Background: The study was conducted in the various districts of Punjab, Pakistan, to ascertain the incidence and epidemiology of SARS-CoV-2 variants circulating in the population during the fifth wave of COVID-19.

Method: A total of 9603 nasopharyngeal swab samples of suspected patients were collected from the different districts of Punjab from December 2021 to April 2022. In the BSL-3 facility, an auto-extractor (Uni-medica) was used to extract the RNA genome. Viral detection and quantification were performed using real-time reverse transcriptase PCR. Multiplex PCR was used to target different mutations of the spike protein in order to identify SARS-CoV-2 variants.

Results: 711 samples were found to be positive from a total of 9603. The Omicron variant of concern (VOC) was the predominant lineage of SARS-CoV-2 circulating at the time of sampling. Ninety-one percent proportion of COVID-19 was caused by the Omicron, followed by the wild variant (3.80%) and Delta (68.11%).

Men were found to have a greater Omicron prevalence (47.96%) than women (42.05%). Furthermore, compared to older individuals (32.07%) and younger kids (10.55%), adults had the higher percentage of Omicron (47.39%).

Conclusion: This study brought attention to the Omicron variant’s circulation in Punjab Province, Pakistan, during the COVID-19 fifth wave.
Introduction

A major issue has arisen with the COVID-19 pandemic caused by SARS-CoV-2 virus, leading to an increase in morbidity and mortality. So far, 704,507,621 people have been infected, and 7,008,354 have died globally. As of March 28, 2024, the virus has infected 1,581,936 people in Pakistan, causing 30,664 fatalities throughout the pandemic [1]. Pakistan observed the 5th SARS-CoV-2 wave in December 2021, driven by Omicron and peaked in late January with around 8,000 cases per day [2].

Punjab—a province of the Islamic Republic of Pakistan—is the most populous of the five provinces. Currently, the population of Punjab province is 109.9 million, with an average growth rate of 2.13%[3]. Of the 1.5 million COVID-19 cases that have been documented in Pakistan, 507,771 cases are reported in Punjab, accounting for 33.8% of total cases reported across Pakistan [4]. The community transmission of COVID-19 cases reported across Punjab is 91% [5]. In Pakistan, the sero-prevalence of SARS-CoV-2 ranges between 31.1% and 48.1%. Punjab, among all provinces, has the highest seroprevalence rate of 44.5% [6].

The coronavirus (CoV) is a member of the order Nidovirales within the family Coronaviridae. It is a positive sense RNA virus that has capsid and single-stranded. It has the largest genome size among all the RNA viruses having 27-32 kilobases [7] [8]. It has infected a wide range of hosts among birds and mammals (like turkeys, chickens, mice, rats and human beings) and caused upper to lower respiratory tract infections and gastrointestinal tract (van der Hoek et al., [9]).

The SARS-CoV-2 is responsible for the lethal respiratory outbreak and is genetically identical to SARS-CoV (nearly 79% resemblance) with enhanced transmissibility and pathogenicity [10,11]. But the most important structural protein, i.e. spike (S) protein, differs slightly between these viruses [12]. The spike protein on the virus’s surface binds with the ACE 2 receptor. It plays a role in determining the host range as well [13]. The S glycoprotein of the virus has two subunits: the S1 and the S2. The former binds to the host cell’s entry receptor (ACE2) later assisting in the virus’s and host cell’s membrane fusion [14]. Being an RNA virus, SARS-CoV-2 is highly prone to mutations. RNA viruses have been reported to have the highest rate of mutations ranging between $10^4 - 10^6$ mutations per base pair [15]. These mutations can change the function of proteins involved in antigenicity, infectivity, transmissibility and pathogenicity. The different mutations resulting in various strains may evade the vaccine response [16].

There are now thirteen different SARS-CoV-2 strains. On the other hand, the Centers for Disease Control and Prevention have designated the Alpha, Beta, Gamma, Delta, and Omicron variants as variant of concern (VOC) [17]. When it comes to genomic surveillance SARS-CoV-2 variant, the major focus is on the spike glycoprotein, as it is responsible for virus attachment with cells and is targeted by antibodies for neutralization [17].

