COVID-19 Case Numbers as a Function of Regional Testing Strategy, Vaccination Coverage, and Vaccine Type

Areg A. Totolian  
Saint Petersburg Pasteur Institute

Viacheslav S. Smirnov  
Saint Petersburg Pasteur Institute

Alexei A. Krasnov  
Saint Petersburg Pasteur Institute

Edward S. Ramsay  
Saint Petersburg Pasteur Institute

Vladimir G. Dedkov (vgdedkov@yandex.ru)  
Saint Petersburg Pasteur Institute

Anna Y. Popova  
Federal Service for Supervision of Consumer Rights Protection and Human Welfare

Research Article

Keywords: COVID-19, SARS-CoV-2 vaccines, SARS-CoV-2 testing coverage, COVID-19 prevalence, SARS-CoV-2 vaccination coverage, SARS-CoV-2 herd immunity

Posted Date: October 20th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2183670/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Introduction

The COVID-19 pandemic that began in 2019 has become a serious challenge for humanity almost everywhere globally. Despite active vaccination around the world, prevalence in different countries varies significantly as of May 2022. The reason may be a combination of demographic, immunological, and epidemiological factors.

The purpose

of this study was to analyze the relationship between COVID-19 prevalence in the population and the types of SARS-CoV-2 vaccines used in different countries globally, taking into account demographic and epidemiological factors.

Materials and methods

An initial database was created of demographic and immunoepidemiological information about the COVID-19 situation in 104 countries, collected from published official source and repository data. The baseline included for each country: population size and density; SARS-CoV-2 testing coverage; vaccination coverage; prevalence; as well as a list of vaccines that were used, including their relative share among all vaccinations. Subsequently, the initial data set was stratified by population and vaccination coverage. The final data set was subjected to statistical processing both in general and taking into account population testing coverage.

Results

After formation of the final data set (including 53 countries), it turned out that reported COVID-19 case numbers correlated most strongly with testing coverage and the proportions of vaccine types used, specifically: mRNA (V1); vector (V2); peptide/protein (V3); and whole-virion/inactivated (V4). Due to the fact that an inverse correlation was found between ‘reported COVID-19 case numbers’ with V2, V3 and V4, these three vaccine types were also combined into one analytic group, ‘non-mRNA group’ vaccines (Vnmg). When the relationship between vaccine type and prevalence was examined, minimum prevalence was noted at V1:Vnmg ratios (%:%) from 0:100 to 30:70. Maximum prevalence was seen with V1:Vnmg from 80:20 to 100:0. On the other hand, we have shown that the number of reported COVID-19 cases in different countries largely depends on testing coverage. To offset this factor, countries with low and extremely high levels of testing were excluded from the data set; it was then confirmed that the largest number of reported COVID-19 cases occurred in countries with a dominance of V1 vaccines. The fewest reported cases were seen in countries with a dominance of Vnmg vaccines.

Conclusion

In this paper, we have shown for the first time that the level of reported COVID-19 prevalence depends not only on SARS-CoV-2 testing and vaccination coverage, which is quite logical, but also on the vaccine types used. With the same vaccination level and testing coverage, those countries that predominantly use vector and whole-virion vaccines feature prevalence that is significantly lower than countries that predominantly use mRNA vaccines.

1. Introduction

Despite large-scale control measures, both medical and non-medical, COVID-19 prevalence in the 3rd year of the pandemic remains high. As of May 13, 2022, total accumulated prevalence amounted to 140.5 million people [1]. The daily number of confirmed COVID-19 cases as of May 13–14, 2022 (the data collection period for this study) was 32.53 cases per 1 million population in Russia. In other regions, the values (new cases per day per million) were: 509.1 in France; 437 in Finland; 345.2 in Belgium; and 439.5 in the European Union (overall). On the other hand, 62.3 cases were noted in Lithuania, but only 19.4 in Turkey, and no new illnesses were detected in Kyrgyzstan or Tajikistan during this period [2]. Although the reference states “Due to limited testing, the number of confirmed cases is lower than the true number of infections,” there is nevertheless a significant gap in the daily number of confirmed cases in different countries. There is reason to believe that there may be several reasons for such significant differences.

A. Sociodemographic factors

In such situations, assigning primary importance to population would be reasonable. It could be assumed that highly populated countries would feature higher prevalence. In the pandemic, however, this factor was not so significant. For example, the population of Russia was 145,478,097 (as of May 13, 2022), while that of France was 67,390,000 [3]. In France, however, the share of daily increase in patients in mid-May was 15.6-fold higher than in Russia.
Population density could be another factor. It is believed that the greater number of contacts within a high-density region can become a factor contributing to more active spread of the pathogen in the population [4]. However, this indicator turned out to be ambiguous. As an example, the neighboring countries of Lithuania and Latvia feature population densities of 40.6 /km² and 28.7 /km², respectively. However, the daily increase in patients in Lithuania was 62.3 people per 1 million, and Latvia's was 126.0 people per 1 million [2].

The third reason for spread of a pathogenic virus can be population mobility. In this regard, it is difficult to give any quantitative estimates, but it has been repeatedly proven that the reason for the rapid spread of SARS-CoV-2 was the active movement of tourists between different countries. In this light, severe restrictive measures were used (border closures, air travel bans between countries) at the beginning of the pandemic in most countries.

Commitment to vaccination has become important in limiting the pandemic spread of SARS-CoV-2 [5, 6]. In parallel with introduction of vaccines against human coronavirus (developed at an unprecedented speed), a fairly strong anti-vaccination movement formed that actively opposed the vaccination campaign. As usual, significant time and effort was required to counter baseless messaging that put lives at further risk.

Last but not least, adherence to PCR testing likely plays a large role. It is clear that widespread use of diagnostic tests contributed to the active detection of overt and asymptomatic forms of infection, with the latter often prevailing among infected individuals [7, 8]. During the COVID-19 pandemic, this has been one of the main means of actively detecting infection.

