Background. To simulate how the number of COVID-19 cases increases versus time, various data sets and different mathematical models can be used. Since there are some differences in statistical data, the results of simulations can be different. Complex mathematical models contain many unknown parameters, the values of which must be determined using a limited number of observations of the disease over time. Even long-term monitoring of the epidemic may not provide reliable estimates of the model parameters due to the constant change of testing conditions, isolation of infected, quarantine conditions, pathogen mutations, vaccinations, etc. Therefore, simpler approaches are necessary. In particular, previous simulations of the COVID-19 epidemic dynamics in Ukraine were based on smoothing of the dependence of the number of cases on time and the generalized SIR (susceptible–infected–removed) model. These approaches allowed detecting the pandemic waves and calculating adequate predictions of their duration and final sizes. In particular, eight waves of the COVID-19 pandemic in Ukraine were investigated.

Objective. We aimed to detect the changes in the pandemic dynamics and present the results of SIR simulations based on Ukrainian national statistics and data reported by Johns Hopkins University (JHU) for Ukraine and Qatar.

Methods. In this study we use the smoothing method for the dependences of the number of cases on time, the generalized SIR model for the dynamics of any epidemic wave, the exact solution of the linear differential equations, and statistical approach for the model parameter identification developed before.

Results. The optimal values of the SIR model parameters were calculated and some predictions about final sizes and durations of the epidemics are presented. Corresponding SIR curves are shown and compared with the real numbers of cases.

Conclusions. Unfortunately, the forecasts are not very optimistic: in Ukraine, new cases will not stop appearing until June–July 2021; in Qatar, new cases are likely to appear throughout 2021. The expected long duration of the pandemic forces us to be careful and in solidarity. Probably the presented results could be useful in order to estimate the efficiency of vaccinations.

Keywords: COVID-19 pandemic; epidemic dynamics in Ukraine; epidemic dynamics in Qatar; mathematical modeling of infection diseases; SIR model; parameter identification; statistical methods.

Introduction

The studies of the COVID-19 pandemic dynamics in Ukraine are presented in [1–3] and summarized in the book [4]. Some SEIR (susceptible–exposed–infected–removed) and SEIRD (susceptible–exposed–infected–removed–dead) simulations of the pandemic dynamics in Qatar can be found in [5, 6]. For Ukraine, different simulation and comparison methods were based on official accumulated number of laboratory confirmed cases [7, 8] (national statistics). These figures coincides with the official WHO data sets [9]. Unfortunately, WHO stopped to provide the daily information in August 2020. In this study we will use also the information from COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [10].

The classical SIR model [11, 12], connecting the number of susceptible $S$, infected and spreading the infection $I$ and removed $R$ persons, was applied in [1, 4] to simulate the first pandemic wave in Ukraine. The unknown parameters of this model were estimated with the use of the cumulative number of cases $V = I + R$ and the statistics-based method of parameter identification developed in [1, 4].

The weakening of quarantine restrictions, changes in the social behavior and the coronavirus activity causes change in SIR characteristics and the epidemic dynamics. To detect these changes, a simple method of numerical differentiations of
accumulated number of cases was proposed in [2, 4]. To simulate these new pandemic waves, the SIR model was generalized in [2–4]. In [2, 4] the results of simulation of the first six epidemic waves in Ukraine are presented with the use of a procedure for sequentially determining the parameters of the model for each epidemic wave, starting with the first one.

This method requires considerable efforts and time. The book [4] introduced a new algorithm for determining the optimal parameter values for a particular epidemic wave without calculating the dynamics of previous waves and presented calculations for the seventh epidemic wave in Ukraine. The eighth epidemic wave in Ukraine were simulated in [3] with the use of this approach. In this paper, we will analyze the dynamics of the epidemic in Ukraine and Qatar in the period from December 1, 2020 to February 20, 2021 and make some predictions.

Materials and Methods

Data
We will use two data sets regarding the accumulated numbers of confirmed COVID-19 cases in Ukraine: the official information from national sources [7, 8] and data set $V_j$ from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [10]. These values and corresponding moments of time $t_j$ (measured in days, zero point is November 30, 2020) are shown in Table 1.

Detection of epidemic waves
To control the changes of epidemic parameters, we can use daily numbers of new cases and their derivatives. Since these values are random, we need some smoothing. For example, we can use the smoothed daily number of accumulated cases proposed in [2–4]:

$$
\bar{V}_i = \frac{1}{7} \sum_{j=i-3}^{j=i+3} V_j. \tag{1}
$$

The first and second derivatives can be estimated with the use of following formulas:

$$
d\bar{V}_i \left|_{t=t_i} \approx \frac{1}{2} (\bar{V}_{i+1} - \bar{V}_{i-1}), \tag{2}
$$

$$
d^2 \bar{V}_i \left|_{t=t_i} \approx \bar{V}_{i+1} - 2\bar{V}_i + \bar{V}_{i-1}. \tag{3}
$$

Generalized SIR model
The classical SIR model for an infectious disease [11, 12] was generalized in [2–4] to simulate different epidemic waves. We suppose that the SIR model parameters are constant for every epidemic wave, i.e. for the time periods:

$$
t_i \leq t \leq t_{i+1}, i = 1, 2, 3,...
$$

Then for every wave we can use the equations, similar to [11, 12]:

$$
\frac{dS}{dt} = -\alpha_i SI, \tag{4}
$$

$$
\frac{dI}{dt} = \alpha_i SI - \rho_i I, \tag{5}
$$

$$
\frac{dR}{dt} = \rho_i I, \tag{6}
$$

where $S$ is the number of susceptible persons (who are sensitive to the pathogen and not protected); $I$ is the number of infected persons (who are sick and spread the infection; please don’t confuse with the number of still ill persons, so known active cases) and $R$ is the number of removed persons (who no longer spread the infection; this number is the sum of isolated, recovered, dead, and infected people who left the region). Parameters $\alpha_i$ and $\rho_i$ are supposed to be constant for every epidemic wave.

