The Study of the COVID-19 Virus: A Review

Marika L. Forsythe

Department of Pathology, NorthShore University HealthSystem, Evanston, Illinois, United States

Abstract:
COVID-19 was declared a pandemic in 2020 and spanned a three-year period, causing devastating effects across the globe. The death toll from the infection rose to millions, with medical experts and government officials worldwide working tirelessly to control its spread. Symptoms from the virus ranged from mild (i.e., fever and cough) to severe (i.e., respiratory failure and multi-organ dysfunction), creating difficulties in tracking its progression and developing appropriate treatments. The aim of this article is to provide a comprehensive review of the COVID-19 virus and its varied characteristics studied throughout the pandemic, including its structure, common comorbidities, and the numerous treatment options made available. Viral and host genetics are described as well, as multiple studies have linked molecular variants to differing degrees of disease severity. Polygenic risk scoring (PRS) has been an approach used for the determination of risk for severe outcomes, assisting with the identification of significant genetic variants and high-risk population groups.

Keywords: COVID-19, Pandemic, Viral structure, Treatment, Vaccine, Genetics, Polygenic risk score.

1. INTRODUCTION
There are multiple types of coronaviruses, each infecting many different animals and involving different modes of host cell entry. They can be divided into three genera (I, II, and III) based on recent genome sequence analysis [1]. Group I includes animal pathogens, as well as the human coronaviruses HCoV-229E and HKU1, which are responsible for various respiratory infections. Group II includes animal pathogens, which are of veterinary relevance, and human coronaviruses OC43 and NL63, which also cause respiratory infections. Group III, as of now, only includes avian coronaviruses [2]. There is currently a debate regarding whether SARS-CoV defines a new group of coronaviruses or whether it is a distant member of Group II (Table 1) [3, 4].

2. THE VIRAL STRUCTURE, MECHANISM OF ENTRY, LABORATORY DIAGNOSTICS, AND DISEASE STAGES
COVID-19, or SARS-CoV-2, is a positive-sense, single-stranded RNA virus surrounded by an envelope made of glycoprotein. Under an electron microscope, it has a crown or “corona”-like morphology, which is what established the name “coronavirus” for this family [4]. The genome is comprised of roughly 30,000 nucleotides, which encode all its necessary proteins, including four structural proteins and numerous non-structural proteins (Fig. 1). The structural proteins include the following:
Table 1. Types of coronaviruses.

<table>
<thead>
<tr>
<th>Group</th>
<th>Virus</th>
<th>Host</th>
<th>Disease caused</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>229E</td>
<td>Human</td>
<td>Respiratory infection</td>
</tr>
<tr>
<td></td>
<td>TGEV</td>
<td>Pig</td>
<td>Respiratory and enteric infection</td>
</tr>
<tr>
<td></td>
<td>PRCoV</td>
<td>Pig</td>
<td>Respiratory infection</td>
</tr>
<tr>
<td></td>
<td>Canine coronavirus</td>
<td>Pig</td>
<td>Enteric infection</td>
</tr>
<tr>
<td></td>
<td>FeCoV</td>
<td>Cat</td>
<td>Enteric infection</td>
</tr>
<tr>
<td></td>
<td>FIPV</td>
<td>Cat</td>
<td>Respiratory, enteric, neurologic infection, and hepatitis</td>
</tr>
<tr>
<td></td>
<td>NL-63</td>
<td>Human</td>
<td>Respiratory infection and croup</td>
</tr>
<tr>
<td>II</td>
<td>OC43</td>
<td>Human</td>
<td>Respiratory infection and possibly enteric infection</td>
</tr>
<tr>
<td></td>
<td>MHV</td>
<td>Mouse</td>
<td>Enteric and neurologic infection and hepatitis</td>
</tr>
<tr>
<td></td>
<td>Sialodacryoadenitis coronavirus</td>
<td>Rat</td>
<td>Neurologic infection</td>
</tr>
<tr>
<td></td>
<td>Hemagglutinating encephalomyocarditis virus</td>
<td>Pig</td>
<td>Respiratory, enteric, and neurologic infection</td>
</tr>
<tr>
<td></td>
<td>BCoV</td>
<td>Cow</td>
<td>Enteric infection</td>
</tr>
<tr>
<td></td>
<td>HKU1</td>
<td>Human</td>
<td>Respiratory infection</td>
</tr>
<tr>
<td></td>
<td>SARS-CoV</td>
<td>Human</td>
<td>Severe acute respiratory syndrome</td>
</tr>
<tr>
<td>III</td>
<td>IBV</td>
<td>Chicken</td>
<td>Respiratory infection, hepatitis, and other</td>
</tr>
<tr>
<td></td>
<td>Turkey coronavirus</td>
<td>Turkey</td>
<td>Respiratory and enteric infection</td>
</tr>
</tbody>
</table>

**Fig. (1).** Schematic representation of the COVID-19 virus structure.

**Nucleocapsid (N) Protein:** Within the protein shell capsid, there is the N-protein, which binds to the viral RNA via the N-terminal of the protein, forming the nucleocapsid. This process allows the virus to hijack human cells during infection and use their cellular machinery to create new viral particles. It is a vital protein involved with viral replication and transcription and ultimately completes viral formation [5].

**Spike (S) Glycoprotein:** The integration of spike (S) glycoprotein over the surface of the virus allows it to mediate its attachment to host cell angiotensin-converting enzyme 2 (ACE2) receptors and fusion between the viral and host cell membranes, facilitating entry [6]. It is made of three identical chains of 1,273 amino acids each and two well-defined protein domain regions: S1 and S2 subunits. These are associated with cell recognition and fusion of the viral and cell membranes, respectively [7]. It has been observed in some coronavirus strains that the expression of the S protein can also mediate cell-to-cell fusion between infected and adjacent, uninfected cells. This forms giant multinucleated cells, or syncytia, which are potentially used as a strategy for the virus to spread directly between cells while avoiding countering antibodies [7, 8].

**Envelope (E) Protein:** This small membrane protein is made of roughly 76-109 amino acids and is a minor component of the virus particle. Nevertheless, it plays an important role in virus assembly, membrane permeability of the host cell membrane, and viral-host cell interactions [9]. The majority of this protein is located at the site of intracellular trafficking, such as the endoplasmic reticulum and the Golgi apparatus, where it assists with viral assembly, budding, and maturation [10].

**Membrane (M) Protein:** This is the most abundant protein found on the viral surface and determines the shape of the viral envelope. It is believed to be the central
organizer for assembly via interactions with the other major proteins [11].

- Interaction between S and M proteins is necessary for retention of the S protein at the site of intracellular trafficking to ensure incorporation into new viral particles [12].
- Binding M with N proteins stabilizes the nucleocapsid as well as the internal core of virions, which ultimately allows the completion of the viral assembly [13].
- M and E proteins together make up the viral envelope, and their interaction allows for the adequate production and release of virus-like particles (VLPs) [14].

Another non-structural protein is the Hemagglutinin-esterase dimer (HE), located on the viral surface. It is not required for replication; however, it may be involved with viral entry into the host cells and appears to be important for infection [15].

The mechanism of viral entry and replication in the host cell involves numerous steps. The spike protein is first attached to the ACE2 receptors located on the surface of many human cells, allowing viral entry. The S protein is then cleaved by host proteases between the S1 and S2 subunits. In a later stage, the S2 domain is cleaved, releasing a fusion peptide that leads to the activation of membrane fusion. Human cells will generally ingest the virus via endocytosis. Once in the endosome, a three-step process has been hypothesized as the method for membrane fusion. This involves receptor-binding, induced conformational changes in the S glycoprotein, and cathepsin L proteolysis within the endosome via intracellular proteases [16]. This is followed by the opening of the endosome to allow the virus entry into the cytoplasm and uncoating of the viral nucleocapsid. A second hypothesis for viral entry involves a two-step method where the virus binds to a host cell surface receptor by its S1 subunit, resulting in the S glycoprotein being cleaved by the host proteases. This causes fusion of the viral and host cell membranes at low pH via the S2 subunit [17]. Once the virus has entered the cytoplasm, the genetic material can then be replicated via the replication/transcription complex (RTC). This complex is part of the viral genome and is made of non-structural proteins. Replicase proteins are translated from the genome, allowing for the generation of full-length negative sense RNA strands. These are then used as templates for new full-length genomes. The required structural proteins are generated, and the newly created virions are exported from the host cell by transportation to the cell membrane in vesicles and secreted via exocytosis. These new virions are then able to infect other host cells (Fig. 2) [18].