### B. Biological factors

Features of pathogen genetics play a role. SARS-CoV-2 belongs to a family of respiratory viruses characterized by high genetic variability. The progenitor Wuhan strain circulated around the world in the first half of 2020. Since December 2020, it was replaced by variant B.1.1.7 (alpha) [9], followed by more pathogenic beta (501.V2/B.1.351) and gamma (B1.128) variants, which had the phenomenon of immunological escape [10, 11]. In February 2020, there was a P.1 strain (S01Y.V3) which spread rapidly, despite the start of clinical trials of the first anti-coronavirus vaccines (based on the mRNA platform) at that time [12]. This was a clear signal that new SARS-CoV-2 vaccines will not necessarily be effective against all mutant variants of the virus. In December 2020, the B.1.617.2 variant (Delta) was detected in India for the first time, displacing other genetic variants of the virus [13, 14, 15]. Finally, the highly pathogenic B.1.1.529 variant (Omicron) was the last to be detected [14, 16]. It was detected in individuals vaccinated against SARS-CoV-2, and it turned out to be 2.8-fold more contagious than the Delta variant [14, 17, 18, 19].

Vaccination coverage, which affects the resistance of populations to infection, is another factor. In addition to circulating infection (with formation of post-infectious immunity), the level of herd immunity is influenced by the vaccination coverage of the population (with formation of post-vaccination immunity). Specific features of vaccines used in different countries can be a significant immunological factor that determines the prevalence in a particular country. All existing SARS-CoV-2 vaccines could be divided into four types according to the technological platform: mRNA vaccines; vector vaccines; whole-virion vaccines; and peptide vaccines. The first three types are the most common globally.

The first vaccine, mRNA-1273, was developed by Moderna in early spring 2020; its clinical trials began in China already on March 16, 2020 [20]. Almost simultaneously, another manufacturer (Pfizer/BioNTech) presented the BNT162b2 preparation based on a similar platform [21, 22]. The common basis of mRNA vaccines suggests similar humoral and cellular responses. This family of vaccine preparations has been shown to promote the production of a limited set of anti-spike antibodies and cellular responses, with active responses of CD4⁺ T cells in particular [23]. An mRNA vaccine forms a stable immune response by the 7th day after the first injection. It reaches a maximum on the 14th day after the second injection. The response most often persists for about three months [24]. At the same time, this immunity can be overcome by a genetically different version of the virus, such as strain B.1.1.529 (Omicron) [17]. This is exactly what happened in Israel. In the context of almost total double-immunization with BNT162b2, a significant increase in COVID-19 was noted during the outbreak caused by B.1.1.529; additional booster vaccinations were required [25].

Another group are vector vaccines, which are constructs consisting of a vector with an embedded fragment of the viral genome. Upon entering a cell, they generate an immunogen (SARS-CoV-2 spike protein) that induces an immune response, including antibody production and cellular immunity, polarized to the side Th1 IFNγ [26]. Vectors used have included: two human adenoviruses, Ad26 and Ad5, in the Gam-COVID-Vac vaccine; or chimpanzee adenovirus ChAdOx1 (ChAdOx1 nCoV-19 vaccine, AZD1222). In many countries, the non-replicating vector vaccines Covivideca (CanSino Biologics) or Janssen COVID-19 Vaccine (Johnson & Johnson) are actively used [27, 28, 29, 30]. Comparative studies indicate similar safety and efficacy of mRNA vaccines and vector vaccines [31]. However, the antibody spectrum and their duration of circulation is longer in vector vaccines than in mRNA vaccines [32].
Whole-virion vaccines containing inactivated virus represent a traditional vaccine development technology. It is no coincidence that vaccines of this type have been developed in most countries. They have a good safety profile and a strong antibody response, but their relatively low immunogenicity requires the use of an adjuvant and multiple vaccinations [33]. Among them, the Chinese vaccines Sinovac/CoronaVac, Sinopharm BBIBP-CorB, and Covaxim are the most common [34, 35, 36, 37]. A whole-virion inactivated vaccine, CovifVac, has been registered in Russia [38]. Similar vaccines have been developed and registered in other countries. Some examples include: QazCovid-in® in Kazakhstan [39]; BBV152/COVAXIN in India [40]; Soberana 02 and Abdala in Cuba [41]; and BIV1-CovIran in Iran [42]. It has been shown that all of these are high-quality prophylactic vaccines capable of generating full-fledged immune protection and forming the widest spectrum of immune response [33, 38, 43, 44].

The purpose of this study was to analyze the relationship between COVID-19 prevalence in the population and SARS-CoV-2 vaccine types used in different countries of the world, taking into account demographic and epidemiological factors.

2. Materials And Methods

2.1. Formation of the initial database

For this analysis, we have created a database in which was collected demographic and immunoepidemiological information about the COVID-19 situation in 104 countries as published in official sources and repositories. Information on prevalence, testing coverage, vaccination coverage, and vaccine ratios used (as of May 10–20, 2022) was taken from: Ritchie et al. (https://ourworldindata.org), accessed 13/05/2022 [2]; Official Coronavirus Statistics (https://gogov.ru/articles/covid-19), accessed 13–15/05/2022 [45]; Coronavirus Monitor (https://coronavirus-monitor.info), accessed 10–20/05/2022 [1]; the World Health Organization (https://covid19.who.int), accessed 10–20/05/2022 [46]; and the Pan American Health Organization (https://www.paho.org), accessed 15–20/05/2022 [47].

When necessary, we used official information from the government websites of a number of states. We also used data on population size and density in different countries provided by the United Nations Department of Economic and Social Affairs (UNDESA, http://population.un.org) [48]. The collected information was brought together in a single database, shown in Table 1S. As no contact with the national populations listed in Table 1S was expected, and analysis used previously-published information in the aforementioned official sources, ethics committee approval was not required.

