Bonn-Aachen International Center for Information Technology

(B-IT)

University of Bonn

Master Program in Life Science Informatics

Master Thesis

**Landscaping of COVID-19 Clinical Trials for the Discovery of Insightful Patterns on Ethnicities**

Submitted by : Ram Kumar Ruppa Surulinathan

First Reviewer : Prof. Dr. Martin Hofmann-Apitius

Second Reviewer : Prof. Dr. Juliane Fluck

Internal Supervisor : Dr. Marc Jacobs

*A thesis submitted in fulfillment of the requirements for the degree of M.Sc. in Life Science Informatics in the*

Bonn-Aachen International Center for Information Technology (B-IT)

*in collaboration with*

Fraunhofer Institute for Algorithms and Scientific Computing (SCAI)

August 2021

**ACKNOWLEDGEMENTS**

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| 1 | Introduction | 1 |
|  |  |  |
|  |  |  |
|  |  |  |
| 2 | Theoretical Background |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 3 | Materials and Methods |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 4 | Results |  |
|  |  |  |
|  |  |  |
|  |  |  |
| 5 | Conclusion |  |
|  |  |  |
|  |  |  |
| 6 | Scientific References |  |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| Figure Number | Figure Title | Page |
|  | Structure of novel coronavirus SARS-CoV-2 |  |
|  |  |  |
|  |  |  |
|  |  |  |

**LIST OF TABLES**

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| SARS-CoV-2 | Severe Acute Respiratory Syndrome Coronavirus-2 |
|  |  |
|  |  |
|  |  |

**ABSTRACT**

**INTRODUCTION**

**THEORETICAL BACKGROUND**

2.1 SARS-CoV-2 a novel coronavirus

The Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) a novel strain and fatal coronavirus was first identified in Wuhan city of China in December 2019. The common symptoms include increased body temperature, dry cough, nausea and body pains. The rapid spread of the virus posed a threat of life to the global environment. According to the records maintained by Worldometer, the top ten most affected nations include USA, India, Brazil, France, Russia, UK, Turkey, Italy, Spain and Germany (*COVID Live Update: Worldometer*, n.d.).

The coronaviruses generally is classified under the family Coronoviridae and subfamily Coronavirinae which is subdivided into four genera namely Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus (Mittal et al., 2020). The SARS-CoV-2, a member of Betacoronavirus genera whose sequence is 96% homologous to the bat coronavirus. Its primary reservoir is considered to be bats and transmitted to human beings through an intermediate host called Pangolin (Zhao et al., 2020).

The SARS-CoV-2 also called by COVID-19 is spherical in structure with positively stranded RNA genome packed inside the nucleocapsid protein (N) and enveloped by the membrane glycoprotein protein (M), envelope protein (E), and the spike protein (S). The typical virus lengths between 26.4 and 31.7 kb with the GC content ranging between 32% and 43% thus indicating to be the largest RNA virus (Mousavizadeh & Ghasemi, 2020).

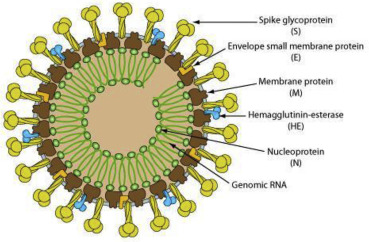


Figure 1 Structure of novel coronavirus SARS-CoV-2 , Adapted from (Mousavizadeh & Ghasemi, 2020)

2.2 Clinical trials

Clinical trials are experimental studies performed using volunteers with the objective of examining different interventions like medical, surgical or behavioral ones. They can be classified into interventional studies and observational studies. These studies are led by a medical doctor assisted by other doctors, nurses carried out in hospitals, universities and research institutes (*Learn About Clinical Studies - ClinicalTrials.Gov*, n.d.).

The United States Food and Drug Administration (FDA) has defined the different stages of clinical trial phases to determine whether the drug could be employed for public usage. They are described as follows:

Phase I: During the initial phase, around 20 to 80 participants are enrolled having no underlying medical complications to evaluate the highest dosage levels that can be administrated without serious side effects.

Phase II: During this phase, around 100 to 300 participants are enrolled to evaluate the effectiveness of the medication along with short time side effects if occurred. This is carried out up to several years.

Phase III: During this phase, around 3000 participants are enrolled to evaluate the effectiveness of the medication across diverse population and varied dosage thus studying both the drug safety and efficacy. The rare and long time effects are observed during this phase. The medications are approved by the FDA if the trial results are positive.

Phase IV: During this final phase, thousands of participants are enrolled to evaluate the long term side effects and efficacy of the approved medications across diverse population (*What Are Clinical Trials and Studies? | National Institute on Aging*, n.d.).

2.3 Clinical Trial Registries

The results of the clinical studies are recorded and published in a repository called Clinical Trial Registries. They are open source and their accession is available for scientific community and public. Each nation has their own trial registry systems centrally maintained by their Government or other approved institution. Some notable registries are described as follows:

2.3.1 ClinicalTrials.gov

ClinicalTrials.gov is an US based trial registry maintained by the National Library of Medicine (NLM) at the National Institute of Health (NIH), made available in February 2000. It holds both interventional and observational studies carried in 50 US states and 220 countries. The records contains information pertaining to the disease under investigation; the type of interventions employed; meta information like ethnicity of the participants, demographic data, inclusion/exclusion criteria for the study, location, list of comorbidities if any, clinical variables and laboratory variables. The results are included sometimes if the trials are subject to Section 801 of FDAAA (FDAAA 801) (*ClinicalTrials.Gov Background - ClinicalTrials.Gov*, n.d.).

The registry actually recorded 5,420 studies for the novel COVID 19 virus as of 21st April 2021.

**SCIENTIFIC REFERENCES**

*ClinicalTrials.gov Background - ClinicalTrials.gov*. (n.d.). Retrieved April 21, 2021, from https://clinicaltrials.gov/ct2/about-site/background

*COVID Live Update: Worldometer*. (n.d.). Retrieved April 21, 2021, from https://www.worldometers.info/coronavirus/

*Learn About Clinical Studies - ClinicalTrials.gov*. (n.d.). Retrieved April 21, 2021, from https://www.clinicaltrials.gov/ct2/about-studies/learn

Mittal, A., Manjunath, K., Kumar, R., Id, R., Kaushikid, S., Kumar, S., & Vermaid, V. (2020). *COVID-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-CoV-2*. https://doi.org/10.1371/journal.ppat.1008762

Mousavizadeh, L., & Ghasemi, S. (2020). Genotype and phenotype of COVID-19: Their roles in pathogenesis. In *Journal of Microbiology, Immunology and Infection*. Elsevier Ltd. https://doi.org/10.1016/j.jmii.2020.03.022

*What Are Clinical Trials and Studies? | National Institute on Aging*. (n.d.). Retrieved April 21, 2021, from https://www.nia.nih.gov/health/what-are-clinical-trials-and-studies

Zhao, J., Cui, W., & Tian, B. (2020). The Potential Intermediate Hosts for SARS-CoV-2. *Frontiers in Microbiology*, *11*, 2400. https://doi.org/10.3389/fmicb.2020.580137