![Fig. (2).](image-files/fig2.png)

**Fig. (2).** The mechanism of COVID-19 entry and replication. The virus enters the host cell via the ACE2 receptor and releases its RNA. The RNA is translated into polypeptides, which then undergo proteolysis to release active proteins. This allows for RNA replication and protein translation, creating new viral particles to be released from the host cell.
Other strains of coronavirus generally only cause mild infection, according to case reports from the Chinese Center for Disease Control and Prevention. COVID-19 symptoms, however, can range from mild to severe to critical. The most commonly experienced COVID-19 symptoms are cough, fever, and myalgia, according to the CDC of the United States (Fig. 3) [12]. In the first part of the pandemic, 81% of patients tended to have mild symptoms, exhibiting either no or mild pneumonia symptoms. Up to 14% of patients had severe disease, exhibiting dyspnea, hypoxia, or >50% lung involvement within the first 2 days of infection. The critical disease was found in 5% of patients who exhibited respiratory failure, shock, or multi-organ dysfunction [19]. Patients were hospitalized when they experienced severe or critical disease. However, it is estimated that around 33% of infected patients remained asymptomatic [20]. Overall fatality rates varied by location and across different risk groups, with an estimation to be between 0.1-6.0% case-fatality by country. It was initially thought to be much higher due to the larger number of asymptomatic cases that were not diagnosed [21, 22]. Fatality rates have progressively dropped due to the availability of new vaccines and treatments; however, severe cases still result in hospitalization with critical disease.

COVID-19 has multiple clinical phases that can be used to monitor its progression. While it was initially believed to be only a respiratory disease, it has become more evident that it is a multi-organ and heterogeneous illness. As the disease progresses and the severity increases, the disease stage can be determined using objective and molecular criteria. Knowing the disease data can assist in clinical decisions for patient management, improved prognosis, and appropriate treatment selections. Cordon-Cardo et al. (2020) described the characteristics of the stages, including the therapeutic interventions (Table 2). The first stage involves diagnostic tests, such as for SARS-CoV-2 detection and viral load (RT-qPCR, qLAMP). The other three stages involve disease monitoring for inflammatory markers, cytokine panels, thrombosis markers, and antibody tests (quantitative, neutralization, etc.) [23].

Table 2. COVID-19 severity staging.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Characteristics</th>
<th>Treatments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1:</td>
<td>Asymptomatic</td>
<td>Viral entry and replication</td>
</tr>
<tr>
<td>Stage 2:</td>
<td>Mild/Moderate</td>
<td>Viral dissemination</td>
</tr>
<tr>
<td>Stage 3:</td>
<td>Severe</td>
<td>Multisystem inflammation</td>
</tr>
<tr>
<td>Stage 4:</td>
<td>Critical</td>
<td>Endothelial damage and thrombosis</td>
</tr>
</tbody>
</table>

COVID-19 diagnosis is typically performed using reverse transcriptase-polymerase chain reaction (RT-PCR) assay, chest CT, or both, depending on access to local viral testing and the time of diagnosis in the course of the disease. Chest CT is most effective in disease monitoring...
for patients who have progressed to pneumonia. In patients with COVID-19 pneumonia, ground-glass opacities were observed on chest CT more frequently than in patients with non-COVID-19 pneumonia [24, 25]. RT-PCR is used to generate cDNA from SARS-CoV-2 RNA extracted from respiratory samples. Common SARS-CoV-2 gene targets include the envelope, nucleocapsid, spike, RNA-dependent RNA polymerase, and ORF1 genes [26]. Multiple different primer-probe sets for detection and test kits have been developed with variable sensitivity [27]. The viral RNA obtained from the samples is detected via PCR, where the RNA undergoes multiple replications until it can be detected with a fluorescent signal. The fewer cycles required to be detected indicate a higher viral RNA load. When the number of cycles is less than 40, the test is considered positive. However, a positive PCR result indicates the presence of viral RNA and is not necessarily a viable virus [26].

Generally, the viral RNA becomes detectable as early as day 1 of symptoms and peaks within the first week of onset. Positivity usually starts to decline by week 3 [28]. It also depends on the type of sample obtained, as a sputum sample may remain positive for longer than a nasopharyngeal swab. Highest positivity is found in bronchoalveolar lavage specimens (93%), followed by sputum (72%), nasal swab (63%), and pharyngeal swab (32%) [29]. An indirect method to detect a positive COVID-19 infection is to measure the immune response. Measuring serology can be useful in cases where the patients present beyond the first 2 weeks of onset due to a milder course. It can also help identify how widespread COVID-19 is within a community. IgM and IgG antibodies can be detected as early as the fourth day of symptom onset, and higher levels are observed in the second and third week of illness. Afterwards, IgM levels begin to decline and almost disappear by week 7, whereas IgG levels will continue to persist beyond 7 weeks (Fig. 4) [30, 31].

Studies have shown that progressive respiratory failure is the primary cause of death in those infected with the COVID-19 virus. This is due to the development of acute respiratory distress syndrome (ARDS)-like symptoms; however, COVID-19 pneumonia is a specific type of respiratory disease. ARDS is an acute, life-threatening inflammation of the lung due to infection, trauma, or other inflammatory conditions. This excessive inflammation leads to alveolar damage and increased permeability of endothelial and epithelial cells, causing decreased respiratory compliance. These cellular changes result in the accumulation of protein-rich fluid in the interstitium and air space, which causes impaired gas exchange and varying levels of hypoxemia. This impairment of the lung microvascular barrier is central to the pathogenesis of ARDS [32, 33]. Chemokines and cytokines have also been shown to contribute to lung injury, as they lead to the recruitment of various types of immune cells, such as neutrophils. These immune cells are important for resolving the cause of the inflammation; however, they can lead to significant damage to the normal lung cells as well [34].

![Fig. (4). Estimated variation over time of diagnostic tests for detection of COVID-19 infection relative to onset of symptoms.](image-url)
ARDS in patients with COVID-19 have been observed to have different physical changes than in patients with ARDS from other causes. One study compared the lungs of seven patients who died from COVID-19 to the lungs of seven patients who died from ARDS secondary to an influenza infection and 10 uninfected control lungs. The lungs of those who died of COVID-19 or influenza-associated respiratory failure were found to have diffuse alveolar damage in the peripheral lung with perivascular T-cell infiltration. However, COVID-19 lungs also showed distinctive vascular features involving severe endothelial injury associated with an intracellular virus and disrupted cell membranes. There was also widespread thrombosis with microangiopathy. Alveolar capillary microthrombi was found to be 9x more prevalent in patients with COVID-19 than those with influenza. There was also a 2.7x increase in new vessel growth among COVID-19 patients, predominantly through intussusceptive angiogenesis [35].

Comorbidities are quite commonly seen in patients infected with COVID-19, with hypertension being the most common, followed by diabetes and coronary heart disease (Table 3). Risk factors, such as older age, high lactate dehydrogenase levels, and a D-dimer level greater than 1 μg/mL, were also indicative of a poor prognosis in the early stage [36, 37]. Other studies also found independent risk factors that were associated with no improvement in these patients, including male sex, a severe COVID-19 condition, expectoration, muscle ache, and decreased albumin [38]. Additionally, other particular laboratory findings have been associated with worse outcomes, including lymphopenia, thrombocytopenia, and elevated liver enzymes (Table 4) [39, 40].

Table 3. Common comorbidities associated with COVID-19.

<table>
<thead>
<tr>
<th>Comorbidities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
</tr>
<tr>
<td>Chronic lung disease</td>
</tr>
<tr>
<td>Chronic liver disease</td>
</tr>
<tr>
<td>Diabetes mellitus type 1 and 2</td>
</tr>
<tr>
<td>Down syndrome</td>
</tr>
<tr>
<td>Heart conditions (ex. coronary artery disease)</td>
</tr>
<tr>
<td>HIV</td>
</tr>
<tr>
<td>Mental health disorders (ex. depression)</td>
</tr>
<tr>
<td>Neurologic conditions (ex. dementia)</td>
</tr>
<tr>
<td>Obesity and overweight</td>
</tr>
<tr>
<td>Pregnancy or recent pregnancy</td>
</tr>
<tr>
<td>Smoker (current or former)</td>
</tr>
<tr>
<td>Sickle cell disease or thalassemia</td>
</tr>
<tr>
<td>Solid organ or blood stem cell transplanation</td>
</tr>
<tr>
<td>Substance use disorders</td>
</tr>
<tr>
<td>Tuberculosis</td>
</tr>
<tr>
<td>Corticosteroids/other immunosuppressive use</td>
</tr>
<tr>
<td>Asthma</td>
</tr>
<tr>
<td>Hypertension</td>
</tr>
<tr>
<td>Other immune deficiencies</td>
</tr>
</tbody>
</table>

Table 4. Laboratory abnormalities associated with severe COVID-19.