The first reports of the SARS-CoV-2 Delta strain (B.1.617.2) came from India in October 2020. The transmission rate of the Delta Variant is higher than that of the wild-type strain. and has higher viral loads in SARS-CoV-2 infected patients [18]. In November 2021, Omicron, another variant of SARS-CoV-2, was identified in South Africa and has been identified as the fifth variant of concern (VOC) [19]. The alarming thing about this VOC was the high number of mutations, mainly in spike protein. The pieces of evidence of a surge in transmission capability and ability to elude neutralizing antibodies are also accumulated. Omicron outpaced the previously dominating variant of concern (Delta variant) in a short time [20]. The prominent mutations reported in S protein in the Omicron variant include E484A, T478K, K417N, del69-70 and P681H [21].

The goal of the present study is to ascertain the frequency of SARS-CoV-2 variants of concern during the COVID-19 pandemic’s fifth wave. This study intended to monitor which variant of concern was dominant during 5th wave of COVID-19 across Punjab province, Pakistan.

Methods

Study Design:

Among suspected COVID-19 patients in Punjab province, this is a prospective observational study evaluating the relationship between SARS-CoV-2 variant, age, and gender. Achieving patient confidentiality during data collection and getting informed consent for sample collection and processing were ensured for ethical considerations.

Data Collection:

The demographic dataset of the patients was collected at the time of sampling. The demographics included age and gender to observe the relation between independent variables (gender and age) and variants of concern of SARS-CoV-2

Sampling:

Nasopharyngeal samples from 9605 suspected cases of COVID-19 were collected from various districts in the province of Punjab. The samples were processed at the BSL-3 Laboratory for Emerging Pathogens, which is
part of the Institute of Microbiology at the University of Veterinary and Animal Sciences in Lahore. The Raosoft sample size calculator was used to determine the sample size. With a margin of error of ±1% and a 95% confidence level, the total estimated size was 9603. Samples were drawn throughout a five-month period, from December 2021 to April 2022, from primary clusters since the beginning of the fifth wave.

Detection of Viral RNA:
RNA was extracted from the samples using Nucleic Acid Extraction Reagent (Uni-Medica) per the manufacturer’s instructions [22]. In order to identify SARS-CoV-2, Novel Coronavirus 2019-nCoV (Uni-Medica) kit that covered three genes (ORF1ab, N, and Internal Control) was used [23]. The Cycle threshold value of each fluorophore was used to determine whether their corresponding targets were tested. The amplification curve of each fluorophore was S-shaped, and the cycle threshold value was ≤ 38. The samples diagnosed with SARS-CoV-2 were separated and preserved at -80°C in a small zipper bag [24].

Determination of SARS-CoV-2 Variants of Concern:
For the identification of variants of concern, Multiplex PCR was executed, and various mutations (E484K, N501Y, HV 69-70del, P681R, L452R, and K417N) were targeted (Zhao et al. 2021). The extracted viral RNA was used for variant identification using multiplex PCR. The Novaplex SARS-CoV-2 Variant IV assay (Seegene) and the Allplex SARS-CoV-2 Variant 1 assay were employed. The RdRp gene, Endo IC, and three mutations (E484K, N501Y, and HV 69/70del) are the targets of the Allplex SARS-CoV-2 Variant 1 Assay. Three distinct mutations (P681R, L452R, and K417N) are the focus of the Novaplex SARS-CoV-2 Variant IV Assay in addition to internal control. Several combinations of mutations were found in order to identify SARS-CoV-2 variants of concern, as indicated in Table 1.

<table>
<thead>
<tr>
<th>Variants of Concern</th>
<th>RdRp gene</th>
<th>Mutation in spike protein</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E484K</td>
<td>N501Y</td>
</tr>
<tr>
<td>Alpha</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Beta</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Gamma</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Delta</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Delta Plus</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Omicron</td>
<td>✓</td>
<td>X</td>
</tr>
</tbody>
</table>

✓: Indicates the presence of the specified mutation in the corresponding gene or spike protein.
X: Indicates the absence of the specified mutation in the corresponding gene or spike protein.

Table 1: Interpretation of the results for variants of SARS-CoV-2. This table provides information on specific mutations found in the RdRp gene and the spike protein of various variants of concern (Alpha, Beta, Gamma, Delta, Delta Plus, and Omicron) associated with SARS-CoV-2 [25].

Statistical Analysis:
The data included COVID-19-positive samples, SARS-CoV-2 variants, age, and gender. These were first input into a Microsoft Excel spreadsheet, from which plots were generated and associations were examined using the Chi-square test using Minitab statistical tool. Data are available in supplement file.