2.2. Study design

The main criteria for inclusion in the initial database were availability of: information about the vaccines used and data on the relative usage of each among all vaccinated. The study's design is schematically presented in Fig. 1. The baseline included the following information for each country: population size and density; SARS-CoV-2 testing coverage (tests per 1,000 population); prevalence (cumulative number of registered cases per 1 million population); vaccination coverage (%); and a list of vaccines used, including the share of each among all vaccinated (%).

Since we initially collected data on all countries for which information was available (vaccines, relative share), further stratification of the initial data set (by population size, vaccination coverage) was necessary. The final data set was subjected to statistical processing, both in general and taking into account testing coverage of the population.

Statistical analysis

Statistical analyses were performed, depending on the nature of the data distribution, using parametric and non-parametric methods in the Excel (2010) and Statistica (ver.12, Windows) applications. The results were presented as median (Me) and interval [Q25 – Q75]. Rank correlation analysis was performed using the Spearman method. Multivariate, discriminant, and regression analyses were performed using Statistica (ver.12, Windows). Calculation of statistically significant differences was carried out according to the Mann-Whitney test. When necessary, Bonferroni correction for multiple comparisons was applied. Differences were designated as significant, unless otherwise indicated, at the $p < 0.05$ level.

3. Results

3.1. Characteristics of the original data set and correction

Figure 2 shows the information collected for 104 countries, including: population size and density; testing coverage (tests per 1,000 population); prevalence (cumulative number of registered cases per 1 million population); vaccination coverage (%); and enumeration of vaccines used, including the share of each among all vaccinated (%). The collected primary information is summarized in Table 1S.
National populations

This indicator varied widely (Fig. 3), with maximum values in China (pop. 1,439,324,000) and India (pop. 1,380,004,000), alongside minimum values in Montserrat (pop. 5,000) and the Falkland Islands (pop. 3,000). Globally, there are more than 200 countries, including groups of large and small nations. United Nations nomenclature [49] designates these as largest (> 100 million), large (40–100 million), medium (20–40 million), small (1.5–20 million), and smallest (0.5–1.5 million). Finally, dwarf microstates (< 0.5 million) are distinguished in a number of cases [50].

According to this classification, 104 countries were divided into the relevant groups (Table 1). According to correlation analysis, population size did not correlate with COVID-19 prevalence in any way (r = -0.144; p > 0.1). For this reason, base adjustment for this indicator was minimal, consisting only of removal of ‘dwarf states’. Hence, for initial adjustment, we excluded countries with a population of 500,000 or less from the study’s data sample. In result, 86 countries remained in the database.

Table 1
Population size groups and countries (UN classification).

<table>
<thead>
<tr>
<th>№</th>
<th>Country group</th>
<th>Population</th>
<th>Number of countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>very large</td>
<td>&gt; 100 million</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>large</td>
<td>40–100 million</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>medium</td>
<td>20–40 million</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>small</td>
<td>1.5–20 million</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>very small</td>
<td>0.5–1.5 million</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>dwarf</td>
<td>&lt; 500 thousand</td>
<td>18</td>
</tr>
</tbody>
</table>

Vaccination coverage

The next indicator, of fundamental importance for the validity of analysis, was vaccination coverage. The number vaccinated relative to overall population (%) was used as an indicator reflecting vaccination level. In the initial data set, coverage fluctuated over a wide range (Table 1S, Fig. 4): from 0.1% (Burundi) to 96.9% (Chile). Therefore, the sample was stratified according to this indicator. The lower threshold for stratification was 50% vaccination coverage. The legitimacy of this approach is primarily due to epidemiological factors. It is known that vaccination coverage thresholds largely determine COVID-19 epidemic spread dynamics [4]. Depending on prevailing conditions, this indicator can vary from 30–85% [51, 52].

Here, we used 50% as the minimum effective threshold for vaccination coverage. This level was also substantiated by a correlation analysis between prevalence and vaccination coverage in the original data set. Significant correlation was found (r = -0.5112) with strong significance (p < 0.00001). Subsequent regression analysis showed that this dependence is described by a logarithmic curve equation of the form y = 8.1469ln(x) - 31.999 (Fig. 5A). The resulting regression curve consists of two branches. The left vertical branch describes the initial phase of immune response formation; it is associated with prevalence level with a probability of p = 0.00001. The choice of this threshold is due to the transition point of the regression curve’s ascending branch to a flat section, established in the process of regression analysis.

In order to exclude the impact of this phase on prevalence, a correction was made to exclude all countries with less than 50% vaccination coverage from the data set. Thus, the second base correction was exclusion of such countries (coverage < 50%) from the set. In result, 33 countries were excluded, leaving 53 countries in the final data set (Table 2S). As such, the range of vaccination levels was significantly narrower, with a maximum in Chile (96.9%) and a minimum in Venezuela (50.2%). Following correction, the logarithmic regression transformed into a linear form, described by the equation y = -9E-06x + 76.044 (Fig. 5B). At the same time, the coefficient of determination value (R²) decreased from 0.4119 (Fig. 5A) to 0.0134 (Fig. 5B). This indicates an absence of any significant relationship between prevalence and vaccination coverage (r = -0.1157; p > 0.1).

Population density

In the initial data set, density also showed high heterogeneity (Table 1S, Fig. 6), with a maximum in Monaco (26,152.3 /km²) and a minimum in Greenland (0.1 /km²). In the final data set, the population density range was significantly narrower (Table 2S), with a maximum in Hong
Kong (7,082.1 /km²) and a minimum in Australia (3.3 /km²). Density also did not correlate with COVID-19 prevalence in any way (r = 0.09; p > 0.1). For this indicator, the base was not adjusted.

