Introduction

The outbreak of COVID-19 has developed into a pandemic where the confirmed cases reached more than 114 million, including more than 2 million deaths reported to the World Health Organization (WHO) till the 5th of March 2021. Clinical presentation of COVID-19 range from symptomless carriers to critical disease with acute hypoxemic respiratory failure, sepsis, or multiple system failure. Around 80% of COVID-19 patients present with a mild illness that usually recover within 2 weeks. The severe disease requiring intensive care (ICU) was nearly encountered in 5% of COVID-19 patients, and 20% of COVID-19 hospitalized patients. Hospital death rates related to COVID-19 are 15% to 20% but, rise to 40% among intensive care patients. Therefore, non-subjective risk assessment of patients is fundamental for early patient management and better medical resource allocation. For most non-severe patients, management includes...
quarantine with symptomatic treatment at home, while for severe patients, urgent referral to tertiary care is needed.3

Phone triage is critical for identification of the vulnerable and high-risk patients, as emergency department’s (ED) overcapacity and delay in treatment in many conditions are correlated with increased mortality.4 Although phone triage helps to identify patients with high-risk, it has limitations related to the providers’ judgment of critical care needs,5 providers’ variations, and patients’ subjective information.5,7

Using machine learning (ML) could be considered a suitable solution. ML is a branch of artificial intelligence that learns from past data to build a prognostic model without the need for further programming8,9 During the last years, it has been developed as a useful tool to analyze large amounts of data from medical records or images.10

Advancements in computing abilities and data availability granted deep learning to be incorporated in many medical arena purposes. Deep learning and neural network have been effectively used in the diagnosis of dermatological tumors and prostate cancer.11 Regarding COVID-19, several modeling studies focused on forecasting the potential international spread of COVID-19.12 Also, deep learning has the potential to guide more efficient triage decisions, risk categorization, and resource allocation.13

Being one of the largest public university hospitals in Egypt; Cairo University Hospitals provide health care services for thousands of patients including all workers in Cairo University. During the pandemic, the family medicine department was in charge of the phone triage system for COVID-19 patients. This study aimed to assess the accuracy of the traditional phone triage system and the phone triage-driven deep learning model in the prediction of positive COVID-19 patients.

Patients/Methods

This is a retrospective study conducted at the family medicine department, Cairo University to detect the accuracy of the traditional phone triage system and phone triage-driven deep learning model in the prediction of positive COVID-19 patients. The study included a dataset from the phone triage system during the first wave of the pandemic in the period between June 1, 2020, and July 18, 2020.

In response to the COVID-19 pandemic, Cairo University hospitals assigned an interactive voice response (IVR) number for phone triaging COVID-19 patients. The service started on the first of June 2020 and received calls from all workers at Cairo University. The phone triage system was run by 20 skilled family medicine physicians and supervised by 3 family medicine consultants. The patients were assessed according to the Egyptian Ministry of Health checklist for acute respiratory illness.13 Accordingly, the patients were classified into urgent cases, suspected cases, or unsuspected cases. According to the protocol of the infection control unit in the hospital, urgent cases underwent the nasopharyngeal swab for COVID-19 diagnosis within 12 to 24h and suspected cases underwent the nasopharyngeal swab within 24 to 48 h. Positive cases received suitable management according to Cairo University hospital’s protocol for COVID-19 management. The unsuspected cases received an explanation of their condition with reassurance and hence, no need to perform a nasopharyngeal swab.

A Microsoft form prepared to collect the data of the suspected COVID-19 patients consisted of 16 questions including:

- Sociodemographic-related variables included name, age, occupation, phone number, and national ID.
- COVID-19 infection-related variables included body temperature over 38°C, severe or persistent cough, severe congestion in the throat, vomiting, diarrhea, fatigue, loss of appetite or muscle pain, and loss of sense of smell or taste.
- Risk factors and comorbidities variables included contact with an acute respiratory infection case, visiting a place or health sector where a positive COVID-19 case was discovered, working in the health sector or isolation hospital, smoking status, and having chronic diseases.

From June 1st, 2020 to July 18th, 2020, the phone triage received 3654 calls through the (IVR) number of Cairo University hospitals. A total of 943 callers were suspected of having COVID-19 disease after the telephone-based screening and triage. They were referred for nasopharyngeal swab and PCR. In addition, the phone triage system received 1211 follow-up calls.

