EasyChair Preprint

№ 9937

COVIData: a Web Platform for Tracking, Classification and Monitoring Cases Suspects of COVID-19

Beatriz Lima Gandolfi, Clarissa S. R. Merino, Vitor Inacio da Silva, Diego S. Costa, Gabriel de M. Fiali, Andre S. Carvalheiro, Luiz Rodrigo C. da Silva, Camila C. Rocha, Giovanna B. Lins, Saul de Castro Leite and Fernanda Nascimento Almeida

April 6, 2023

COVIData: a web platform for tracking, classification and monitoring cases suspects of COVID-19

Beatriz L. Gandolfi[1[0009−0002−2111−0471]], Clarissa S. R. Merino[1[0009−0000−5418−1837]] Vitor I. da Silva[2[0009−0001−1736−2521]], Diego S. Costa[2[0009−0008−0233−1029]], Gabriel de M. Fiali[2[0009−0004−5714−0095]], André S. Carvalheiro[1[0009−0008−9415−2131]], Luiz R. C. da Silva[2[0009−0001−7749−7655]], Camila C. Rocha[2[0009−0001−7676−458X]], Giovanna B. Lins[3[0000−0003−2600−9225]], Saul C. Leite[2[0000−0002−1793−2280]], and Fernanda N. Almeida[1[0000−0003−3086−454X]]

[1] Center for Engeneering, Modeling and Applied Social Sciences, Federal University of ABC, São Bernardo do Campo, Brazil
[2] Center for Mathematics, Computing and Cognition, Federal University of ABC, Santo André, Brazil
[3] Center for Natural and Human Sciences, Federal University of ABC, Santo André, Brazil
fernanda.almeida@ufabc.edu.br

Abstract. The high speed spread of SARS-CoV-2 trough out the entire globe has ignited the warning about the importance and need to collect data from patients possibly infected, on a massive scale, in order to understand spreading dynamics of this particular disease. Having stated that, this article aims to present the entire development of the web platform COVIData, created by students and researches of the Federal University of ABC (UFABC) while in partnership with Inter-municipal Consortium of ABC, resulting in a tool directed to people' self-screening their symptoms and being able to have and immediate identification of whether they are or not possibly infected.The tool consists of a detailed questionnaire based on scientific data on the most common symptoms of the disease. The questionnaire has been validated by healthcare professionals to verify the correlation of symptoms described by individuals using COVIData with SARS-CoV-2 infection. Furthermore, as a result, data analysis may be made and enhanced viewing the possibility to discuss, develop and implement public policies that help facing the disease.

Keywords: COVID-19, self-screening, health monitoring, public health, decision support system, database

1 Introduction

Since the coronavirus outbreak in China, late 2019, to the present time (August 2022), more than 593 million people have been reported as infected with SARS-CoV-2 around the globe [13]. The sudden spread of the disease has struck up

a global alert about the need to understand its different manifestation forms, as well as the many spreading ways it has and is transmitted through society, bearing in mind not only to detect the entire behavior spectrum of the new virus, but also the need for public health and public management to make decisions and face the challenges and consequences it shall up-bring by facing this pandemic.

Having lasted over 02 years and being something population has to deal with on a daily basis, meanwhile going back to their "normal lives", COVID-19 is still a topic on demand and many studies are still carried on in this regard in order to make sure we will not face its pandemic status once again.

Once the panorama of increasing cases accelerating over time, the concern with under-reportation of cases was born, especially in places with lack of information access, such as not so much developed countries or even rural or low income areas. These factors revealed the huge challenge of lack of communication spread which lead to lack of knowledge, from these citizens, on the pandemic itself and also on prevention methods, vaccination and follow up in cases of infection.

Contemplating the previously stated facts, on regards of finding ways to identify possible infected and contaminating patients without conducting clinical examinations has become an urgent and extremely important matter so that one may to try to curb the geographical spread of the virus and to advise these patients to perform proper social isolation and take compatible hygiene measures. Moreover, once the self-assessment, to be further explained, is performed and gives as a result a possible positive to COVID-19 infection, the individual receives indications of what to do and what not to do next, avoiding the spread of COVID-19.