**COVID-19 prevalence**

The cumulative number of reported cases per 1 million population was used as an integral indicator of prevalence. In the initial data set, this indicator varied over a wide range (Table 1S, Fig. 7), from 664,914 (Faroe Islands) to 237 (Cook Islands). In the final data set, the prevalence range was significantly narrower (Table 2S), with a maximum in Cyprus (552,200) and a minimum in China (617).

**SARS-CoV-2 testing coverage**

Testing coverage was assessed as the cumulative number of tests per 1,000 population. In the initial data set, this information was not available for 25 countries. For the remaining 79 countries, this indicator varied widely (Fig. 8), from 32,860 (Cyprus) to 10 (Niger). In the final data set, information on this indicator was missing for only 3 countries, and the minimum value increased to 52 (Sierra Leone). Testing coverage was strongly correlated with COVID-19 prevalence. In the initial data set, the relationship was strong (r = 0.5949; p < 0.000001). The relationship was also confirmed in the final set (r = 0.5637; p < 0.001). For this indicator, the base was not adjusted. However, the presence of a stable, highly significant correlation with prevalence justified the need for further analysis taking it into account.

**Vaccines used in different countries**

During preliminary work, it was found that different vaccines (created on four main platforms) were used in different countries. In some, vaccines from the mRNA platform were fully or partially used. In other countries, various platforms (vector, protein/peptide, whole-virion) were used. In most countries, several vaccines created on different platforms were used. Unfortunately, there was no detailed quantitative information on the vaccines used for many countries, which significantly complicated analysis. All vaccines used were divided into four groups.

The first group included preparations based on the mRNA platform (mRNA-1273, BNT162b2). Among them, the Pfizer vaccine accounted for 80.3%, and the Moderna vaccine accounted for 19.7%. This group was abbreviated as V1.

In the 2nd group, all vector vaccines were combined: Johnson & Johnson; ChAdOx1 (Oxford-AstraZeneca, Covishield); Gam-COVID-Vac (Sputnik V); and Convidecia (CanSino Biologics). In the final data set, the Oxford/AstraZeneca vaccine accounted for the largest share (75.6%). Gam-COVID-Vac (Sputnik V) represented 12.3%, and the Johnson & Johnson vaccine represented 10.7%. This group was abbreviated as V2.

Peptide and protein vaccines were combined into group V3 (Novavax, EpiVacCorona, Abdala); their prevalence was less than 1%. Whole-virion vaccines were grouped as V4 (Sinovac, Sinopharm, Soberana, COViran Barekat, Covaxin, FAKHRAVAC, Corbevax, QazVac, IMBCAMS, KCONVAC, CoviVac). In terms of the number of vaccines developed, V4 is the largest group. In it, Sinopharm stood out (54.3%), alongside Sinovac (40.6%) and Covaxin (4%).

When information was available on each of the four vaccine types for a country, their share of total vaccines was calculated. In terms of the prevalence of vaccine types (platforms used), the leaders were V1 (59.08%), V2 (27.75%), and V4 (12.89%). Type V3 accounted for only 0.28%.

**3.2. Characteristics of the final data set**

The final data set included 53 countries, the distribution of which (by group) is presented in Table 2. General characteristics of the indicators used (population size and density, SARS-CoV-2 testing coverage, reported COVID-19 case numbers, vaccination coverage, enumeration of vaccines used) have been presented above (section 3.1). Correction made it possible to establish the final structure of correlations in the set (Table 2). Multiple correlation analysis revealed a number of strong statistical relationships between ‘reported COVID-19 case numbers’ and other sample parameters (Table 2).
Table 2
Distribution of countries in the final data set by size group.

<table>
<thead>
<tr>
<th>№</th>
<th>Country group</th>
<th>Group size</th>
<th>Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>very large (&gt;100 million people)</td>
<td>5</td>
<td>India, China, Russia, USA, Japan</td>
</tr>
<tr>
<td>2</td>
<td>large (40–100 million people)</td>
<td>9</td>
<td>Argentina, Great Britain, Germany, Iran, Spain, Italy, France, South Africa, South Korea</td>
</tr>
<tr>
<td>3</td>
<td>medium (20–40 million people)</td>
<td>6</td>
<td>Australia, Venezuela, Canada, Nepal, Peru, Poland</td>
</tr>
<tr>
<td>№</td>
<td>Country group</td>
<td>Group size</td>
<td>Countries</td>
</tr>
<tr>
<td>---</td>
<td>---------------------------</td>
<td>------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>small (1.5–20 million people)</td>
<td>27</td>
<td>Austria, Belarus, Belgium, Hungary, Hong Kong, Denmark, Israel, Ireland, Kazakhstan, Kyrgyzstan, Cuba, Lithuania, The Netherlands, Norway, Portugal, Senegal, Slovakia, Slovenia, Sierra Leone, Togo, Finland, Croatia, Czechia, Chile, Switzerland, Sweden, Ecuador</td>
</tr>
<tr>
<td>5</td>
<td>very small (0.5–1.5 million people)</td>
<td>6</td>
<td>Bhutan, Cyprus, Luxembourg, Mauritius, Malta, Estonia</td>
</tr>
</tbody>
</table>

In the pair ‘reported COVID-19 case numbers’ – ‘population size’, a weak inverse correlation was found ($r = -0.285; p < 0.05$). Correlations were not found in two other pairs: ‘reported COVID-19 case numbers’ – ‘population density’; and ‘reported COVID-19 case numbers’ – ‘vaccination coverage’. As mentioned earlier, a strong correlation was found in the pair ‘reported COVID-19 case numbers’ – ‘testing coverage’ ($r = 0.5637; p < 0.001$). Therefore, further analysis was carried out taking into account the number of tests per 1,000 population.