Results
SARS-CoV-2 variant prevalence during the COVID-19 fifth wave:

![Figure 1: Percentage of SARS-CoV-2 mutations of concern during the COVID-19 fifth wave.](image)

Out of 9603 COVID-19 suspected samples processed, 711 were positive. Among 711 positive samples separated, the Omicron variant was prevalent (90.01%). Delta variant was observed at 6.188%, followed by the wild variant (the virus originated from Wuhan, China), 3.798% of total COVID-19-positive samples as shown in Fig. 1. While no other variant of concern was observed during this period of 5th wave of COVID-19.

Gender-wise distribution of Variants of Concern during the fifth wave of COVID-19:
According to the results, the prevalence of the Omicron variant was high in males (47.96%) compared to females (42.05%) of COVID-19-positive samples. The Delta variant was equally prevalent in males and females (3.09%) as shown in Fig 2. The wild variant has a high percentage of 2.53 in males compared to females during the fifth wave. The p-value for the association between gender and variants of concern is more than 0.05, which means that the incidence of COVID-19 is not gender-based.

<table>
<thead>
<tr>
<th>VOC SARS-CoV-2</th>
<th>F</th>
<th>M</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta</td>
<td>22</td>
<td>22</td>
<td>44</td>
</tr>
<tr>
<td>Omicron</td>
<td>299</td>
<td>341</td>
<td>640</td>
</tr>
<tr>
<td>Wild Variant</td>
<td>9</td>
<td>18</td>
<td>27</td>
</tr>
<tr>
<td>All</td>
<td>350</td>
<td>381</td>
<td>731</td>
</tr>
</tbody>
</table>

Table 2: Gender-specific Incidence of SARS-CoV-2 VOCs in the 5th COVID-19 Wave.
Age-based analysis of COVID-19’s fifth wave’s variant distribution concerning:

The patients were grouped as teenage (9-19 years), adults (20-39 years) and elderly (40-79 years). The age group with the least %age of reported variants of SARS-CoV-2 was the teenage ones during the fifth wave. The observed variants during the fifth wave were Omicron, Delta and Wild variants (the virus originated from Wuhan, China). The %age of these variants in the teenage group was 10.54 %, 1.27 % and 0.42 % of COVID-19 positive samples, respectively. Among the elders, the Omicron was reported to be more dominant compared to other observed variants, with 32.07 % of COVID-19-positive samples, while Delta and Wild variants were observed to be 1.82 % and 1.42 %, respectively. The highest %age (47.40 %) of Omicron was reported in adults, while Delta and Wild variants were reported to be 3.09 % and 1.40 %, respectively as shown in Fig 3. The p-value for the association between age and observed variants of concern is more than 0.05, which means that the result is insignificant. Thus, the age factor does not affect the incidence of COVID-19 disease.

Discussion

A 711 COVID-19-positive samples provided critical insights into the demographic factors, including the age and gender of people infected with strains that were in circulation in Pakistan during the fifth wave. We draw attention to the Variants of concern (VOCs) that were still being transmitted. The variant with the highest number of cases is the Omicron variant, which was being transmitted between November 20, 2021, and February 02, 2022. This demonstrates the omicron variants’ high transmissibility, corroborating reports from around the world that this quickly spreading VOC is becoming the predominant variant in nations like the USA, UK, and others [26,27].

Omicron was the most mutated variant that had affected the transmissibility of the virus, severity of disease and immune evasiveness [28]. Following a spike in SARS-CoV-2 positive samples, which resulted in the fifth wave of COVID-19 in Pakistan from December 2021 to March 2022, the first case of Omicron was identified on December 13, 2021. According to this study, the Omicron strain of SARS-CoV-2 predominated during the COVID-19 fifth wave, with 90.01 % of COVID-19-positive samples followed by some cases (3.79 %) reported because of the wild variant and 6.19 % cases were reported because of Delta variant. In a previous study, the data showed a surge in the number of cases since the mid of December 2021 because of the Omicron variant. During the fifth wave of the pandemic in Pakistan, this variant was linked with an increase in daily infections. The fatalities reported during the fifth wave of pandemic lessened when Omicron was documented as a variant of concern. The reason behind this might be the higher vaccination rate compared to previous waves, pre-existing immunity and intrinsic feature of the virus [29].

The higher COVID-19 cases during the fifth wave were reported because of the three important substitutions in the virus’s Receptor Binding Domain (RBD). The reported substitutions were HV69-70del, K417N and N501Y. These substitutions increased the binding affinity of the virus with the host receptor ACE2 and increased the antibody escaping capability of the virus [30]. The mutation N501Y has been reported in the receptor-binding domain and is linked to the surge in the binding affinity of the virus to the ACE2 receptor of host cells [31,32]. Furthermore, due to their association to escape from the humoral immune response, the spike mutations HV69-70del and K417N are labelled potential escape mutations Sabino et al. [33],[34].