One way to identify these possible cases of COVID-19 is to offer tools that allow screening the symptoms experienced by the population, as an attempt to avoid the fast-paced spread of this disease, especially given that tests and masks are not mandatory for most situations and not all are vaccinated still.

In addition to being possible to classify people as suspects based on protocols of the World Health Organization and the Health Ministry of Brazil, tools such as the one presented may also offer the possibility to collect quality information in real-time about the characteristics of the population such as age, gender and body mass index (BMI), their pre-existing diseases, geographical location and how the disease manifests itself in different ways according to these factors.

Once we are able to track and have these quality data in hand, it is possible to generate solid insights into the disease behavior in a predetermined region to boost government measures to tackle the disease and its uncontrolled spread in the population, which can lead to the collapse of the health system resulting in deaths due to lack of adequate care.

Last, with the advances of social media usage and podcast, we were able to share useful and valuable information, even fighting back the so called "Fake News" (false or misleading information) and spread topics that were relevant for the population including: vaccination, health safety measures, mental health

helping tools as well as UFABC updates on returning to in person classes and also community updates and news.

Many initiatives addressing the mechanisms previously mentioned have been launched all over the world, such as the CDC screening system in partnership with Apple Inc. [5] and the application 'Corona Virus - SUS' [16] concomitant with the Brazilian framework are solid examples. Nonetheless, both tools are exclusively dedicated to doing a self-assessment of symptoms and do not expand its functionalities to analyze the community transmission of the virus in the regions where the screenings are carried out and the symptoms particularities.

Thus, the COVIData web platform is a tool for self-screening the population, easy to use and access, estimating a possible infection and suggesting next steps to be taken, also helping the debate, development and implementation of public policies that help to face the pandemic we are experiencing.

The platform itself aims to reach the public with the report of symptoms of COVID-19, in real time, through a detailed questionnaire, based on scientific knowledge and on the most common symptoms of the disease. This questionnaire was validated by world health professionals. The platform allowing the tracking of the geographic location of identified as symptomatic individuals and helping the early identification of cases, leading to help in blocking the spread of the disease.

In addition, through social media and podcast, the platform seeks constant communication with the scientific community and society, aiming at broad knowledge and dissemination of true information about the pandemic and the like. The COVIData emerged amid concerned within the notification and monitoring of suspected cases of COVID-19 in the ABC Paulista area (metropolitan region of São Paulo).

Efforts were established addressing the creation of a simple and intuitive web application that could be accessed in any location from any personal device, being those computers or mobiles. The web platform is concerned and aiming to identify potentially infected citizens using a self-assessment of patients' symptoms and pre-existing health conditions, also tracking geographically people who have already perform the screening. By these means, it is possible to create a database with all the information provided by users to understand how the disease spreads and manifests itself in the ABC population.

2 Methods and Approach

2.1 Development of the Screening Method

At first sight, we foccus on creating a web platform which counts with a complete screening flowchart that understood symptoms that are being recently identified in clinical research around the world, updated and evidence-based. The questionnaire - result of the flowchart - was applied through an electronic web form, hosted on the UFABC web server. The user, before being directed to the platform's questionnaire, must accept to participate in the survey (the acceptance

form was also made available electronically). The Terms of Use of the tracing tool was registred in the Ethics Committee's with the protocol number CAEE: 33797120.4.0000.5594.

During the period in which the web platform was implemented and given the lack of knowledge about the dissemination aspects of the SARS-CoV-2 virus, analyzing and understanding the behavior of this virus in clinical cases of infected patients was followed by several studies [18, 3]. Through these studies and the follow-up of cases, we were able to notice as studies on clinical cases are treated segmentally around the world, clinical manifestations may vary according to the group of patients analyzed.