In addition to those already listed, correlations were established between ‘reported COVID-19 case numbers’ and the share of vaccine types (V1, V2, V3, V4) used. The strongest, significantly positive, relationship between was found for V1 ($r = 0.7251; p < 0.0001$). An inverse
correlation was found for other vaccine types (significant for V2, V4), and for V3 this relationship was in the nature of a trend. Given the unidirectional correlation coefficients between ‘reported COVID-19 case numbers’ with V2, V3 and V4, we considered it reasonable to combine these three vaccine types into one group, designated ‘non-mRNA group vaccines’ (Vnmg), for further analysis as shown in Table 2S.

### 3.3. Influence of share V1 and Vnmg on ‘reported COVID-19 case numbers’ in different countries globally

As already shown, group V1 and Vnmg vaccines featured opposite relationships with ‘reported COVID-19 case numbers’ (Table 3). In different countries, their ratios varied from 0 to 100% (Table 2S). In this regard, we hypothesized that a certain factor influencing ‘reported COVID-19 case numbers’ could be the ratio of vaccine types used in different countries. To test this assumption, a continuous axis of V1:Vnmg ratios was plotted from 0 to 100% (Fig. 9). Analysis of these ratios (%:%) permitted distribution of the sample to: countries in which V1 vaccination prevailed (80:20, 90:10, 100:0); countries where V1 and Vnmg usage was nearly equal (60:40, 50:50, 40:60); and countries in which V2 vaccination predominated (0:100, 30:70). The lowest ‘reported COVID-19 case numbers’ were found at V1:Vnmg ratios from 0:100 to 30:70; the highest case numbers were found at ratios from 80:20 to 100:0.

**Table 3.** Correlation analysis of the adjusted data set (n=53).

<table>
<thead>
<tr>
<th>Dependent variable, significance</th>
<th>Independent Sample Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>reported COVID-19 case numbers</td>
<td>-0.285</td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

Note. Critical value of the correlation coefficient: 0.233 at $p=0.1$; 0.276 at $p=0.05$; 0.358 at $p=0.01$; and 0.447 at $p=0.001$

Based on the results, three subgroups of countries were formed. The 1st subgroup (V1 > Vnmg) included countries with V1:Vnmg ratios from 70:30 to 100:0 (n = 30). The 2nd subgroup (V1 ≈ Vnmg) with the V1:Vnmg ratios from 40:60 to 60:40 included 6 countries. The 3rd subgroup (V1 < Vnmg) with V1:Vnmg ratios from 0:100 to 30:70 included 17 countries (Fig. 10).

The highest ‘reported COVID-19 case numbers’ (hereafter per million, unless noted) were noted in the 1st subgroup where V1 > Vnmg (349,700 [246,024–436,413]). The lowest level was in the 3rd subgroup with V1 < Vnmg (49,900 [10,672–113,641]). Moreover, this indicator in the V1 < Vnmg subgroup was significantly lower than in the V1 > Vnmg subgroup ($p = 0.000001$) and the V1 ≈ Vnmg subgroup ($p = 0.0099$). The V1 ≈ Vnmg subgroup had an intermediate level of ‘reported COVID-19 case numbers’ (178,420 [103,420–282,861]). Statistical analysis of the significance of differences (between the compared groups according to three statistical criteria) is given in Table 3S.

At the same time, there were no statistically significant differences in ‘reported COVID-19 case numbers’ between subgroups V1 > Vnmg and V1 ≈ Vnmg ($p = 0.0357$). Taking into account the multiplicity of intergroup comparisons, Bonferroni’s correction for multiple comparisons was applied to determine the critical $p$ value. Given that each group was compared twice, the critical $p$ value after applying the correction was $p \leq 0.025$. The $p$ values in both comparison cases were less than the critical value obtained using the Bonferroni correction. Thus, reported COVID-19 case numbers in the V1 < Vnmg subgroup (adjusted for multiple comparisons) were significantly lower than in the other two groups. The data obtained confirmed the initial assertion that V1 and Vnmg were in opposite relationships with ‘reported COVID-19 case numbers’ in the studied populations.

### 3.4. Dependence of ‘reported COVID-19 case numbers’ on vaccine type, taking into account SARS-CoV-2 testing adherence

Reported COVID-19 case numbers in different countries can largely depend on testing coverage. This is particularly evidenced by a strong correlation between these indicators ($r = 0.5637; p < 0.001$) with a high level of determination, $R^2 = 0.3177$ (Table 3). To determine the legitimacy of such a dependence, we analyzed the exponential curve shown in Fig. 11. Based on characteristics of the trend line, which reflects the dependence of prevalence on testing coverage, the study sample was empirically divided into three groups. Grouping validity was checked using discriminant analysis, which included two independent variables: ‘testing coverage’ and ‘reported COVID-19 case numbers’. As
a result of a stepwise change in the structure of the groups, it was possible to formulate a discriminant function with a high degree of reliability for both independent variables: for the variable 'testing coverage' \(p<0.000001\); and for the variable 'prevalence' \(p=0.000014\).

When assessing classifying ability, this discriminant model made the least number of classification errors: 7 errors out of 50 classification acts. Discriminant analysis data are presented in Table 4S. As a result of discriminant analysis, three groups were identified with the following testing coverage ranges (tests per 1,000 people): group 1 with \(\leq 972\) \((n=13)\); group 2 with from 1007 to 7348 \((n=33)\); and group 3 with \(\geq 9380\) \((n=4)\). Although the number of tests is significant, it is not, however, the only indicator that influences 'reported COVID-19 case numbers'. Another factor may be the type of vaccines used. In each group, the effect of \(V1:Vnmg\) vaccine ratio on 'reported COVID-19 case numbers' was assessed according to the subgrouping described earlier (section 3.1).

In the first group, which included countries with a low level of testing, division into \(V1 > Vnmg\) and \(V1 < Vnmg\) did not reveal statistically significant differences, which seems quite logical for countries where there is less than 1 test per resident. The third group (countries with an extremely high level of testing) was excluded from further analysis due to its small size.

