEDGEMENT

The authors thanked Ethiopia Ministry of Health Organization for sharing COVID-19 daily cases on their Facebook and telegram pages. <https://www.facebook.com/EthiopiaFMoH/>.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Wu F, et al., Pei 465 YY et al: A new coronavirus associated with human respiratory disease in 466. China. Nature. 2020;579(7798):265-269.
2. Wang Y, et al., Abnormal respiratory patterns classifier may contribute to large-scale screening of people infected with COVID-19 in an accurate and unobtrusive manner. arXiv preprint arXiv:2002.05534, 2020.
3. Hu, Z., et al., Artificial intelligence forecasting of covid-19 in china. arXiv preprint arXiv:2002.07112, 2020.
4. WHO, WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020.
5. Cases C, Worldometer. Retrieved on, 2020. 30.
6. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. Malawi medical journal, 2012;24(3):69-71.
7. Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. 2021: John Wiley & Sons.
8. Montalban EJ, et al. Correlational Analysis of COVID-19 Deaths and Recoveries per Country. Available at SSRN 3624543;2020.
9. Odhiambo JO, et al. Modelling of covid-19 transmission in kenya using compound poisson regression model. Journal of Advances in Mathematics and Computer Science. 2020;101-111.
10. Sharma VK, Nigam U. Modelling of Covid-19 cases in India using Regression and Time Series models. medRxiv;2020.
11. Ghosal S, et al. Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases-March 14th 2020). Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 2020;14(4):311-315.
12. Mahanty C, et al. Prediction of COVID- 19 active cases using exponential and non- linear growth models. Expert Systems, 2020;e12648.
13. Pandey G., et al., SEIR and Regression Model based COVID-19 outbreak predictions in India. arXiv preprint arXiv:2004.00958, 2020.
14. Ayyoubzadeh SM, et al., Predicting COVID-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. JMIR public health and surveillance. 2020;6(2):e18828.
15. Rath S, Tripathy A, Tripathy AR. Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 2020;14(5):1467-1474.
16. To T, et al. Correlation of ambient temperature and COVID-19 incidence in Canada. Science of the Total Environment. 2021;750:141484.
17. Lee H., et al., The relationship between trends in COVID-19 prevalence and traffic levels in South Korea. International Journal of Infectious Diseases. 2020;96:399-407.
18. Sansa NA. The Correlation between COVID-19 Confirmed and Recovered Cases in China: Simple Regression Linear Model Evidence. Electronic Research Journal of Social Sciences and Humanities. 2020;2.

© 2021 Argawu et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here:
<https://www.sdiarticle4.com/review-history/76078>