In spite of the fact that the most well-known symptoms of the coronavirus infection being dry cough and shortness of breath associated with fever, studies indicate that other symptoms such as diarrhea, headaches, and myalgia [20, 1, 8] are quite common in COVID-19 patients, and may as well go unnoticed or mistaken as symptoms for different diseases other than COVID-19, for example, flu.

Another study presented by the University of Mons has shown that a relevant part of the individuals afflicted by loss of smell and taste [9, 10]. Others works published alongside Europe and China also have shown mental confusion and migraine [19, 11, 15], and dermatological manifestations [12, 17] as possible symptoms of COVID-19.

In addition to symptoms that manifest with the disease, variables like how characteristics gender, age and comorbidities, for example, diabetes, cardiovascular diseases, immunosuppressive diseases are related to the number of infected and coronavirus results are also discussed [21, 4].

Using the information contained in these studies, as well as the COVID-19 management protocol of the Health Ministry of Brazil [14], a flowchart was developed (Fig. 1), consisting in questions about the most widespread symptoms about the coronavirus. This symptoms are similiar as a respiratory syndrome and were also related with symptoms that were identified by researchers as previously cited. We wera also include in the COVIData the pre-existing illnesses and personal characteristics (Fig. 1).

The workflows were built based on a scoring scheme. These score has been added to each question marked as 'yes' by the user. The final sum of the screening questions score served to classify the user as 'not suspicious', 'mild suspicious', 'average suspicious', 'high suspicious'. In those cases, in which the suspicion of COVID-19 infection is confirmed, the website informs user important recommendations on social isolation, hygiene and guidelines for medical advice seeking and provides the address with telephone of the health centers and hospitals closest to the user for assistance. As the screening tool uses the user's geographic location, it was possible to indicate the nearest hospitals and health centers to the user.

An utterly important fact that COVIData is not intended to act as a diagnosis web platform. The result of the questionnaire indicates a suggestion about the possibility of SARS-CoV-2 virus infection.

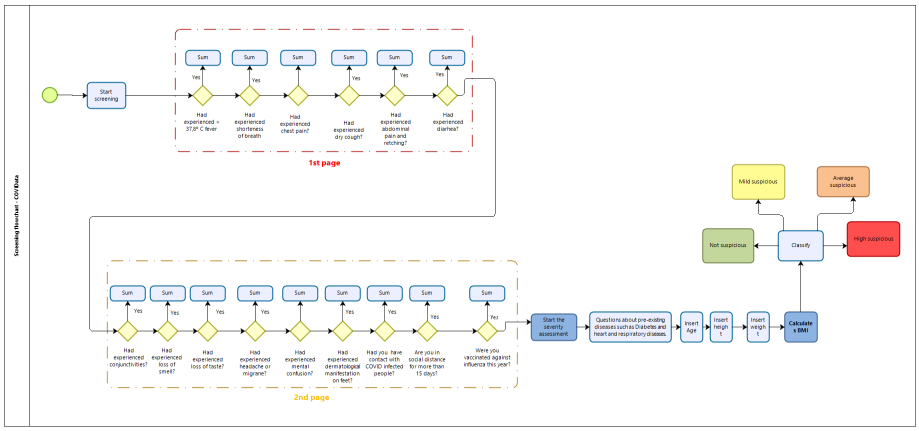

Fig. 1: COVIData simplified screening workflow.

2.2 Partnership and Additional Features

The implementation of the platform made it possible to generate several public health and population management insights, so the initiative started at the Federal University of ABC could also count on a partnership with the Intermunicipal Consortium of ABC, which is currently responsible for planning and articulating regional actions in seven cities of the metropolitan region of São Paulo, Brazil. Relying on this partnership, it was possible to direct the geographical analysis of the geo-location data captured through the IP and GPS of the devices used to access the platform combined to clinical data obtained to positively impact the area close to the UFABC and provide information to supplement the updating of the number of COVID-19 in these cities.