The second group included countries with testing levels ranging from 1,007 to 7,348 tests per 1,000 people (Fig. 12A). This group represented the main sample, within which the correlation coefficient between 'reported COVID-19 case numbers' and the level of testing was \(r=0.156\) \((p>0.1)\); this proves that there is no relationship between these two indicators.

In this main sample group, the 'reported COVID-19 case numbers' indicator (Fig. 12B) was 342,300 \((261,365–408,949)\) for countries in subgroup 1 and 122,871 \((104,411–177,508)\) for countries in subgroup 3, with a high level of significance \((p=0.000244)\) for difference between subgroups 1 and 3 (Table 5S).

For subgroup 2 countries with \(V1 = Vnmg\), this indicator occupied a middle position, amounting to 195,400 \((161,440–266,595)\). Differences between subgroups 1 and 2 were significant \((p=0.027391)\); they were absent between subgroups 2 and 3 \((p=0.21)\). Thus, when the important variable 'testing coverage' is excluded from the data set, higher 'reported COVID-19 case numbers' are noted among countries with predominant use of \(V1\) vaccines. Among the populations of countries where \(Vnmg\) vaccines were more widely used, 'reported COVID-19 case numbers' were lower by an average of 2.8-fold.

Since Russia was in the 2nd group, where 'reported COVID-19 case numbers' were several times lower than in neighboring European countries, we separately analyzed countries comparable to Russia in terms of testing coverage. To this end, a sample of countries was formed such that with Russia's testing coverage occupied the median value \((2001\) tests per 1,000 people); the interquartile interval was 1697–2602. This interval included 13 countries (Fig. 13A). Similar to analysis of the entire sample, significant differences \((p=0.041259)\) were found between subgroup 1 \((311,400 [247,232–432,100])\) and subgroup 3 \((155,386 [122,871 – 187,900])\) (Fig. 13B). Moreover, it should be noted that in Russia, 'reported COVID-19 case numbers' were 1.6-fold lower than in Finland, which features the lowest level of this indicator among subgroup 1 countries.

During formation of the data set, two aspects were not analyzed: duration of the post-vaccination period; and specifics of viral genetic variants circulating in the population (with varying ability to evade immunity). We can only note the observed decrease in 'reported COVID-19 case numbers' in countries where vector and whole-virion vaccines were used. Possible reasons for this phenomenon should probably be the subject of further research.

**Discussion**

The COVID-19 pandemic has become one of the biggest challenges of the 21st century. Suddenly appearing on December 31, 2019 in Wuhan (PRC), it covered most of the world within a matter of weeks. The lack of any effective antiviral drugs and specific vaccines prompted most countries globally to respond in the form of unprecedented restrictive measures. On the other hand, the lack of effective therapeutic and prophylactic agents contributed to the unprecedented activity of the scientific community. This resulted in the rapid development and implementation of a whole series of vaccines on various platforms, with no parallels in the history of medicine [33, 53, 54]. Four of them are the most widespread: messenger RNA-based vaccines (designated as group \(V1\) in this work) [55, 56]; vector vaccines (\(V2\)) [57, 58]; peptide and protein vaccines (\(V3\)); and inactivated, whole-virion vaccines (\(V4\)) [42, 43, 44].

It should be noted that most vaccines were created on the basis of the progenitor Wuhan strain [59]. Meanwhile already in 2020, new genetic variants of the virus began to appear, characterized by more pathogenic properties. In result, data appeared on the ability of new viral variants to overcome the adaptive immunity created by vaccines [60, 61]. In addition, it has been shown that various vaccines form post-vaccination immunity differently; they can differ both in duration and in conferred resistance to new viral genetic variants [62]. All of the above may have contributed to the varying COVID-19 prevalence around the world.
In this regard, it seemed relevant to assess the relationship between COVID-19 prevalence in the population and SARS-CoV-2 vaccine types used in different countries globally, taking into account demographic (population size, density) and immunobiological (testing and vaccination coverage) factors. The work was sequentially performed, according to algorithm, in several stages (Fig. 1). The main results are presented in Fig. 2. To this end, the necessary information was collected from available sources for 104 countries. The main inclusion criterion was availability of information on the vaccines used, including their proportional contributions to overall vaccination.

The collected information was subsequently summarized in a general table (Table 1S), and initial analysis showed high heterogeneity in all indicators. Heterogeneity in terms of demographic indicators was seen by the presence of all country groups in the initial data set (according to UN classification); only dwarf states were subsequently excluded. For the goals set in the research, vaccination coverage was a critical indicator, initially ranging from 0.1% in Burundi to 96.6% in Chile and South Africa (Table 1S). For valid analysis, countries with vaccination coverage below 50% were further excluded. The People's Republic of China (PRC) was included in the data set despite local peculiarities of controlling the COVID-19 epidemic, specifically the PRC's overall 'zero COVID-19 strategy' in the event of outbreaks. This strategy is essentially unique and is not applied in other countries of the world.

Following preparation of a final data set of 53 countries, reported COVID-19 case numbers were found to be most strongly correlated with testing coverage and proportion of vaccine types used. Due to the fact that an inverse correlation was found between 'reported COVID-19 case numbers' with V2, V3 and V4, we considered it possible to combine these three types of vaccines into one group (non-mRNA group vaccines, Vnmg) for further analysis. Moreover, to offset the impact of SARS-CoV-2 testing on reported COVID-19 case numbers, countries with low testing rates (< 1 test/person) and countries with extremely high testing rates were excluded. The resulting analyzed sample represented 33 countries. When analyzing reported COVID-19 case numbers within this sample, it was shown that the highest numbers occur in countries with V1 vaccine (mRNA) dominance. The lowest was seen in countries with Vnmg (vector, peptide/protein, whole-virion/inactivated) vaccine dominance (Fig. 12).