<table>
<thead>
<tr>
<th>Laboratory Feature</th>
<th>Abnormal Value</th>
<th>Normal Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-dimer</td>
<td>&gt;1000 ng/mL</td>
<td>&lt;500 ng/mL</td>
</tr>
<tr>
<td>CRP</td>
<td>&gt;100 mcg/L</td>
<td>&lt;8.0 mg/L</td>
</tr>
<tr>
<td>LDH</td>
<td>&gt;245 units/L</td>
<td>110-210 units/L</td>
</tr>
<tr>
<td>Troponin</td>
<td>&gt;2x the upper limit of normal</td>
<td>Females: 0 - 9 ng/L</td>
</tr>
<tr>
<td>Ferritin</td>
<td>&gt;500 mcg/L</td>
<td>Females: 10-200 mcg/L</td>
</tr>
<tr>
<td>CPK</td>
<td>&gt;2x the upper limit of normal</td>
<td>40-150 units/L</td>
</tr>
<tr>
<td>Absolute lymphocyte count</td>
<td>&lt;800/μL</td>
<td>1800-7000/μL</td>
</tr>
<tr>
<td>Platelet count</td>
<td>&lt;23-31 x 10^9/L</td>
<td>150-450 x 10^9/L</td>
</tr>
<tr>
<td>AST and ALT</td>
<td>&gt;3x the upper limit of normal</td>
<td>AST: 10-40 units/L</td>
</tr>
</tbody>
</table>

While initially believed to be exclusively a respiratory disease, COVID-19 has been observed to affect more than just the throat and the lungs. The virus has been observed to have a significant impact on the cardiovascular system, leading to various conditions, including myocarditis, myocardial injury, scar formation, arrhythmias, heart block, and vascular occlusion due to local thrombus formation or embolism, which can lead to cardiac death [41-43]. Other potential systems include the brain and the kidneys, resulting in multi-organ failure. It has been hypothesized that, especially in the later complicated stages of the disease, COVID-19 can be considered an endothelial disease. The reasoning for this takes into consideration how the vascular endothelium cells not only direct the circulating blood to the tissues but also assist with maintaining homeostasis in the body. This includes regulating an array of functions, such as hemostasis, fibrinolysis, inflammation, oxidative stress, vascular permeability, and vasculature structure. These functions normally contribute to coordinating host defenses and maintaining overall homeostasis; however, they can also contribute to disease if defense mechanisms overreach during an illness and turn against the host. Such is the case of COVID-19, which triggers an increase in cytokines and protein proinflammatory mediators, triggering endothelial functions towards a defensive mode. Continuing down this path can lead to a cytokine storm with positive feedback loops maintaining the cytokine production that overpowers the counter-regulatory mechanisms, resulting in significant damage to the host [44].

COVID-19 has been shown to bind to the receptor ACE2 on cell surfaces to gain access to the host cells. This receptor plays a functional role in the renin-angiotensin-aldosterone system (RAAS), which is a signalling pathway involved with the homeostatic regulation of vascular functions. This includes control over systemic and local blood flow, blood pressure, natriuresis, and trophic responses to many different stimuli. Angiotensin II (Ang II) is the main effector molecule of the system and binds to the ACE2 receptor [45, 46]. COVID-19 competes with Ang...
Il for this receptor. ACE2 expression is seen in numerous types of cells, including respiratory epithelial cells, myocardial cells, oesophagus epithelial cells, tongue epithelial cells, and enterocytes from the ileum, to name a few [47-49]. This variety in cell types expressing ACE2 can help explain why patients exhibit a vast array of symptoms when infected with COVID-19 due to the ability of the virus to diffuse throughout the body.

3. Past/current treatments and vaccines

By March 2020, COVID-19 had become so widespread it was characterized as a global pandemic. As lockdowns increased to slow the spread, efforts increased to develop treatment options for infected patients. Initially, no specific treatments showed high efficacy for treating the infection. Beyond the use of interventional therapies to treat the respiratory issues associated with the infection (described in Table 5) [50], classes of drugs mainly used as potential treatments included antiviral agents, immunomodulators, inflammation inhibitors, low-molecular-weight heparins, plasma, and hyperimmune immunoglobulins. Examples of these treatments include the following:

**Lopinavir/ritonavir (Kaletra):** These antivirals are protease inhibitors that inhibit viral replication (Fig. 5). They were mainly used in COVID-19 patients in the early stages with mild symptoms, being managed at home or in the hospital. Previous experiences with SAS-CoV-1 and MERS infections have suggested that this drug may improve the outcomes in some patients. However, a randomized trial study for severe COVID-19 cases showed no significant clinical improvement with the use of lopinavir/ritonavir therapy [51, 52].

**Table 5. Therapeutic options for severe acute respiratory distress syndrome (ARDS) associated with COVID-19.**

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-flow nasal oxygen</td>
<td>Might prevent or delay the need for intubation</td>
</tr>
<tr>
<td>Tidal volume</td>
<td>Use 6 mL/kg per predicted bodyweight</td>
</tr>
<tr>
<td>Plateau airway pressure</td>
<td>Maintain at &lt;30 cm H₂O, is possible</td>
</tr>
<tr>
<td>Positive end-expiratory pressure (PEEP)</td>
<td>Consider moderate-high levels, if needed</td>
</tr>
<tr>
<td>Neuromuscular blockade</td>
<td>For ventilator dyssynchrony, increased airway pressure, or hypoxemia</td>
</tr>
<tr>
<td>Prone position</td>
<td>For worsening hypoxemia, PaO₂:FiO₂ &lt;100-150 mm Hg</td>
</tr>
<tr>
<td>Inhaled NO</td>
<td>Use 5-20 ppm</td>
</tr>
<tr>
<td>Fluid level management</td>
<td>Aim for a negative fluid balance of 0.5-1.0 L/day</td>
</tr>
<tr>
<td>Renal replacement therapy</td>
<td>For oliguric renal failure, acid-base management, and negative fluid balance</td>
</tr>
<tr>
<td>Extracorporeal membrane oxygenation</td>
<td>Used the EOLIA trial criteria [71]</td>
</tr>
</tbody>
</table>

**Remdesivir (Veklury):** This antiviral belongs to the class of nucleotide analogues and can be used in moderate and severe cases of COVID-19 (Fig. 6). Multiple studies have been conducted to examine its efficacy, which have concluded the following results:

![Fig. (5). Lopinavir/ritonavir function as protease inhibitors, inhibiting proteolysis of newly formed polypeptides, which prevents viral replication.](image)

![Fig. (6). Remdesivir functions as an RNA-dependent RNA polymerase inhibitor, preventing the translation of newly formed viral RNA.](image)
Grein et al. (2020) analyzed the use of remdesivir in a small cohort of severe COVID-19 patients. A total of 68% of patients who followed up after treatment were found to have clinical improvement [53].

Goldman et al. (2020) determined the first results of the phase 3 trial on the use of remdesivir in hospitalized patients, where patients with severe COVID-19 were administered the drug for either 5 or 10 days. By day 14 after initiation of treatment, clinical improvement was found in 64% of patients in the 5-day treatment group and 54% in the 10-day treatment group. While there was no great distinction between a 5-day and 10-day treatment course, it is suggested that patients undergoing mechanical ventilation could benefit from the 10-day treatment. However, further studies on high-risk groups are needed to determine the shortest period of therapy [54].

Beigel et al. (2020) did a randomized control trial with 1,059 patients, with 538 assigned to remdesivir and 512 on placebos. Patients assigned remdesivir were shown to have an average hospitalization time of 11 days with a mortality estimate of 7.1%, compared to 15 days with an 11.9% mortality estimate among those on placebos [55].

These studies have suggested that the use of remdesivir can prevent the progression of a COVID-19 infection, as well as shorten the recovery times for hospitalized patients.

**New Molecules:** Two molecules were developed by Dai et al. (2020) that are capable of blocking molecules 11a and 11b, which are protease enzymes that allow replication of SARS-CoV-2. Both molecules showed anti-COVID-19 activity in cell culture, and neither were observed to cause significant cytotoxicity. Further investigation was conducted in animal experiments to determine extended pharmacological potential. Both showed good pharmacokinetic properties, such as high bioavailability and half-life, and further studies on molecule 11a showed no obvious toxicity when conducted on animals, suggesting it might be a good candidate for human clinical trials [56].

**Tocilizumab (Actemra):** An antibody directed against the interleukin-6 (IL-6) receptor, tocilizumab was previously the most used drug against COVID-19 (Fig. 7). Multiple studies were conducted to determine its efficacy, and the results were as follows:

- A study was conducted by the Agenzia Italiana de Farmaco (AIFA) to determine its safety and efficacy. Results suggested tocilizumab could significantly reduce mortality at one month but had less of an impact on early mortality (i.e., 14 days) [57].
- Guaraldi et al. (2020) conducted a retrospective observational cohort study that included adults with severe COVID-19 pneumonia who were hospitalized at a tertiary center. The study evaluated the efficacy of tocilizumab treatment when combined with usual care in reducing mortality and its impact on the likelihood of invasive mechanical ventilation when compared with those receiving standard treatment. It was concluded that administration of tocilizumab, both intravenous and subcutaneous, may help reduce the risk of invasive mechanical ventilation or death in patients with severe COVID-19 pneumonia [58].
- Another study published on the AIFA website was conducted to determine the effectiveness of early administration of tocilizumab in patients with COVID-19 pneumonia. There were no significant differences observed in the number of ICU admissions and 30-day mortality rate, and it was therefore concluded that early administration of the drug did not provide any clinical benefit for the patients [59].