The dataset for real-time polymerase chain reaction (RT-PCR) results was retrieved and checked for completeness. Complete data for 440 participants out of the 943 suspected COVID-19 patients were enrolled in the deep learning model. We prepared a PCR-dependent and phone triage-driven deep learning model for automated classifications of natural human responses collected by the phone triage system. This was dependent on the extensive analysis of the collected dataset of 440 participants with PCR results available. This deep learning model consisted of 3 stages, that is, the preprocessing stage which aimed at performing feature-wise normalization of input and turning integer categorical features into a one-hot vector, then the feature extraction stage which aimed at converting the preprocessed phone responses into PCR-dependent features, and the classification stage which aimed at classifying the PCR-dependent features extracted from the previous stage.

A structural illustration of the proposed deep learning model is given in (Figure 1). As seen, the preprocessing
Figure 1. Structure illustration of the deep learning model.
stage took the raw responses as inputs and fed its output to a sequence of 3 layers, each consisting of 4 operations, including Dense, Batch Normalization, LeakyReLU, and Dropout operation. This sequence of layers was used to extract class-dependent features and fed them to the classification stages which utilized the Softmax activation function. The Softmax converted the extracted feature vector to a vector of 2 categorical probabilities, namely positive PCR, and negative PCR.

**Ethical Considerations**

Ethical approvals were obtained from the Research Ethics Committee at Kasralainy school of medicine and the research committee of the Family medicine department at Cairo University.

**Statistical Analysis of Data**

Statistical analysis was done using Stata software (version 16). All categorical variables were presented in numbers and percentages, while numerical variables in mean and standard deviation (SD). Accuracy testing was done for the symptoms, deep learning model, and phone triage protocol compared to RT-PCR results. We analyzed specificity, sensitivity, area under the receiver operating characteristic (ROC) curve, positive predictive value, and negative predictive value.

**Results**

For the 943 suspected COVID-19 patients by the phone triage, the mean age (±SD) was 41.42 (±14.6), and 55.9% were females.

Regarding the risk factors and comorbidities of participants, nearly half of them were healthcare workers. More than two-thirds (77.7%) were contacts of confirmed COVID-19 cases. Forty-seven percent of the participants visited places with confirmed cases. About 10% were smokers. Only 2.3% of females were pregnant. Three percent of participants were immunocompromised. More than 33% suffered from chronic diseases such as hypertension (14%), diabetes (11.5%), pulmonary diseases (6%), cardiovascular diseases (4.7%), chronic liver disease (1.4%), chronic kidney disease (1.3%), cancer (1.17%), and other chronic diseases (3.4%), as shown in Table 1.

Regarding the most common prevalent symptoms recorded by the COVID-19 suspected patients, myalgia was the most common recorded symptom (84%) followed by fever (73%) (Figure 2). Based on the RT-PCR results, we found that myalgia, fever, and contact with a case with respiratory symptoms had the highest sensitivity among the symptoms/risk factors that were asked during the phone calls (86.3%, 77.5%, and 75.1%, respectively).

| Table 1. Participant’s Characteristics and Risk Factors (n=943). |
| --- | --- | --- |
| | Yes | No |
| n (%) | n (%) |
| **Sex** | | |
| Female | 528 (55.99) | - |
| Male | 415 (44.01) | - |
| **Job** | | |
| Employee | 557 (61.55) | - |
| Faculty member | 102 (11.27) | - |
| Nurse | 37 (4.09) | - |
| Relative | 194 (21.44) | - |
| Resident | 15 (1.65) | - |
| **Contacted a case with respiratory symptoms** | 733 (77.73) | 210 (22.27) |
| **Visited a place with COVID-19 case** | 443 (46.98) | 500 (53.02) |
| **Working in healthcare or isolation area** | 441 (46.77) | 502 (53.23) |
| **Smoking** | 87 (9.88) | 794 (90.12) |
| **Pregnancy** | 12 (2.28) | 515 (97.72) |
| **Immunodeficiency diseases or drugs** | 29 (3.08) | 914 (96.92) |
| **Presence of comorbidities** | 313 (33.19) | 630 (66.81) |
| **Hypertension** | 136 (14.42) | 809 (85.58) |
| **Diabetes mellitus** | 109 (11.56) | 834 (88.44) |
| **Cardiovascular diseases** | 44 (4.67) | 899 (95.33) |
| **Chronic liver disease** | 13 (1.38) | 930 (98.62) |
| **Pulmonary disease** | 58 (6.15) | 885 (93.85) |
| **Chronic kidney disease** | 12 (1.27) | 931 (98.73) |
| **Cancer** | 11 (1.17) | 932 (98.83) |
| **Others** | 32 (3.39) | 911 (96.61) |

**Figure 2.** Frequency distribution of different recorded symptoms among COVID-19 suspected patients.

While immunodeficiency, smoking, and loss of smell or taste had the highest specificity (96.9%, 83.6%, and 74.0%, respectively) as revealed in Table 2.