For the most appropriate comparison of Russia with other countries, a data set was formed with a specific median level of 'testing coverage' (2001 per 1,000 people). Analysis showed that, like analysis of the overall sample, there were significant differences between countries with V1 vaccine (mRNA) dominance and countries with Vnmg dominance.

Moreover, it should be noted that, despite exactly the same testing levels, 'reported COVID-19 case numbers' for Russia (with Vnmg usage exclusively) were 1.6-fold lower than, for example, in Finland. Russia's low numbers are remarkable given the fact that Finland itself already represents with the lowest case numbers among 10 countries with the dominant use of mRNA vaccines (Fig. 13). The results show that in countries where non-mRNA vaccines have been used, 'reported COVID-19 case numbers' are significantly lower than in countries with dominant use of mRNA vaccines.

A chart summarizing information for all 53 countries regarding reported COVID-19 case numbers, and their dependence on vaccine types used, is presented in Fig. 14. In addition to polar groups (nations dominated by V1 or Vnmg), the group of countries with equal proportions of mRNA and non-mRNA vaccines is of particular interest. As expected, reported COVID-19 case numbers for these countries are middle values.

The South Asian country of Bhutan stands apart because unique experience was gained through combined use of different vaccine types. The first immunization was carried out with an mRNA vaccine; the second immunization used a vector vaccine [63]. At the same time, high vaccination coverage and the lowest 'reported COVID-19 case numbers' were achieved among countries in this group. However, case numbers in Bhutan were higher than in the neighboring regional countries of India and Nepal. To be fair, in the latter two countries, testing coverage was less than 1 test per person.

A similar situation has developed in East Asian countries. The maximum prevalence was registered in South Korea, which used mainly mRNA vaccines. The minimum was registered in China with the dominant use of whole-virion vaccines. Hong Kong, with an equal proportion of mRNA and non-mRNA vaccines, featured an intermediate prevalence. This is despite the fact that: the minimum testing coverage was in South Korea; and in China and Hong Kong the level of testing was equally high. Interestingly, in Japan, the prevalence was the lowest among countries in the region; this was combined with a very low level of testing for highly developed countries. On the other hand, in Israel, where testing coverage was even slightly lower than in Hong Kong and China, 'reported COVID-19 case numbers' were 2.8-fold higher than in Hong Kong and three orders of magnitude higher than in China.

In a number of European countries in which mRNA vaccines dominated (France, Portugal, Luxembourg, Estonia, Slovenia, Switzerland), prevalence was 18–45% higher than in the UK. The UK had an even vaccine-type ratio (mRNA vs. non-mRNA), and testing coverage was equal to, or higher (1.5 to 3-fold) than, the listed countries.

In another group of four European countries with the same testing level, maximum prevalence was registered in Germany and Croatia, which mainly used mRNA vaccines. The minimum was in Belarus with dominant use of vector and whole-virion vaccines (2.6 to 3-fold less than...
Germany and Croatia). An intermediate level was seen in Hungary with an equal ratio of mRNA and non-mRNA vaccines (1.4 to 1.6-fold less than in Germany and Croatia).

The presented statistical calculations inevitably raise the question of what is the underlying cause of the difference noted. Comparative studies have convincingly shown the high efficacy of all currently approved vaccines [61, 64]. However, it is impossible not to notice some differences that can affect prevalence. Regarding mRNA vaccines, one can agree with the opinion of a number of researchers who have shown that they are able to form rapid immunity already in the early stages after immunization, persisting for 3 months, following which, use of a booster dose may be required [21]. As for vector vaccines (assigned to the Vmng group), full-fledged immunity is formed by the 14th day after the second dose, but it exists for at least 6 months [28, 30, 65, 66]. Regarding whole-virion vaccines, it is worth noting their lowest immunogenicity, alongside their longest elicited immunity [67, 68], which is closest to a post-infectious response.

The key parameter is probably the formation of stable herd immunity. Unfortunately, there are currently no published studies that have been conducted according to a single methodology with different countries globally when examining epidemic process dynamics as was done here. In our experience, assessment of SARS-CoV-2 collective immunity, carried out according to a single methodology [69] at different stages of the epidemic in Russia [8, 70], Belarus [71], and Kyrgyzstan [72], showed that there is successful formation of herd immunity in those countries. Those countries were discussed in our earlier work, and usage of vector and whole-virion vaccines usage dominates in them.

On the other hand, in the work of Morens et al. [73], it was suggested that it is impossible to achieve long-term herd immunity with COVID-19 and therefore regular booster immunization is necessary. This is probably true primarily for mRNA vaccines, whose distribution currently dominates globally. In our opinion, the main reasons for this are: breadth (spectrum) of the immune response; and constant viral variability, through which new genetic variants appear regularly [74].

Vector vaccines, and even more so whole-virion vaccines, induce a significantly wider range of post-vaccination antibodies than mRNA vaccines [33, 75]. Minimal diversity and maximal specificity were features embraced in the initially ideology of mRNA vaccine development. Undoubtedly, it is a very progressive technology that makes it possible to induce a narrow spectrum of post-vaccination antibodies, resulting in a decrease in the share of post-vaccination adverse reactions, including those of an autoimmune nature [76]. In conditions of high viral variability, however, the technology likely has a number of limitations, and the formation of a narrow Ab spectrum is more of a disadvantage than an advantage. In result, the level of post-vaccination immunity persists for a short time, and its restoration is impossible without the introduction of a second vaccine dose.

In contrast, an immune response is formed to a much wider range of antigens and their epitopes with the use of vector and whole-virion vaccines. In result, when post-vaccination immunity is attenuated, a repeat encounter with the virus, even a new genetic variant, leads to activation of the secondary immune response. As such, regular booster immunization is not required. A similar situation is observed with healthcare workers. Often, they have been ill only once, but regular restoration of post-infectious immunity ensues through periodic contact with infected carriers.