![Fig. (7). Tocilizumab is an antibody that acts as an IL-6 inhibitor, preventing it from binding to its receptor.](image-url)
Anakinra (Kineret): Anakinra is an inhibitor of IL-α and IL-β proinflammatory cytokines (Fig. 8). It was suggested that treating the hyperinflammation of COVID-19 patients could improve mortality rates. King et al. (2020) reviewed many studies that used anakinra to target hyperinflammation in COVID-19. They found some success in treating the macrophage activation syndrome caused by inflammatory conditions overall. It was important to take the dose and administration of anakinra into consideration due to its short plasma half-life. Studies have shown that using the subcutaneous route, for example, could guarantee an adequate plasma concentration with a bioavailability ranging from 80 to 95% [60].

Baricitinib (Olumiant): It is an inhibitor of Janus kinases (JAK), which are enzymes involved in the transduction of the cytokine-mediated signal (Fig. 9). Overstimulation of this pathway can potentially lead to hyperinflammation, as seen in many COVID-19 patients. Cantini et al. (2020) conducted an observational study in hospitals with moderate COVID-19 pneumonia patients to evaluate the effectiveness and safety of 2-week treatments with baricitinib plus antivirals in comparison with the standard of care therapy. It was observed that the 2-week case fatality rate was significantly lower in patients treated with baricitinib than the controls. This treatment may have also reduced the number of ICU admissions and deaths [61].

Corticosteroids: They are immunosuppressants that can be used in disease states where the body is hypersensitive and attacking the cells of the host (Fig. 10). A study was conducted in the United Kingdom called the randomised RECOVERY trial to determine which drugs were the most effective for treating adults hospitalised with COVID-19. The drugs included in the study included low-dose dexamethasone (a corticosteroid), lopinavir-ritonavir, hydroxychloroquine (an immunomodulator), azithromycin (an antibiotic), and tocilizumab (a monoclonal antibody). When examining the control group, who only received standard treatments, it was observed that 28-day mortality was higher in those who needed ventilation and intermediate in those who only required oxygen. In the dexamethasone treatment group, mortality was observed as one-third lower in ventilated patients and one-fifth lower in those requiring only oxygen [62]. The World Health Organization (WHO) recommended that corticosteroids should not be used in treatment for non-severe COVID-19 cases as they provide no benefit [63].

Fig. (8). Anakinra is an immunosuppressive drug. It functions by inhibiting the actions of IL-α and IL-β proinflammatory cytokines on cellular receptors.
Fig. (9). Baricitinib is a JAK inhibitor, preventing the increased production of cytokines that leads to hyperinflammation.

Fig. (10). Corticosteroids are immunosuppressants that inhibit the body's innate immune response. For a COVID-19 infection, corticosteroids can be given to prevent the increased production of cytokines and inflammation of the lungs, reducing the risk for pneumonia, fluid retention, and development of ARDS and fibrosis.
Anticoagulants: These are medicines that help prevent blood clots. Progressive alteration of certain inflammatory and coagulative parameters can occur in the late stages of COVID-19 infections. These include increased degradation of fibrin, leading to increased levels of D-dimer; consumption of coagulation factors, and thrombocytopenia, to name a few. Therefore, at this stage of the disease, when it becomes increasingly important to contain hyperinflammation due to the risk of blood clots, anticoagulant drugs, such as non-fractionated low-molecular-weight-heparin (LMWH) or unfractionated heparin can be used to reduce this risk. A retrospective analysis was conducted in 2020 that examined many cases of severe COVID-19 pneumonia patients. For those who had observable activation of their coagulation cascade, if they were administered heparin for at least 7 days, it was suggested to be an advantage in terms of survival. LMWH can also be used as a prophylaxis in the earlier stages of the disease when the patient suffers from hypomobility due to the illness to prevent the formation of venous thromboembolism [64].

Therapeutic Antibodies: They are antibodies taken from blood samples of patients who had recovered from COVID-19 to serve as an alternative therapy for patients who were subsequently infected. It was estimated that the required dose of antibodies to treat infected patients required the removal of antibodies from at least three recovered individuals to reach the optimal concentration. There is limited data on the effectiveness of these hyperimmune immunoglobulins as a treatment option due to the short period of time between the beginning of the pandemic and when this treatment was under consideration. The following studies were conducted with limited results:

- One of the first studies in 2020 used plasma in the treatment of infected patients. The main results initially reported improvement in the clinical and survival of the patients after the end of additional administration of plasma and hyperimmune immunoglobulins. However, these findings needed to be confirmed through randomized clinical trials (RCTs) [65].
- A trial evaluated the safety and efficacy of using donated plasma with COVID-19 antibodies as a treatment option for hospitalized patients. The study was terminated due to insufficient eligible participants; however, the results obtained did not show a significant improvement from the use of this treatment option [66].

Overall, the effectiveness of using the plasma of recovered COVID-19 patients to treat patients subsequently infected was not shown to be significant. This treatment, plus the previously described options, indicated the necessity of developing newer, more effective treatment options.

Molnupavir (Lagevrio): It is a nucleoside analog that inhibits COVID-19 replication. Multiple studies have been performed on the effectiveness of this drug as a treatment option, showing conflicting results. One study by Butler et al. (2023) involved 25,000 non-hospitalized patients with early-stage COVID-19 who were started on molnupavir. This treatment did not reduce the risk of hospitalization or death; however, there was a reduced time to recovery (9 vs. 15 days) [67]. Other studies showed similar results, suggesting molnupavir is an effective alternative treatment for COVID-19.

Nirmatrelvir/Ritonavir (Paxlovid): It is currently the preferred treatment choice against COVID-19. It is an antiviral agent (nirmatrelvir) combined with a cytochrome P-450 inhibitor (ritonavir), the latter of which increases the concentration of the former to target therapeutic range. The antiviral component functions by blocking the processing of the viral polypeptides, halting the replication of new virion particles similar to Lopinavir/ritonavir (Fig. 5). Many studies have been performed on its efficacy, demonstrating the following results:

- Hammond et al. (2022) conducted a randomized controlled trial with 2246 patients, where 1120 were assigned to the treatment group and 1126 to the placebo group. The results of the study demonstrated that treatment with nirmatrelvir-ritonavir resulted in a risk of progression to severe COVID-19 that was 89% lower than the risk from the placebo [68].
- Schwartz et al. (2023) conducted a population-based cohort study of adults who had a positive COVID-19 test between April 4th-August 31st, 2022, comparing patients who were treated with nirmatrelvir-ritonavir and those who were not. Of the 177,545 patients in the final cohort, even though only 5% (8,876) were treated with nirmatrelvir-ritonavir, there were still significantly reduced odds among these patients of hospital admission and death from COVID-19 [69].
- Dryden-Peterson et al. (2023) conducted a population-based cohort study to assess whether nirmatrelvir-ritonavir reduces the risk of hospitalization and death among outpatients with early-stage COVID-19. A total of 44,551 patients were included: 12,541 who received nirmatrelvir-ritonavir and 32,010 who did not. While the overall risk was already significantly low (1%) in this population, treatment with nirmatrelvir-ritonavir demonstrated to reduce this risk even further [70].

While this treatment has proven effective, there are significant drug interactions that may require dose adjustment or avoidance. Special care must also be taken in patients with kidney impairment, as this can lead to toxicity levels of the drug.

Many pharmaceutical companies worldwide have developed vaccines to counter the COVID-19 virus. Specifically in the United States, the three main developed vaccines include Pfizer-BioNTech (BNT162b2 or Comirnaty), Moderna (mRNA-1273 or Spikevax), and Janssen (also referred to as Johnson & Johnson; Ad26.COV2.S or Jcovden). The first two are mRNA vaccines, while the third one is an adenoviral vector vaccine. BNT162b2 is indicated for individuals 5 years or older, while mRNA-1273 and Ad26.COV2.S are indicated for individuals 18 years old and older [72, 73].
The mRNA vaccines are more favourable over Ad26.COV2.S due to the preferred risk-benefit ratio associated with the mRNA vaccines. Ad26.COV2.S has been associated with thrombosis, thrombocytopenia, and possible Guillain-Barre syndrome, while the mRNA viruses have only been associated with myocarditis. The benefits outweigh the risks for all the vaccines; however, the risks associated with the mRNA are deemed less severe \[74\]. There have also been many observational studies that determined vaccine effectiveness is higher from two doses of the mRNA vaccines compared to one dose of Ad26.COV2.S. A case-control study examined many immunocompetent adults who were hospitalized with COVID-19, and it was estimated that effectiveness against COVID-19-related hospitalization was 93% and 88% for mRNA-1273 and BNT162b2, respectively, in comparison to 71% for Ad26.COV2.S \[75\]. There are differences between the three vaccines in terms of immunogenicity (Table 6), efficacy (Table 7), and safety (Table 8) \[76-78\].