The total retrieved PCR results were 440 cases: 213 of them were positive for COVID-19. The Positive predictive
value of phone triage was 48.4%. The classification and prediction accuracy achieved by the PCR-dependent and phone triage-driven deep learning model was 66%. While the positive predictive value was (70.5%) with a sensitivity and specificity of 67.4% and 63.9%, respectively (Table 2).

**Discussion**

The worrying trend of rapid transmission, sudden progression to critical illness, and death accompanied by the COVID-19 pandemic in presence of a broad spectrum of clinical presentations and non-specific initial symptoms of the infection led to a sharp increase in the numbers of patients seeking advice in hospitals and medical centers. So, there is an urgent need for a way that can help the community to overcome this epidemic and can triage the patients with maintaining the social distance.

Remote triage has many benefits; it reduces the workload on the medical team, reduces the travel burden, and improves resource utilization of emergency department services. Within the current pandemic, it helps to apply social distancing which is the main preventive measure to limit the spread of COVID-19.\(^1\)

In our experience, the family physicians (FPs) were able to provide effective phone triage, remote screening, and referral services. This triage is based on the symptoms such as epidemic screenings. We faced many calls unrelated to triage (to know clinics’ schedules, obtain test results, ask questions about medications or symptoms, and sick leaves). So, we continuously had to increase the number of physicians per shift from 4 to 8 physicians to cope with increasing calls. This led to reducing medical workload, and infection exposure and gave chance for patients in need to be properly managed.

Based on RT-PCR results of suspected patients on phone triage, it was found that more than half of those who did not complain of myalgia were negative for COVID-19 infection. It yields that the highest sensitivity and the highest negative predictive value were for myalgia among all the recorded symptoms. This high negative predictive value was also revealed in other studies reaching up to 80% in the study of Clemency et al.\(^14\)

Despite being the most specific, loss of smell and/or taste was recorded only in one-third of cases with the least sensitivity among all the recorded symptoms, as anosmia and ageusia commonly appear late in the course of the illness. That goes hand in hand with the systematic review which assessed the diagnostic accuracy of signs and symptoms to determine if a person presenting to the clinical care setting has COVID-19. It showed that anosmia had a pooled sensitivity of 28.0% and a specificity of 93.4% while ageusia had a pooled sensitivity of 24.8% and a specificity of 91.4%.\(^15\)

Despite being the second least common symptom, vomiting and diarrhea demonstrated high specificity second to loss of smell and/or taste. Clemency et al also revealed high specificity (74%) related to diarrhea.\(^14\)

(RT-PCR)—the validated diagnosis of SARS-CoV-2 infection—is expensive, and not easy to be conducted in all health services to find a proper and quick method for screening the patient before referral for PCR.\(^16-18\)

In this study, out of the total 440 suspected cases sent for RT-PCR testing through the phone triage service, almost half showed positive results for COVID-19 infection yielding a positive predictive value of 48.4% for the phone triage accuracy. While the use of RT-PCR-dependent and phone triage-driven deep learning model had been shown a positive predictive value of (70.5%). The accuracy achieved

### Table 2. Accuracy and Predictive Values of COVID19 Symptoms, Phone Triage, and Deep Learning Model.

| Symptom                                | Sensitivity (%) | Specificity (%) | AUC  | PPV (%) | NPV (%) |
|----------------------------------------|----------------|----------------|------|---------|---------|
| Contacted a case with respiratory symptoms | 75.1           | 15.9           | 0.46 | 45.7    | 40.4    |
| Visited a place with COVID-19 case     | 45.2           | 43.6           | 0.44 | 42.8    | 46.0    |
| Working in healthcare or isolation area | 48.1           | 45.8           | 0.47 | 45.3    | 48.6    |
| Fever                                  | 77.5           | 25.3           | 0.51 | 49.5    | 54.3    |
| Cough                                  | 65.1           | 34.1           | 0.50 | 48.1    | 51.0    |
| Sore throat                            | 56.1           | 32.3           | 0.44 | 43.8    | 44.0    |
| Vomiting or diarrhea                   | 40.0           | 54.2           | 0.47 | 44.7    | 49.4    |
| Myalgia                                | 86.3           | 18.1           | 0.52 | 49.5    | 58.6    |
| Loss of smell or taste                 | 36.2           | 74.0           | 0.55 | 56.6    | 55.3    |
| Smoking                                | 8.3            | 83.6           | 0.46 | 32.7    | 48.9    |
| Comorbidities                          | 37.6           | 67.8           | 0.53 | 52.3    | 53.7    |
| Immunodeficiency                       | 4.3            | 96.9           | 0.51 | 56.3    | 51.9    |
| Deep learning model                    | 67.4           | 63.9           | 0.66 | 70.5    | 60.5    |
| Phone triage service                   | 48.4           |                |      |         |         |