What is more, in addition to the screening functionality, once the user is classified as suspect, the tool provides a registration tab that is customized according to the ABC Inter-Municipal Consortium membership cities' data of interest. From this registration, it is possible to drive services in collaboration with this municipalities and their Health Departments. Were developed, 3 functional fronts: registration for tele-monitoring carried out by the health departments, registration to address the patient for clinical testing in partnership with the Federal University of ABC, and a personalized study on the prevalence of the disease in one of the consortium's member city.

On the first registration front that directs patients for tele-monitoring, those who had been classified as 'suspects by SARS-CoV-2 infection' have the possibility to register. When the user requests registration, the following data is requested: address, phone number, ethnic group, occupation and other information that may be important for the development of public policies. In this case, city halls and health departments are responsible by the tele-monitoring structure. The data collected in these tele-consultation will return to the COVI-

Data platform, to create more robust information about population experiencing coronavirus symptoms.

2.3 Information Spreading

Since the beginning of the pandemic, there has been an increase in the spread of so-called fake news, which not only led to the alienation of the population, but also to a poor management of the pandemic and aggravation of the unfavorable situation in which we find ourselves numerous times with the collapse of the health system.

People who didn't understand well what the virus was, if there really was a virus, how to prevent it, what treatments or measures to take when infected population, not always with great accuracy or veracity.

Thus, it was decided to create a communication and social media group on the platform, in order to search for new relevant information about COVID-19 and the like (variants, vaccination, tests, etc.) and transmit it in a simple and cohesive way to as many people as possible.

Thus, this group was responsible for creating the social networks of COVIData UFABC (Instagram and Facebook), in addition to communicating with the platform's followers through their message portals and institutional email.

After the creation of social media, the issue regarding the inclusion of their users was raised. Thus, we started to insert tools such and enabling the option of reading text and comments on Facebook and Instagram, avoiding the use of special characters, in addition to creating a Podcast with monthly updates of the project in which the topics covered in the publications are taken spoken to the population.

Finally, with the unexpected extension of the pandemic for more than a year, issues of Mental Health and Well-Being were raised, leading to the prioritization of publications and interaction with the public aimed at this point.

3 The COVIData Implementation

The Web platform COVIData was developed using mostly *JavaScript* often abbreviated JS. This language is a high-level programming language primarily designed to run in browsers and manipulate web page behaviors. The JS is one of the most important technologies aimed at the front-end and, joining the trio HTML, CSS and PHP, they form a group of languages that cover practically all the requirements of the development of a complete, dynamic and with good performance page. With its scripts it is possible to include, in a static page, dynamic elements such as maps, forms, numerical operations, animations, interactive infographics and much more.

The JS frameworks we used for development were different for the frontend and backend. For the frontend, React JS was used. The Front-end is closely related to the graphical interface of the project. That is, it is where the application

is developed with which the user will interact directly, whether in software, websites, applications, etc. Therefore, it is essential that the developer has a concern for the user experience.

React (also known as React.js or ReactJS) is a free and open-source front-end JavaScript library for building user interfaces based on UI components. It is maintained by Meta (formerly Facebook) and a community of individual developers and companies. React can be used as a basis in the development of single-page, mobile, or server-rendered applications with frameworks like Next.js. However, React is only concerned with state management and rendering that state to the DOM, so creating React applications usually requires the use of additional libraries for routing, as well as certain client-side functionality.

For backend, we used NodeJS with Express. The backend is the structure that enables the system to operate, while the front-end is responsible for the visual part, such as presentation, design, languages, colors, among others. Even though they have different roles, these applications are closely linked so that electronic environments operate in sync.

Node.js can be defined as a server-side Javascript execution environment. This means that with Node.js it is possible to create Javascript applications to run as a standalone application on a machine, not depending on a browser for execution, as we are used to. Despite being recent, Node.js is already used by large companies in the technology market, such as Netflix, Uber and LinkedIn.

The main reason for its adoption is its high scalability. In addition, its architecture, flexibility and low cost make it a good choice for implementing Microservices and Serverless architecture components. Even the main providers of Cloud products and services already support the development of scalable solutions using Node.js.