Numerous vaccines have been produced outside of the United States across the globe. These include University of Oxford, AstraZeneca, and the Serum Institute of India (ChAdOx1 nCoV-19/AZD1222 or Vaxzevria), Novavax (NVX-CoV2373, Covovax or Nuvaxovid), CanSino Biologics (AD5-nCOV or Convidecia), Gamaleya Institute (GamCOVID-Vac or Sputnik V), Sinopharm (BBIBP-CorV or Vero Cell), Sinovac (CoronaVac), Bharat Biotech/Indian Council of Medical Research (BBV152 or Covaxin), and Zyduz Cadila (ZyCoV-D). Each vaccine has been authorized for usage in at least one country.

### Table 6. The immunogenicity of BNT162b2, mRNA-1273, and Ad26.COV2.S.

<table>
<thead>
<tr>
<th>Immune Modality</th>
<th>BNT162b2 (Comirnaty)</th>
<th>mRNA-1273 (Spikevax)</th>
<th>Ad26.COV2.S (Jcovden)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immunogenicity</td>
<td>Demonstrated binding and neutralizing antibody response comparable to those in convalescent plasma of individuals aged 18-85 years who had an asymptomatic or moderate COVID-19 infection.</td>
<td>Demonstrated binding and neutralizing antibody response comparable to those seen in convalescent plasma with vaccination in healthy individuals aged 18-85 years.</td>
<td>Demonstrated binding and neutralizing antibody response in individuals aged 18-85 years after a single dose, although slightly lower than those in convalescent plasma.</td>
</tr>
<tr>
<td></td>
<td>Antibodies induced in individuals aged 12-15 years and 5-11 years (with a lower dose) were higher than in those aged 16-25 years.</td>
<td>Antibodies induced in individuals aged 12-15 years and 5-11 years (with a lower dose) were higher than in those aged 16-25 years.</td>
<td>Antibodies induced in individuals aged 12-15 years and 5-11 years (with a lower dose) were higher than in those aged 16-25 years.</td>
</tr>
<tr>
<td></td>
<td>Neutralizing antibodies levels generated are lower against Delta and even lower against Omicron.</td>
<td>Neutralizing antibodies levels generated are lower against Delta and even lower against Omicron.</td>
<td>Neutralizing antibodies levels generated are lower against Delta and even lower against Omicron.</td>
</tr>
</tbody>
</table>

### Table 7. The efficacy of BNT162b2, mRNA-1273, and Ad26.COV2.S.

<table>
<thead>
<tr>
<th>Efficacy</th>
<th>BNT162b2 (Comirnaty)</th>
<th>mRNA-1273 (Spikevax)</th>
<th>Ad26.COV2.S (Jcovden)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy</td>
<td>95% vaccine efficacy (95% CI 90.3-97.6) in preventing symptomatic COVID-19 at 7 days and onwards following the second dose in individuals aged 16 and older.</td>
<td>94.1% vaccine efficacy (95% CI 89.3-96.8) in preventing symptomatic COVID-19 at 14 days and onwards following the second dose in individuals aged 18 and older.</td>
<td>The regular single-dose regimen had a 66.9% vaccine efficacy (95% CI 59-73.4) in preventing moderate to severe COVID-19 at 14 days and onwards in individuals aged 18 and older.</td>
</tr>
<tr>
<td></td>
<td>Individuals aged 65 years and up with comorbidities or obesity observe an efficacy of 91.7% (95% CI 44.2-99.8).</td>
<td>Individuals aged 65 years and up observe an efficacy of 86.4% (95% CI 61.4-95.5).</td>
<td>Vaccine efficacy against severe cases trended higher at 76% and 85% after 14- and 28-days post-vaccination, respectively.</td>
</tr>
<tr>
<td></td>
<td>Individuals aged 12-15 years without evidence of prior infection observe an efficacy of 100% (95% CI 75.3-100).</td>
<td>Individuals aged ≤11 years without evidence of prior infection observe an efficacy of a lower vaccine dose of 91% (95% CI 68-98%).</td>
<td>The trial of two dose regimen showed preliminary efficacy rates against symptomatic and severe COVID-19 at 75% and 100%, respectively.</td>
</tr>
</tbody>
</table>

### Table 8. The safety of BNT162b2, mRNA-1273, and Ad26.COV2.S.

<table>
<thead>
<tr>
<th>Safety</th>
<th>BNT162b2 (Comirnaty)</th>
<th>mRNA-1273 (Spikevax)</th>
<th>Ad26.COV2.S (Jcovden)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety</td>
<td>Local and systemic side effects are common, especially after the second dose. Generally mild to moderate symptoms.</td>
<td>Local and systemic side effects are common, especially after the second dose. Generally mild to moderate symptoms.</td>
<td>Local and systemic side effects are common, especially after the second dose. Generally mild to moderate symptoms.</td>
</tr>
<tr>
<td></td>
<td>- Reported symptoms include reaction at the injection site, usually pain (in 56% after either dose), fatigue (20% after first dose vs. 48% after second dose), headache (25% vs. 40%), myalgia (17% vs. 37%), fever, chills and joint pain (20% after second dose).</td>
<td>- Reported symptoms include reaction at the injection site, usually pain (in 74% after the first dose vs. 82% after the second), fatigue (33% vs. 60%), headache (27% vs. 53%), myalgia (21% vs. 51%), fever and chills (40% after second dose), and joint pain (32% after second dose).</td>
<td>- Others reported anxiety-related symptoms, including tachycardia, hyperventilation, light-headedness, and syncope.</td>
</tr>
<tr>
<td></td>
<td>- Myocarditis and pericarditis have also been reported following receipt of the vaccine.</td>
<td>- Myocarditis and pericarditis have also been reported following receipt of the vaccine.</td>
<td>- It has been associated with a specific syndrome of thrombosis with thrombocytopenia, and possibly associated with Guillain-Barre syndrome.</td>
</tr>
</tbody>
</table>
ChAdOx1 nCoV-19/AZD1222 (Vaxzevria): This vaccine has a two-dose regimen and is based on a replication-incompetent chimpanzee adenovirus vector that expresses the spike protein. Its vaccine efficacy after the second dose is 70-76% (95% CI 54.8-80.6) at 14 days and onwards, although it is suspected to wane over time. It was proven ineffective against the Delta and Omicron variants. Fatigue, headache, and fever were common side effects after receiving the vaccine and were severe in up to 8% of recipients [79-90].

NVX-CoV2373 (Covovax or Nuvaxovid): This vaccine is a recombinant protein nanoparticle vaccine comprising trimeric spike glycoproteins and a Matrix-M1 adjuvant. In the United States and Mexico clinical trials, efficacy was found to be 90.4% (95% CI 82.9-94.6) in preventing symptomatic COVID-19 in individuals aged 18-84 years who have never been infected. A clinical trial in the United Kingdom found a lower but significant efficacy (82.7%, 95% CI 73.3-88.8) [91, 92].

AD5-nCoV (Convidecia): This vaccine has a single-dose regimen and is based on a replication-incompetent adenovirus 5 vector that expresses the spike protein. It was shown to have a vaccine efficacy of 57.5% (95% CI 39.7-70) for symptomatic infection and 91.7% (95% CI 36.1-99) for severe infection [93].

Gam-COVID-Vac (Sputnik V): This vaccine has a double-dose regimen and uses two replication-incompetent adenovirus vectors that express a full-length spike glycoprotein. It has an efficacy of 91.6% (95% CI 85.6-95.2) in preventing symptomatic COVID-19 at the time of the second dose. The most common side effects were local and systemic flu-like reactions [94, 95].

BBIBP-CorV (Vero Cell): Two inactivated, whole viruses based on two different SARS-CoV-2 isolates from patients in China, WIV04 and HB02, were selected as potential COVID-19 vaccines. Each had an aluminum hydroxide adjuvant and had a double-dose regimen. Vaccine efficacy was shown to be around 73% (95% CI 58-82) for WIV04 and 78% (95% CI 65-86) for HB02 [96]. The HB02 stain was shown for the later approved BBIBP-CorV vaccine [97].