Abbreviations: AUC, area under the receiver operating characteristic (ROC) curve; PPV, positive predictive value; NPV, negative predictive value.
through this model provides a base for larger studies and clinical implementations as it focuses on fundamental clinical evidence that is present in the early stage of the disease. The study achieved its proposed aim to digitalize the patient histories received from phone triage using a patient history detailed checklist. Similar promising results have been presented in other studies as the study of Zoabi et al.19 and Langer et al.20

The model automates triage with machine learning (ML) techniques and categorizes the patients through PCR-dependent and phone triage-driven deep learning to minimize workloads for primary care. To produce reliable predictive algorithms, AI requires large quantities of processing data.21 Large amounts of data might not be available during the initial phases of outbreaks, which is possibly the period when the prediction is most needed. In addition to the amount of data, it is especially impossible to receive high-quality data. Several researchers have used special data forms to perform AI-driven functions, such as using only radiological images to diagnose COVID-19 example.22,23 Thus, Attempts are required to move AI dependency from costly techniques to cheaper and more easily available alternatives, such as chest X-rays, or just statistics on clinical symptoms and vital signs alone.24

Limitations and Strengths
To our knowledge to the current date few studies, if any, have compared the accuracy of phone triage to that of deep learning models in the diagnosis of positive COVID-19 infection. Also, compared to other similar studies, the current study has included a relatively large number of cases. One of the limitations of the current study was the inability to calculate sensitivity or specificity values for the phone triage results since not all the PCR testing results were available.

The Implication for Practice and Future Research
Finally using the deep learning models for symptoms screening will help to save the limited clinical resources for those who need more precise clinical judgment and follow up and will help to provide the proper medical care as early as possible for those at a higher risk of developing severe illness paving the way for a more efficient allocation of the scanty health resources.

Using the deep learning models in making diagnostic decisions may miss the clinical judgment in patients with multiple risk factors or those with a history of chronic pulmonary diseases whose symptoms are similar to that of patients with COVID-19 infection increasing the percentage of false-positive results. So, future research may be needed to detect its accuracy in these patients.

Conclusion
Phone triage driven by a family physician can be effective in decreasing hospital visits, workloads, and ultimately COVID-19 infection exposure. Deep learning is a promising backup method for phone triage in screening COVID suspected patients. It has good accuracy that can help efficient triage decisions and resource allocation.

Acknowledgments
The authors would like to acknowledge all members of the Family Medicine department, Cairo university, for their great effort at the phone triage service during the first wave of the COVID-19 pandemic especially; Dr. Fatma Fathy and Dr. Lamiaa Adel, assistant lecturers of Family Medicine, Family Medicine department, Cairo University and our dedicated residents; Dr. Sara Ahmed, Dr. Salma Zedan, Dr. Radwa Barakat, Dr. Rania Ahmed, Dr. Somaya Nasser, Dr. Nesma Mostafa, and Dr. Meral Ahmed.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs
Marwa M. Ahmed https://orcid.org/0000-0002-5461-7861
Marwa D. Hasan https://orcid.org/0000-0003-0422-109X
Mai D. Sarhan https://orcid.org/0000-0002-8445-1970
Radwa M. Elsayed https://orcid.org/0000-0002-3219-3703
Asmaa A. Sayed https://orcid.org/0000-0001-7324-5736
Rehab M. Yassen https://orcid.org/0000-0002-0991-8032
Hala A. Hussein https://orcid.org/0000-0001-6357-7892
Dalia M. Gaber https://orcid.org/0000-0002-7847-8886
Samar Fares https://orcid.org/0000-0002-3438-1329

Data Availability
The source code is publicly available for downloading at https://github.com/ahfares/Phone-triage-driven-deep-learning-model