When analyzing the assembled country database, we did not have data on the timing of primary or booster vaccinations of course. Therefore, we could not in any way assess post-vaccination immunity level at the time of information collection. Another factor is the uneven spread across countries of new genetic variants of the virus, which may feature different abilities to evade the immune response. Answering these and other questions would likely permit more accurate determination of the nature of post-vaccination morbidity. Perhaps this will be the subject of a separate study someday.

Thus, we have shown for the first time that ‘reported COVID-19 case numbers’ (per million population) depend not only on SARS-CoV-2 testing coverage and vaccination coverage, which is quite logical, but also on the vaccine types used. With the same level of vaccination and testing coverage, countries using predominantly vector and whole-virion vaccines experienced significantly lower prevalence than countries predominantly using mRNA vaccines.

**Declarations**

**Contributors**

AAT- conceptualisation, data analysis, manuscript preparation, supervision; VSS- data analysis, manuscript preparation; AAK - statistical analysis, manuscript preparation; ERS – reviewing and editing; VGD- reviewing and editing, decision to submit for publication, AYP- supervision

**Declaration of Competing Interest**
The authors declare that they have no competing interests.

Acknowledgments
None.

Funding
The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Ethical considerations
Not required.

References
1. Coronavirus-monitor. URL: https://coronavirus-monitor.info/ Accessed 13/05/2022


**Figures**

![Figure 1](image)

Research design workflow
Figure 2

Study profile

Figure 3
Distribution of countries (n=104) included in the original data set (Table 1S) by population. Countries with populations below 500,000 were excluded from the adjusted data set (Table 2S) and are marked in red. Source: United Nations Department of Economic and Social Affairs (UNDESA) (http://population.un.org).

Figure 4

Distribution of countries (n=104) included in the original data set (Table 1S) by vaccination coverage (%). Countries marked in red (<50% vaccination coverage) are not included in the adjusted set (Table 2S). Source: Our World in Data (https://ourworldindata.org/explorers/coronavirus-data-explorer?facet=none&uniformYAxis=0&Metric=People+vaccinated&Interval=Cumulative&Relative+to+Population=true&Color+by+test+positivity=false).
Figure 5

Relationship between reported COVID-19 case numbers and vaccination coverage. Key: A – initial data set (n=104); B – final data set (n=53). The regression equations are highlighted in red in the lower right.
Figure 6

Distribution of countries (n=104) included in the original data set (Table 1S) by population density (per km²). Source: United Nations Department of Economic and Social Affairs (UNDESA) (http://population.un.org).

Figure 7

Prevalence (per 1 million population)
Distribution of countries (n=104) included in the original data set (Table 1S) by COVID-19 prevalence (cumulative confirmed cases per 1 million population). Source: Our World in Data (https://ourworldindata.org/explorers/coronavirus-data-explorer?facet=none&uniformYAxis=0&Metric=Confirmed+cases&Interval=Cumulative&Relative+to+Population=true&Color+by+test+positivity=false).

Figure 8

Distribution of countries (n=79) included in the original data set (Table 1S) by testing coverage (cumulative number of tests per 1,000 people). Note: No data available for 25 out of 104 countries. Source: Our World in Data (https://ourworldindata.org/explorers/coronavirus-data-explorer?facet=none&uniformYAxis=0&Metric=Tests&Interval=Cumulative&Relative+to+Population=true&Color+by+test+positivity=false).
Figure 9

Reported COVID-19 case numbers in groups with different vaccine usage ratios (V1:Vnmg). Note: x-axis – V1: Vnmg ratio (%:%); y-axis – cases per 1 million population; numbers in boxes – median values.
Figure 10

Reported COVID-19 case numbers in countries with different V1:Vnmg vaccine usage ratios (n=53). Key: A shows the correlation between prevalence and share Vnmg ($r = 0.7251; p<0.0001$). B shows the median (Me) and interquartile interval (Q25 - Q75) in subgroup 1 (V1>Vnmg, n=30), subgroup 2 (V1=Vnmg, n=6), and subgroup 3 (V1<Vnmg, n=17). Significance ($p$) of differences: $p(1-2) = 0.036; p(2-3) = 0.01$; and $p(1-3) = 0.000001$. 
Figure 11

Relationship between SARS-CoV-2 testing coverage (tests per 1,000 people) and 'reported COVID-19 case numbers' per 1 million population. The correlation coefficient and regression equation are shown in the upper left.
Figure 12

Reported COVID-19 case numbers in countries with different V1:Vnmg vaccine usage ratios and testing coverage from 1007 to 7348 (per 1,000 people) (n=33). Key: A is distribution of countries by testing coverage. B is median (Me) and interquartile interval (Q25 - Q75) in groups depending on vaccine type dominance: subgroup 1 (V1>Vnmg, red); subgroup 2 (V1=Vnmg, green); and subgroup 3 (V1<Vnmg, blue).
Figure 13

Reported COVID-19 case numbers in countries (n=13) with different V1:Vnmg vaccine usage ratios and testing coverage near the 2001 level (1697 to 2602 tests per 1,000 population). Key: A shows countries comparable with Russia in terms of testing coverage. B is median (Me) and interquartile interval (Q25 – Q75) in: subgroup 1 (V1>Vnmg, n=11, red); subgroup 2 (V1=Vnmg, n=1, green); and subgroup 3 (V1<Vnmg, n=2, blue).
Figure 14

Distribution of countries by 'number of reported COVID-19 cases' and vaccine types used. Note: red diamonds – countries with mRNA vaccine dominance; green squares – countries with near equal use of mRNA and non-mRNA vaccines; blue triangles – countries dominated by non-mRNA vaccines (vector, peptide/protein, whole-virion/inactivated).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplimentary.docx