CoronaVac: This inactivated COVID-19 vaccine has an aluminum hydroxide adjuvant and a double-dose regimen. Vaccine efficacy from a clinical trial in Turkey was shown to be 83.5% (95% CI 85.4-92.1); however, lower rates were reported from trials in other countries. In an observational study in Chile, effectiveness was estimated to be 70% for preventing COVID-19 and 86-88% for preventing hospitalization or death. Another study in Brazil reported lower effectiveness among adults aged 70 years and older, estimating 47% for preventing COVID-19, 56% for preventing hospitalization, and 61% for preventing death [98-101].

BBV152 (Covaxin): This inactivated COVID-19 vaccine has aluminum hydroxide and a toll-like receptor agonist adjuvant with a double-dose regimen. It has been shown to have a vaccine efficacy of 78% (95% CI 65-86) against symptomatic COVID-19. There have been no serious adverse events related to the vaccine except for one possible case of immune thrombocytopenic purpura [102].

ZyCoV-D: This is the first DNA COVID-19 vaccine. It is delivered through a needle-free device with a high-pressure stream. A clinical trial among 28,000 participants showed that efficacy against symptomatic COVID-19 was reportedly 67% following three doses; however, the trial details required for critical review of these results have not been made public [103].

The effectiveness of the COVID-19 vaccines has been demonstrated through multiple clinical trials and published reports. However, concerns have risen over the health risks associated with the use of synthetic mRNA in some of their development [104, 105]. Many studies have stated that the injection of a vaccine containing synthetic mRNA has no potential risk of integration into the host genome [106-109]. This is due to the fact that in cells, transcribed mRNA is transported across the nuclear membrane to the cytoplasm for translation into proteins. This transportation is facilitated by nuclear pore complexes, which are proteins embedded in the membrane that regulate the transportation of molecules in and out of the nucleus. It is this selective function that prevents mRNA from re-entering the nucleus once translation is complete and is the basis for the claims made concerning the safety of mRNA vaccines. Furthermore, the integration of mRNA into the host genome would also require reverse transcription to create complementary or cDNA. This process requires a reverse transcriptase, a DNA polymerase enzyme, which is not normally present in healthy somatic cells.

Other than retroviruses, such as the Human deficiency virus (HIV), the genome from viruses does not typically integrate into the infected host genome. However, documented cases have been reported. In the human genome, roughly 17% encodes for Long Interspersed Element-1 (LINE1) elements, which are a type of autonomous retrotransposons. In somatic cells, LINE1 is normally repressed as it can contribute to the development of disease and cancers. However, derepression can be seen in ageing cells, cancerous tissues, or virus-infected cells [110]. In the case of COVID-19 infection, studies have found evidence that large fractions of the viral RNA could be reversed, transcribed, and integrated into the genome of the infected cells, later being expressed as viral-host chimeric transcripts. This was achieved through the activation of LINE1 elements, acting as the reverse transcriptase, and was hypothesized to contribute to PCR-positive results being observed weeks after the initial infection. However, as only fractions of the genetic sequences were integrated into the host genome, full viral replication was prevented [111-112].

There have been limited studies examining the risks surrounding the genomic integration of vaccine-injected mRNA. One study in 2022 examined mRNA integration in vitro in a human liver cell line, Huh7, after vaccination with the Pfizer BNT162b2 vaccine. The results demonstrated increased LINE1 activity with significant
mRNA reverse transcription into cDNA [113]. However, a few points are important to note. Firstly, the cell line used for the study was a cancerous liver cell line. LINE1 activity was shown to increase in cancerous tissues, which limits the accurate representation of the general, healthy population. In fact, another study that observed COVID-19 genetic integration into the host genome of infected cells reported no integration when the viral mRNA was transfected into healthy cells. This is attributed to the increased LINE1 activation in the infected cells that would not be seen in the transfected ones [112]. Additionally, the paper did not identify positive host genome integration of the cDNA, only commenting on the fast uptake of the vaccine in the cells and intracellular reverse transcription. Further investigations are required to further explore the risk for vaccine-delivered mRNA integration into the host genome in the general population and the potential impact and health risks.

4. VIRUS VARIANTS, GENETICS, AND HOST GENETIC IMPLICATION IN INFECTION AND DISEASE SEVERITY

COVID-19 disease severity has been shown to be highly polygenic, with several studies linking multiple genetic variants with severity. The first strain was detected in Wuhan, China, in December 2019 and was called the L strain. By the end of 2020, the L strain had gone through multiple mutations, producing the S, V, and G strains. Throughout the pandemic, multiple variants of the original strain were identified and are described in Table 9. The five main variants that were monitored included Alpha, Beta, Delta, Gamma, and Omicron, with eight additional lesser monitored variants [114-117]. The most severe variant was the Delta variant, which was believed to be one of the most transmissible of the variants and led to more severe illness and hospitalizations. This was followed by the Alpha variant, which was determined to be more contagious than the original strain but with a lower severity of illness in comparison to the other variants. The beta and gamma variants were also believed to be more transmissible than the original strain, and they have an increased ability to evade immune responses. However, with the Alpha variant, the disease course was generally less severe. Finally, the Omicron variant was considered to have the least severe disease course with decreased hospitalizations and mortality compared to the other variants. Since its detection, it has been divided into multiple subvariants. In terms of vaccination efficacy per variant type, the developed vaccines had the highest efficacy against the original COVID-19 variants. However, additional vaccine boosters were shown to provide increased immunity for the newer variants, although these effects might only be temporary [118-121].

By early 2023, the COVID-19-associated death count in the United States alone was over 1.1 million. The federal COVID-19 Public Health Emergency (PHE) declaration ended soon on May 11, 2023, marking the official end of the pandemic. Currently, only the Omicron subvariants are variants of concern (VOC) that are still in circulation.

As there is variation in the severity of COVID-19 symptoms, one study sought to examine the molecular mechanisms that increase its severity. Results identified an association between the risk of severe disease and a multigene locus on chromosome 3p21.31 and the ABO blood group locus on chromosome 9q34.2 [122]. The HLA region did not show any association signal. Individuals with blood group A were determined to have an increased risk, while those with blood group O had a decreased risk. A strong associative signal was observed at the rs11385942 insertion-deletion GA or G variant at the locus on chromosome 3p21.31. The GA allele was observed to have a higher frequency among patients on mechanical ventilation than those who were only receiving supplemental oxygen. This could suggest that this risk allele causes a predisposition to the severe forms of COVID-19. There was no elevated risk-allele frequency observed for the ABO locus [123, 124].

The gene cluster on chromosome 3p21.31 has been identified as the major genetic risk locus for severe symptoms following COVID-19 infection. The genomic segment is a 50-kilobase locus and comprises multiple genes, including LZTFL1, SLC6A20, FYCO1, CCR9, CXCR6, and XCR1. The risk haplotype is inherited from Neanderthals and is carried by around 50% of people in South Asia, around 16% of people in Europe, and 9% of admixed American individuals carry at least one copy. The highest carrier frequency is seen in Bangladesh, where 63% of the population carries at least one copy, and 13% are homozygous for the risk haplotype. It is almost completely absent in the African population. Multiple susceptibility-related and cytokine regulation-related genes in immune cells have been assessed and described in Table 10. Also included is the pLI score, which is the probability of being loss-of-function intolerant, representing how much a single gene can tolerate mutational variants (e.g., frameshift, stop-gained, etc.). A pLI score of 0.9 or greater indicates the gene can be highly constrained, with mutational variants not frequently observed [125, 126].

Table 9. A list of the COVID-19 variants and molecular data.

<table>
<thead>
<tr>
<th>Variant type</th>
<th>First outbreak</th>
<th>Earliest detection</th>
<th>Associated mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epsilon (B1.427/B.1.429)</td>
<td>Southern California</td>
<td>July 2020</td>
<td>D1183Y, S131, W152C and L452R</td>
</tr>
</tbody>
</table>
Study of the COVID-19 Virus

<table>
<thead>
<tr>
<th>Variant type</th>
<th>First outbreak</th>
<th>Earliest detection</th>
<th>Associated mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambda (C.37)</td>
<td>Peru</td>
<td>December 2020</td>
<td>G75V, T76I, L452Q, F490S, D614G, and T859N</td>
</tr>
</tbody>
</table>

Table 10. Measurement of intolerance to loss-of-function variants of genes related to COVID-19 susceptibility, hyperinflammation, and severity. pLI scores indicate how constrained the gene is. The bolded genes are not intolerant to loss-of-function variants.
The genetic risk segment includes genes of several chemokine receptors, such as CCR9, CXCR6, and XCR1 (Table 10), which play critical roles in immune responses and pathogenesis. They can be classified into several functional groups, including “inflammatory”, “homeostatic”, and “dual function”. Inflammatory chemokines are upregulated until inflammatory conditions are met with leukocyte recruitment to inflamed tissues. Homeostatic chemokines are continuously expressed in lymphoids and other organs, mediating the homeostatic migration of various cell types. Dual-function chemokines overlap in both fields. Unlike inflammatory chemokine receptors, homeostatic and dual-function chemokine receptors are more restricted, only binding with up to two ligands [127].