For the database we use MongoDB (*https://www.mongodb.com/*). This source-available cross-platform is an opensource, high performance and flexible database, being considered the main NoSQL database. NoSQL databases have some advantages over other types, especially when we need scalability, flexibility, good performance and ease of queries.

MongoDB is document-oriented, that is, data is stored as documents, unlike relational model databases, where we work with records in rows and columns. Documents can be described as data in key-value format, in this case, using JSON (JavaScript Object Notation) format (*https://www.json.org/*).

4 Results

The COVIData tool (https://covidata.ufabc.edu.br/) was made available on April 2020 . In the first two weeks of use (from launching to May 2020), 7012 screenings were carried out throughout Brazil. The database was then processed and filtered to select the ones carried out in the Grande ABC region and to remove duplicated or inconsistent data. Altogether, after this treatment, 2672 screenings were considered as interest data. From these, 1106 users were classified as suspected of COVID infection according to their responses. The distribution

of the levels of suspicion was: 59% without suspicion, 6% mild suspicion, 4% medium suspicion, 31% severe suspicion.

Using the geolocation information linked to these screenings, it was possible to track which cities owned the highest number of suspected cases. Using the Leaflet [2] package was generated an interactive graph with the estimated density of screenings classified as suspected of being infected by COVID-19 (Fig. 2).

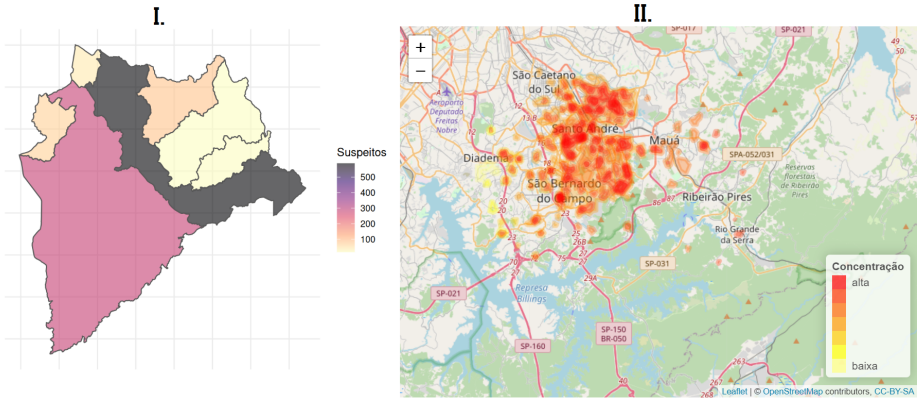

Fig. 2: I. Suspected cases identified for the ABC region and II. Interactive graph of the estimated density of screens classified as suspicious

At this very first moment right after the availability of the platform, the aim of the tool was to provide simple, quick and easy information to both the consortium and the population - who has access to reports directly on the platform's website. The graphics presented by COVIData show statistics about of the number of cases and their distribution over age groups and gender, as well as the frequency with which symptoms were reported for the different classifications of suspicion (Fig. 3) and generated. From the age groups graphic, it is visible the suspicion is concentrated among individuals between 20 and 60 years old. Although studies have [1] shown that elderly patients have more symptoms and have a higher risk of death, the sample demonstrates that the elderly are not the main focus of COVID-19 in the region.

Regarding gender, the suspicion of infection was more prevalent among women, representing 59% of the total. Again, the sample did not corroborate with published research that showed the prevalence among men and women is the same despite the fact that the disease has higher mortality among men [6].

Through the graphs which show symptoms, it was also possible to observe that for individuals classified as serious suspects, the frequency in which the different symptoms are reported was more balanced, while for users without suspicion and with mild suspicion the symptoms "headache" and "coughs" are reported much more frequently than other symptoms. Despite the balance in the frequency of symptoms reported by those classified as serious suspects, it was

possible to identify that symptoms such as respiratory distress, chest pain, and loss of taste and smell become more relevant in patients with the highest levels of suspicion.