References
1. WHO. COVID-19 Dashboard. Accessed March 5, 2021. https://www.google.com.eg/url?sa=t&rct=j&q=&esrc=s&source=w eb&cd=&cad=rja&uact=8&ved=2ahUKEwiJ37iJ3ZjvAhX MEcAKHW6jCjEQFJAegQlEXeAE&url=http%3A%2F%2F covid19.who.int%2F&usg=AOvVaw22XfLm9xex_cPbXQ-q5vKUr
2. WHO. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). 2020. Accessed March 7, 2021. https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf
Ahmed et al

3. Wu G, Yang P, Xie Y, et al. Development of a clinical decision support system for severity risk prediction and triage of COVID-19 patients at hospital admission: an international multicentre study. *Eur Respir J*. 2020;56(2):56.

4. Bernstein SL, Aronsky D, Duseja R, et al. The effect of emergency department crowding on clinically oriented outcomes. *Acad Emerg Med*. 2009;16(1):1-10.

5. Christ M, Grossmann F, Winter D, Bingisser R, Platz E. Modern triage in the emergency department. *Dtsch Arztebl Int*. 2010;107(50):892-898.

6. Farrohknia N, Castrén M, Ehrenberg A, et al. Emergency department triage scales and their components: a systematic review of the scientific evidence. *Scand J Trauma Resusc Emerg Med*. 2011;19(1):42.

7. Welch S, Davidson S. Exploring new intake models for the emergency department. *Am J Med Qual*. 2010;25(3):172-180.

8. Kwon JM, Lee Y, Lee Y, Lee S, Park H, Park J. Validation of deep-learning-based triage and acuity score using a large national dataset. *PLoS One*. 2018;13(10):e0205836.

9. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380(14):1347-1358.

10. van der Schaar M, Alaa AM, Floto A, et al. How artificial intelligence and machine learning can help healthcare systems respond to COVID-19. *Mach Learn*. 2021;110(1):1-14.

11. Jiang X, Coffee M, Bari A, et al. Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. *Comput Mater Contin*. 2020;62(3):537-551.

12. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*. 2020;395(10225):689-697.

13. Masoud HH, Elaissal G, Zaky S, et al. Management protocol for COVID-19 patients version 1.4/30th May 2020 Ministry of health and population (MOHP), Egypt. In: *Coronavirus Disease 2019 (COVID-19), SARS COV2 management guideline*. Ministry of Health and Population. Accessed May 29, 2021. http://www.mohp.gov.eg/JobsDetails.aspx?job_id=3061

14. Clemency BM, Varughese R, Scheafer DK, et al. Symptom criteria for COVID-19 testing of heath care workers. *Acad Emerg Med*. 2020;27(6):469-474.

15. Struyf T, Deeks JJ, Dinnes J, et al. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. *Cochrane Database Syst Rev*. 2022;5(7):CD013665.

16. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus--infected pneumonia. *N Engl J Med*. 2020;382(13):1199-1207.

17. Guo L, Ren L, Yang S, et al. Profiling early humoral response to diagnose novel Coronavirus disease (COVID-19). *Clin Infect Dis*. 2020;71(15):778-785.

18. Lippi G, Simundic AM, Plebani M. Potential preanalytical and analytical vulnerabilities in the laboratory diagnosis of coronavirus disease 2019 (COVID-19). *Clin Chem Lab Med*. 2020;58(7):1070-1076.

19. Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *npj Dig Med*. 2021;4(1):3-5.

20. Langer T, Favaro M, Giudici R, et al. Development of machine learning models to predict RT-PCR results for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in patients with influenza-like symptoms using only basic clinical data. *Scand J Trauma Resusc Emerg Med*. 2020;28(1):113.

21. Santosh KC. AI-Driven tools for Coronavirus outbreak: need of active learning and cross-population train/Test models on multidimensional/multimodal data. *J Med Syst*. 2020;44(5):93.

22. Li D, Wang D, Dong J, et al. False-negative results of real-time reverse-transcriptase polymerase chain reaction for severe Acute Respiratory Syndrome Coronavirus 2: role of deep-learning-based CT diagnosis and insights from two cases. *Korean J Radiol*. 2020;21(4):505-508.

23. Li L, Qin L, Xu Z, et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology*. 2020;296(2):E65-E71.

24. Maru DS, Schwarz R, Jason A, Basu S, Sharma A, Moore C. Turning a blind eye: the mobilization of radiology services in resource-poor regions. *Global Health*. 2010;6(1):18.