Regarding the six candidate genes on chromosome 3p21.31, LZTFL1 is one of the most significant, with the rs11385942 variant located within it. The LZTFL1 gene encodes a protein involved in protein trafficking and primary cilia. In T lymphocytes, this gene participates in the immunologic synapse with antigen-presenting cells, such as dendritic cells. LZTFL1 encodes Leucine zipper transcription factor-like 1 that is associated with the immune synapse between an antigen-presenting cell or target cell and a lymphocyte, such as a T/B cell or natural killer (NK) cell. Ultimately, it was observed to modulate T-cell activation as well as increase IL-5 levels, which is a key roles in antigen presentation and inflammasome activation [128-130].

A genome-wide association study was performed for genetic variations that may identify mechanistic targets for therapeutic treatments against COVID-19. Additional associations have been shown on chromosome 12q24.13 (rs10735079, p=1.65 × 10⁻⁸) in a gene cluster encoding antiviral restriction enzyme activators (OAS1, OAS2 and OAS3), on chromosome 19p13.2 (rs2109069, p=2.3 × 10⁻¹²) near the gene encoding tyrosine kinase 2 (TYK2), on chromosome 19p13.3 (rs2109069, p=3.98 × 10⁻¹²) within the gene encoding dipeptidyl peptidase 9 (DPP9), and on chromosome 21q22.1 (rs2236757, p=4.99 × 10⁻⁹) in the interferon receptor gene IFNAR2. These genetic signals relate to key host antiviral defence mechanisms and mediators of inflammatory organ damage in COVID-19, both of which can be targeted for treatment [131].

The IFNAR2 and OAS genes are involved with innate antiviral defences, which are important in early disease. In contrast, DPP9 and TYK2 genes are associated with host-driven inflammatory lung injury, which impacts late, life-threatening COVID-19 infections. These four genes are described below:

**IFNAR2:** The increased expression of this gene was associated with reduced odds of severe COVID-19, which was in line with the beneficial role of type I interferons. This suggests a protective role against the virus. However, exogenous interferon treatment was not observed to reduce mortality in hospitalized patients, suggesting that its beneficial effect may be mediated during the early stages of the infection when the viral load is high [132].

**OAS gene cluster:** These genes encode enzymes that produce 2’5’-oligoadenylate (2-5A), which is a host antiviral mediator. Its role is to trigger an effector enzyme, RNase L, which degrades double-stranded RNA, a replication intermediate of COVID-19 and other coronaviruses. This offers a potential therapeutic treatment, as endogenous phosphodiesterase 12 (PDE-12) degrades 2-5A in the body. The use of PDE-12 inhibitors can, therefore, be used to augment OAS-mediated antiviral activity [133].

**DPP9:** This gene encodes a serine protease that is involved with many intracellular functions, including antiviral activities. Some examples are the cleavage of CXCL10, an antiviral signalling mediator, as well as having key roles in antigen presentation and inflammasome activation [134].

**TYK2:** This gene is a member of the Janus family and is involved with cytokine signaling and antiviral immunity. It has been chosen as one of four gene targets for JAK inhibitors, such as baricitinib, due to the association between TYK2 expression and critical illness [135].

Other important genes include SLC6A20, which is expressed in the gastrointestinal tract and encodes sodium-imino acid (proline) transporter 1 (SIT1), reportedly regulated by ACE2 receptors. This gives evidence as to why individuals infected with COVID-19 can experience gastrointestinal symptoms, such as nausea, vomiting, and diarrhea [136, 137]. FYCO1 encodes the FYVE and coiled-coil domain containing 1 (FYCO1), which is an autophagy adaptor protein. Studies have found associations between this region and clinical characteristics found in individuals with severe COVID-19 infections. This is due to the likelihood of it being a key mediator linking the primary site of coronavirus replication, ER-derived double-membrane vesicles, with the microtubule network in the host. This offers a potential target for therapeutic treatment against COVID-19, as downregulating expression of FYCO1 could

<table>
<thead>
<tr>
<th>Signal</th>
<th>Gene name</th>
<th>pLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytokine storm</td>
<td>IL-1β</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>IL-6</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>IL-8</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>IL-10</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>IL-18</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>IL-1RA</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>TNF</td>
<td>0.803</td>
</tr>
<tr>
<td></td>
<td>IFN-γ</td>
<td>0.472</td>
</tr>
</tbody>
</table>
potentially decrease replication and infectivity of the virus [138].

Many studies were performed at Mount Sinai Health System in New York City regarding the molecular profile of COVID-19. One study sought to explore the genetic variations seen in children presenting with multisystem inflammatory syndrome following a diagnosis of COVID-19 [139]. Multisystem inflammatory syndrome in children (MIS-C) presents with symptoms including fever, inflammation, and widespread pathology in multiple organs, similar to those seen in Kawasaki disease [140, 141]. The study involved RNA sequencing of 30 blood specimens obtained from MIS-C patients, pediatric COVID-19 cases, and healthy control genes while implementing different statistical approaches, including co-expression and Bayesian probabilistic-causal networks, to identify disease-associated genes. Among the MIS-C population, the results showed that there was a decrease in NK cells, cytotoxic (CD8+) T-cells, and a downregulated module of genes associated with mature CD8+ (Tex) cells and CD56dimCD57+ NK cells. CD8+ T cells are capable of regulating one another when confronted with a viral infection, and a depletion of circulating NK cells can lead to CD8+ T cell exhaustion. This exhaustion causes specific CD8+ T cells to have poor cytokine output, cytolytic activity, and inhibition of proliferative capacity. This can lead to severe T cell immunopathology following a viral infection, where, on the contrary, its presence can improve the symptoms [142, 143]. Nine key regulators were implicated in these findings, including TGFB3, TBX21, C10orf21, Sipr5, PRF1, MYBL1, KLRD1, SH2D1B, and GZMA. All of these genes have known associations with NK cells and CD8+ T cells, as well as illnesses similar to MIS-C [144, 145]. The conclusion indicated that an aberration in these aforementioned genes could lead to a decrease in NK cells with an accompanied lack in CD8+ T cell exhaustion, which may lead to severe inflammatory disease similar to MIS-C. TBX21 has also been determined to be the most promising therapeutic target of the nine genes due to its coding T-bet, which is a biological switch in the transition of Texdim to Texint during Tex differentiation [146]. However, future studies would be needed to determine why the resolution of a COVID-19 infection would lead to the development of MIS-C.

Another study conducted at Mount Sinai Health System evaluated the role of differing genes among the COVID-19 variants when developing rapid diagnostic tests. As new variants emerged as the pandemic progressed, there was an increased risk for “target dropout” due to the insufficient amplification of the target, leading to false negatives caused by primer/probe binding site (PBS) mismatches. One diagnostic test is the Agena MassARRAY® SARS-CoV-2 Panel, which utilized probes for five targets across the N and ORF1ab genes, allowing for a broad platform to accommodate for PBS mismatches. The study used this diagnostic test to determine target results and sequence data for 1262 positive cases of COVID-19. Not surprisingly, the PBS mismatches were higher in specimens with target dropout. It was also observed that among specimens with N3 target dropout, 57% contained an A28095T substitution that is very specific for the Alpha variant. Furthermore, redundancy in target design can be beneficial in preventing false negatives, and understanding target performance can help explain the activity of the numerous COVID-19 variants [147].

There was limited information on COVID-19 when it was introduced to the United States, as most data were concerned about its spread. Respiratory pathogen-negative nasopharyngeal specimens from 3,040 patients were obtained from the Mount Sinai Health System during the first ten weeks of 2020 for a study attempting to understand the viral origin and its transmission better. These samples were negative for diagnostic molecular amplification testing for routine respiratory pathogens. They were evaluated for the presence of two COVID-19 viral markers, the SARS-CoV-2 specific ORF1ab genes and the pan-Sarbecovirus envelope E gene. Results from this study indicated that the COVID-19 infection was present in NYC at least 6 to 8 weeks before the first official wave of the pandemic and was overlooked as another type of viral respiratory infection [148].