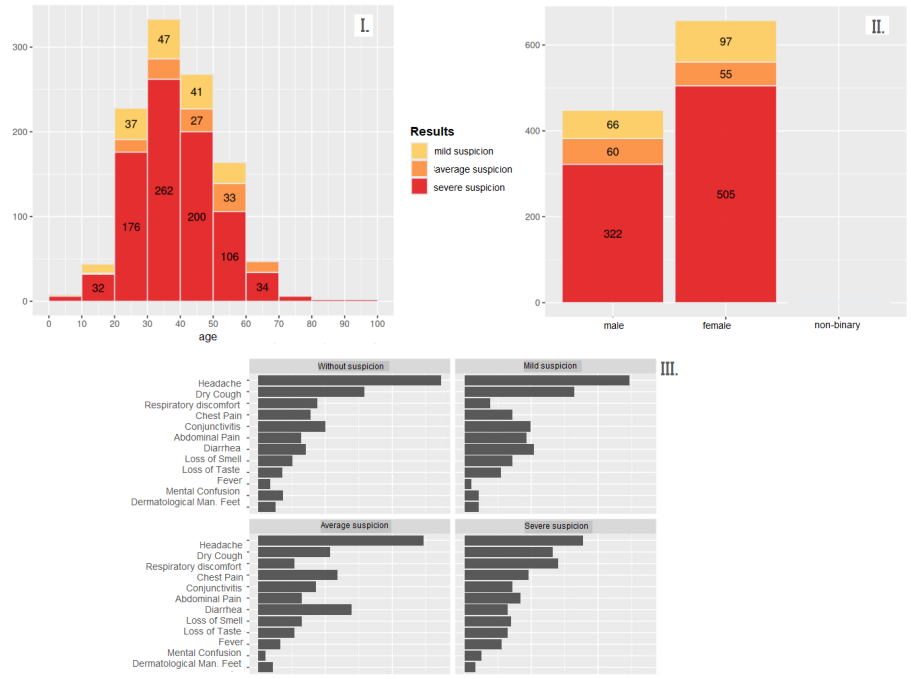

Fig. 3: I. Prevalence of cases by age, II. Prevalence of cases by gender, III. Frequency of symptoms according to risk classification

Finally, after much study and analysis of topics related to social media, the creation of arts, videos and podcast was carried out, with weekly publications on days and times of greater access, in addition to the use of tools such as stories to reach a greater number of people, even with the possible fatigue that the COVID-19 issue has caused after so long a pandemic [7].

5 Conclusion

By these means, we may state that both a accurate diagnosis and follow-up and immediate responses have been of immeasurable need while we attempt to fight back and direct our efforts on the non-spread of Coronavirus disease.

It's also important to enhance the fact that, when we bear in mind facing pandemics like the one we've been facing for the last months (COVID-19) having

tools which may be of immediate response is a must as the disease spreads more and more as the time passes.

Withal, the purposed tool targets not only a pre-diagnosis of COVID-19 but also helps people out on how to prevent themselves, what to do next concerning their infection suspicious level and also build up many analysis within the multiple data collected; ending up in benefits conjointly of use to many health organizations and governmental attitudes to be taken.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgements

The authors would like to thank the Federal University of ABC to the support.

References

[1] "Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series". In: *BMJ* 368 (2020). Ed. by. DOI: 10.1136/bmj.m792. eprint: https://www.bmj.com/content/368/bmj.m792.full.pdf. URL: https://www.bmj.com/content/368/bmj.m792.

[2] Joe Cheng, Bhaskar Karambelkar, and Yihui Xie. *leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library*. R package version 2.1.1. 2022. URL: https://CRAN.R-project.org/package=leaflet.

[3] X Cheng et al. "Symptom Clustering Patterns and Population Characteristics of COVID-19 Based on Text Clustering Method". In: *Front Public Health* (2022). DOI: 10.3389/fpubh.2022.795734.

[4] Lei Fang, George Karakiulakis, and Michael Roth. "Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection?" In: *The Lancet. Respiratory Medicine* 8.4 (2020), e21.