The studies conducted in the recent past revealed that in some regions across the globe, the incidence of COVID-19 was not gender-based [35,36]. However, in
some regions, the prevalence of COVID-19 was significantly higher in males than in females [37,38]. This study revealed that during the fifth wave of COVID-19 in Pakistan, the Omicron variant was most prevalent in males (49.4 %) compared to females (44.2 %). But this study was not aligned with another study in which 60 % of Omicron was reported in females compared to 40 % in males. This might be due to the small sample size of that study [39]. Another study conducted in Denmark revealed that a higher % of Omicron is present in females (51.1 % as compared to males (48.1 %), which is not aligned with the results of this study [40]. In a previous study in Canada, it was reported that the Omicron variant was more dominant in females (50.3 %) as compared to males (49.7 %) during the fifth wave [41]. The reason for the higher %age of Omicron variant in males compared to females is that the ACE2 receptors are present in abundance in the lungs and heart of males compared to females, making them more exposed to SARS-CoV-2 [42]. Males have higher levels of immunosuppressive steroid hormones than females, which is another reason for the higher infection rate in men Klein et al. [43]. As a result, gender is a risk factor for increased susceptibility towards SARS-CoV-2. The gender component, as well as the higher rate of drinking and smoking habits in males, led to higher rates of COVID-19 in males compared to females [44].

This study also revealed that during the fifth wave, the highest %age of Omicron was reported in the age group 20–39 years old followed by the group 40–79 years old and the 9–19 years old patients. The association between the SARS-CoV-2 variant that warrant concern and age revealed that incident of VOC is not linked to any age group. Another study revealed that the highest % of the Omicron variant was reported in age group 21-49 Fonager et al. [40]. These results are associated with the current study. In the study conducted in Ontario, Canada, the highest %age of Omicron was reported in the age group of 21-44 years, which is nearly associated with the result of this study [41].

Compared to other molecular techniques, the real-time polymerase chain reaction, or RT-PCR, is less expensive, including NGS (next-generation sequencing). This molecular technique can be applied in most nation-states, especially third-world countries, as a screening method to detect these mutations in the spike protein of SARS-CoV-2 that have biologically pathogenic consequences. The RT-PCR may have several epidemiological advantages compared to existing national policies, which rely on samples sent to certain institutions for sequencing. This molecular technique was implied in the past to observe the viral changes that are clinically significant such as those reported in the Influenza virus [45-47].

The study’s findings showed that during the COVID-19 fifth wave in Pakistan’s Punjab province, the Omicron variant of concern predominated. Males were found to have a higher percentage of cases attributed to the Omicron variant than females. The age group of 20–39 reported the highest percentage of Omicron, followed by the 40–79 age group and the 9–19 age group. RT-PCR is a molecular approach that can be used as a screening tool to identify SARS-CoV-2 spike protein mutations that have biologically harmful implications.

Author Contributions
Muhammad Mubashar Beig and Muhammad Asad Ali designed the experiments. Adnan Mehmood, Hamza Raza, Awais Masud, Rida Basharat, Nadia Mukhtar and Waqar Aziz helped in the data analysis and interpretation. Muhammad Mubashar Beig also wrote the main manuscript text and prepared figures. Hasnain Javed and Sohail Raza arranged the PCR kits. Aftab Ahmad Anjum, Muhammad Adnan Ashraf, Muhammad Abu Bakar Shabbir, Muhammad Asad Ali, Muhammad Imran Najeeb, Asim Khalid Mehmood, and Muhammad Nawaz reviewed the manuscript. Tahir Yaqub and Muhammad Asad Ali supervised the work and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest
The authors declare that there is no conflict of interest regarding the publication of this paper.

References
27. You're reading

Geospatial Dynamics of SARS-CoV-2 Variants during the Fifth Wave of COVID-19 in Pakistan


45. Mai PHV, Hong TUT, Le Khanh HN, Thanh TN, Le Thi T, et al. Missed detections of influenza A (H1) pdm09 by real-time RT-PCR assay due to haemagglutinin sequence mutation, December 2017 to March 2018, northern Viet
Nam. Western Pacific surveillance and response journal (2019); 10(1): 32.