One other study at Mount Sinai Health System attempted to understand the molecular etiology behind the post-acute sequelae of the COVID-19 infection. Blood samples were obtained for whole blood RNA sequencing and serology from 567 patients who were being followed into the post-acute period. The study sought to determine if the relationship between acute phase COVID-19 and the development of post-acute sequelae could be observed in blood gene expression. In addition to the RNA sequencing, cell-type specific changes in gene expression were assessed for association with the commonly observed symptoms of the infection. Results showed at least two etiologies for different sets of symptoms, one dependent and one independent of the antibody response. Processes leading to post-acute sequelae had already begun during patient hospitalization, indicating that study designs often missed an important time period to explore the pathogenesis of post-acute sequelae. Additionally, the molecular processes leading to post-acute sequelae were also not simply explained by acute severity. In regard to inflammatory cell response, plasma cells were found to play a key role in this process. The downregulation of genes associated with antibody production that also play a functioning role in pulmonary symptoms were independent of the antibody titers, whereas the upregulation of genes involved in the same processes causing other COVID-19 symptoms (i.e., sleep problems, nausea, skin rash, etc.) were highly dependent on the antibody response. The latter suggests these symptoms are linked to the immune system’s response to the virus [149].

A final study at Mount Sinai Health System was performed using autopsy tissues to examine the molecular profiling of COVID-19-induced damage further. Rapid
Polygenic risk scores (PRS) with COVID-19

Polygenic risk scores (PRS) have been widely applied in clinical studies investigating genetic variants associated with complex diseases that have a polygenic architecture. They are particularly useful in discerning the association between genetic scores and disease status, especially in cohorts where there is a higher prior probability of disease. For example, they can be used to assist in diagnosis or to inform treatment choices. A polygenic risk score is calculated by computing the sum of risk alleles per individual, weighted by the risk allele effect sizes as estimated by a genome-wide association study (GWAS) on the phenotype. It is described as a single-value estimate of an individual’s genetic liability to a phenotype. It can be calculated as the sum of their genome-wide genotypes and weighted by corresponding genotype effect size estimates derived from the GWAS summary statistic data (Fig. 11) [151, 152].

$$\text{PRS} = \sum_{i=1}^{m} \beta_i \left( \sum_{j=0}^{2} \omega_{ij} \times j \right)$$

Fig. (11). Equation to calculate polygenic risk scores.

Polygenic risk scoring is one approach to associating the risk of severe outcomes in genome-wide association studies that have clearly shown that common complex disorders have a polygenic architecture. This has enabled researchers to identify genetic variants associated with diseases. The variety of host genetic variants that determine susceptibility to COVID-19 infection and severity can be combined in a PRS. Studies have been conducted to compute genetic risk scores in different populations to determine high-risk groups. One study evaluated how well a genetic risk score based on chromosomal-scale length variation and machine-learning classification algorithms can predict the severity of a COVID-19 infection. Three groups of patients with severe COVID-19 infections were selected, one including patients less than 90 years old, the second less than 80 years old, and the third less than 70 years old. These patients were compared against age-matched people deemed normal (i.e., not infected). A significant difference was observed between the three age groups and their corresponding controls. This indicates that the germ line genetics of the infected patients is correlated with COVID-19 severity. However, while it was shown that genetics play a significant role, it is still too early to develop a useful genetic test to predict severity due to limited data [153, 154].

Another study analyzed PRS and history of chronic obstructive pulmonary disease (COPD) in patients and their association with severe COVID-19 disease. The PRS for the participants was calculated using 112 single nucleotide polymorphisms that were believed to be related to COVID-19 disease severity. The history of COPD was factored into the overall risk. The study then used logistic regression models to examine associations of genetic risk with or without COPD and the severity of COVID-19 disease. About 712 of the 430,582 participants in the study developed severe COVID-19, and 19.8% of these patients had pre-existing COPD. When examining the PRS, those with lower risk scores were observed to have a lower chance of developing severe COVID-19 when compared with others who had intermediate or high PRS [155].

A third study examined the causality of specific coagulation factors, such as D-dimers, prothrombin time (PT), von Willebrand factor (VWF), platelet count, and fibrinogen in the incidence of COVID-19 severity. To investigate the causal relationships, SNPs associated with 12 coagulation factors were selected, including VWF, ADAMTS13, tPA, FX, PAI-1, D-dimer, FVII, FVIII, FXI, aPTT, ETP, and PT. Results showed that genetic predisposition to the antigen level of VWF and the activity level of its cleaving protease, ADAMTS13, were causally associated with the incidence of COVID-19 severity. However, their effects were displayed in opposite directions, with plasma VWF antigen level showing a positive association while ADAMTS13 displayed a negative association. Furthermore, the statistical significance of ADAMTS13 could not always be reproduced with repeated tests, which weakens the association. The predictive ability of PRS derived from the VWF-associated genetic variants was then explored together with other critical risk factors, including age, sex, body mass index (BMI), coronary artery disease (CAD), systolic blood pressure (SBP), type 2 diabetes mellitus (T2DM), and COPD. It was determined that PRS of VWF was an independent risk factor in distinguishing severe COVID-19 infections from healthy controls, with an observed 16% higher risk.

Furthermore, when investigating the contribution of VWF PRS in relation to the other critical risk factors, age was determined to be the most important risk factor for COVID-19 severity. However, VWF PRS showed a larger normalized effect size than SBP, which emphasizes that while it is not the most important factor, its predictive value is still significant in determining prevention and a personalized treatment plan. Therefore, monitoring plasma levels of VWF can assist with developing strategies...
to control severity as well as associated thrombotic complications [156].

CONCLUSION
The COVID-19 virus has been shown to have significant complexity, eliciting symptoms of differing severity from person to person. This increased complexity led to the development of multiple different treatment types, all with varying levels of success. The previously performed studies surrounding the virus have shown that while further research is required, exploring the molecular makeup across different populations can assist in better understanding pre-existing risks for severe COVID-19 infections and their prognosis. This can further help with developing treatment and targeted therapies, combat the disease more efficiently, and continue to help move past this pandemic worldwide.

CONSENT FOR PUBLICATION
Not applicable.

FUNDING
None.

CONFLICT OF INTEREST
The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS
Declared none.

REFERENCES


http://dx.doi.org/10.1001/jama.2020.8259

http://dx.doi.org/10.1056/nejmoa2015432 PMID: 32374370


http://dx.doi.org/10.1126/science.abb4499 PMID: 32321856


http://dx.doi.org/10.1111/eci.13554


http://dx.doi.org/10.1038/s41598-020-6049-y PMID: 32370560


http://dx.doi.org/10.1038/s41591-020-1051-9


http://dx.doi.org/10.1001/jama.2020.5114

http://dx.doi.org/10.1056/nejmoa2015301

http://dx.doi.org/10.1056/nejmoa2015432 PMID: 32374370


http://dx.doi.org/10.1097/CCM.000000000003406 PMID: 30247273


http://dx.doi.org/10.1001/jama.2020.57975

http://dx.doi.org/10.1016/j.jcm.2020.07.002 PMID: 32569214


http://dx.doi.org/10.1111/cmi.2020.04.012 PMID: 32304745


http://dx.doi.org/10.1111/jce.13554

http://dx.doi.org/10.1016/j.jcm.2020.07.002 PMID: 32569214


http://dx.doi.org/10.1111/jce.13554 PMID: 33768536

http://dx.doi.org/10.1056/nejmoa2015301

http://dx.doi.org/10.1056/nejmoa2015432 PMID: 32374370


http://dx.doi.org/10.1097/CCM.000000000003406 PMID: 30247273

Study of the COVID-19 Virus


[81] Interim recommendations for use of the AZD1222 (ChAdOx1-S recombinant) vaccine against COVID-19 developed by Oxford University and AstraZeneca. 2022. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-vaccines

http://dx.doi.org/10.1016/S0140-6736(20)32661-4 PMID: 33506989

http://dx.doi.org/10.1038/s41586-021-03777-9 PMID: 34237773

http://dx.doi.org/10.1016/S0140-6736(20)31604-8 PMID: 32702298

http://dx.doi.org/10.1056/NEJMoa2116185 PMID: 34910859

http://dx.doi.org/10.1093/cid/ciac083 PMID: 36210481

http://dx.doi.org/10.1016/S0140-6736(21)02753-7 PMID: 34953526


http://dx.doi.org/10.1073/pnas.2105968118 PMID: 34774196

http://dx.doi.org/10.1038/d41586-021-02385-x PMID: 34475553

http://dx.doi.org/10.3390/genes13050719 PMID: 35627104


http://dx.doi.org/10.1016/j.lrha.2023.100487 PMID: 37155483

http://dx.doi.org/10.1016/S0140-6736(21)02009-6 PMID: 34774196

http://dx.doi.org/10.1001/jama.2021.8565 PMID: 34037666

http://dx.doi.org/10.1016/j.vaccine.2022.02.039 PMID: 35193703

http://dx.doi.org/10.1016/S0140-6736(21)01429-X PMID: 34246358


DISCLAIMER: The above article has been published, as is, ahead-of-print, to provide early visibility but is not the final version. Major publication processes like copyediting, proofing, typesetting and further review are still to be done and may lead to changes in the final published version, if it is eventually published. All legal disclaimers that apply to the final published article also apply to this ahead-of-print version.