[5] Apple Inc., Centers of Disease Control, and Prevention CDC. *Coronavirus (COVID-19)*. 2020. URL: https://www.apple.com/covid19/.

[6] Jian-Min Jin et al. "Gender Differences in Patients With COVID-19: Focus on Severity and Mortality". In: *Frontiers in Public Health* 8 (2020), p. 152. ISSN: 2296-2565. URL: https://www.frontiersin.org/article/10.3389/fpubh.2020.00152.

[7] Omar Chávez-Martínez y Rosa Avila-Malpica y Lissette Gómez-Rivera y José Franco-Rico. "Los servicios de información ante la pandemia por COVID-19 / Information services facing the COVID-19 pandemic". In: *Revista Médica del Instituto Mexicano del Seguro Social* 60.1 (2022), pp. 1–3. ISSN: 2448-5667. URL: http://revistamedica.imss.gob.mx/editorial/index.php/revista_medica/article/view/4480.

[8] Sophie A M van Kessel et al. "Post-acute and long-COVID-19 symptoms in patients with mild diseases: a systematic review". In: *Family Practice* 39.1 (July 2021), pp. 159–167. ISSN: 1460-2229. DOI: 10.1093/fampra/cmab076. eprint: https://academic.oup.com/fampra/article-pdf/39/1/159/42243810/cmab076.pdf. URL: https://doi.org/10.1093/fampra/cmab076.

[9] Jerome R Lechien et al. "Objective olfactory testing in patients presenting with sudden onset olfactory dysfunction as the first manifestation of confirmed COVID-19. infection". In: *Medrxiv* (2020).

[10] Jerome R Lechien et al. "Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study". In: *European Archives of Oto-Rhino-Laryngology* (2020), pp. 1–11.

[11] Yanan Li et al. "Acute cerebrovascular disease following COVID-19: a single center, retrospective, observational study". In: (2020).

[12] F Mazzotta and T Troccoli. "Acute acro-ischemia in the child at the time of COVID-19". In: *Dermatologia Pediatrica, Bari* (2020).

[13] World Health Organization. *WHO - Coronavirus Disease (COVID-19) Dashboard*. 2022. URL: https://covid19.who.int/.

[14] *Protocolo de Manejo Clínico da Covid-19 na Atenção Especializada*. 2020. URL: https://bvsms.saude.gov.br/bvs/publicacoes/manejo_clinico_covid-19_atencao_especializada.pdf.

[15] Pedro Augusto Sampaio Rocha-Filho et al. "Headache, anosmia, ageusia and other neurological symptoms in COVID-19: a cross-sectional study". In: *The Journal of Headache and Pain* (Mar. 2022). DOI: https://doi.org/10.1186/s10194-021-01367-8.

[16] Ministério de Saúde Brasil. *Coronavirus - SUS*. 2020. URL: https://play.google.com/store/apps/details?id=br.gov.datasus.guardioes&hl=pt_BR.

[17] Gholizadeh Mesgarha M Pour Mohammad A-Shaka Z Goodarzi Seirafianpour F Pourriyahi H. "A. A systematic review on mucocutaneous presentations after COVID-19 vaccination and expert recommendations about vaccination of important immune-mediated dermatologic disorders". In: *Dermatol Ther* (2022). DOI: 10.1111/dth.15461.

[18] BRO Silva et al. "Clinical-Epidemiology Aspect of Inpatients With Moderate or Severe COVID-19 in a Brazilian Macroregion: Disease and Countermeasures". In: *Front Cell Infect Microbiol* (2022).

[19] Pauline Vetter et al. *Clinical features of covid-19*. 2020.

[20] Xiao-Wei Xu et al. "Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series". In: *BMJ* 368 (2020). DOI: 10.1136/bmj.m606. eprint: https://www.bmj.com/content/368/bmj.m606.full.pdf. URL: https://www.bmj.com/content/368/bmj.m606.

[21] Fei Zhou et al. "Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study". In: *The lancet* (2020).