Parameters $\alpha_i$ show how quick the susceptible persons become infected (see (4)). Large values of this parameter correspond to severe epidemics with many victims. These parameters accumulate many characteristics. First they shows how strong (viral) is the pathogen and what is the way of its spreading. Parameters $\alpha_i$ accumulate also the frequency of contacts and the way of contacting. In order to decrease the values of $\alpha_i$, we have to minimize the number of our contacts and change our contacting habits. For example, we have to avoid the public places and use masks there, minimize or cancel traveling. We have to change our contact habits: to avoid handshakes and kisses. First, all these simple things are very useful to protect yourself. In addition, if most people follow these recommendations, we have chance to diminish the values of $\alpha_i$ and reduce the negative effects of the pandemic.

The parameters $\rho_i$ characterize the patient removal rates, since eq. (6) demonstrates the increase rate of $R$. The inverse values $1/\rho_i$ are the
Table 1: Cumulative numbers of confirmed Covid-19 cases in Ukraine and Qatar

| Day in December 2020, \( t \) | NC in Ukraine, National statistics \([7, 8] V_j\) | NC in Ukraine, JHU [10] \( V_j \) | NC in Qatar, JHU [10] \( V_j \) | Day in January 2021 | NC in Ukraine, National statistics \([7, 8] V_j\) | NC in Ukraine, JHU [10] \( V_j \) | NC in Qatar, JHU [10] \( V_j \) | Day in February 2021 | NC in Ukraine, National statistics \([7, 8] V_j\) | NC in Qatar, JHU [10] \( V_j \) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 758264 | 765117 | 139901 | 1 | 1069517 | 1096855 | 144042 | 1 | 1223879 | 151720 |
| 2 | 772760 | 778560 | 139256 | 2 | 1074093 | 1102526 | 144240 | 2 | 1227164 | 152095 |
| 3 | 787891 | 793372 | 139477 | 3 | 1078251 | 1107137 | 144437 | 3 | 1232246 | 152491 |
| 4 | 801716 | 808828 | 139643 | 4 | 1083585 | 1117256 | 144644 | 4 | 1237169 | 152898 |
| 5 | 813306 | 822985 | 139783 | 5 | 1090496 | 1117256 | 144852 | 5 | 1241479 | 153296 |
| 6 | 821947 | 834913 | 139908 | 6 | 1099493 | 1124482 | 145061 | 6 | 1244849 | 153690 |
| 7 | 832758 | 843898 | 140086 | 7 | 1105169 | 1133802 | 145271 | 7 | 1246990 | 154098 |
| 8 | 845343 | 853504 | 140203 | 8 | 1110015 | 1139800 | 145466 | 8 | 1249646 | 154525 |
| 9 | 858774 | 867991 | 140353 | 9 | 1115026 | 1144943 | 145672 | 9 | 1253055 | 155002 |
| 10 | 872228 | 881727 | 140516 | 10 | 1119314 | 1150265 | 145865 | 10 | 1258094 | 155453 |
| 11 | 885039 | 895620 | 140680 | 11 | 1124430 | 1154850 | 146068 | 11 | 1262867 | 155901 |
| 12 | 894215 | 908839 | 140827 | 12 | 1130839 | 1160243 | 146279 | 12 | 1268049 | 156351 |
| 13 | 900666 | 918444 | 140961 | 13 | 1138764 | 1166958 | 146480 | 13 | 1271143 | 156804 |
| 14 | 909082 | 925321 | 141121 | 14 | 1146963 | 1173543 | 146689 | 14 | 1273475 | 157244 |
| 15 | 919704 | 934161 | 141272 | 15 | 1154692 | 1189363 | 146885 | 15 | 1276618 | 158132 |
| 16 | 931751 | 945218 | 141417 | 16 | 1160682 | 1192114 | 147089 | 16 | 1280904 | 158138 |
| 17 | 944381 | 957692 | 141557 | 17 | 1163716 | 1195152 | 147277 | 17 | 1287141 | 158591 |
| 18 | 956123 | 970758 | 141716 | 18 | 1167655 | 1201894 | 147504 | 18 | 1293672 | 159053 |
| 19 | 964448 | 982937 | 141858 | 19 | 1172038 | 1206125 | 147729 | 19 | 1299967 | 159518 |
| 20 | 970993 | 991700 | 142001 | 20 | 1177621 | 1210854 | 148000 | 20 | 1304456 | 159967 |
| 21 | 979506 | 998678 | 142159 | 21 | 1182699 | 1216780 | 148258 | 21 | – | – |
| 22 | 989642 | 1007627 | 142308 | 22 | 1187897 | 1222459 | 148521 | 22 | – | – |
| 23 | 1001132 | 1018199 | 142448 | 23 | 1191812 | 1227723 | 148772 | 23 | – | – |
| 24 | 1012167 | 1030125 | 142605 | 24 | 1194328 | 1231965 | 149019 | 24 | – | – |
| 25 | 1019876 | 1041583 | 142734 | 25 | 1197107 | 1234772 | 149296 | 25 | – | – |
| 26 | 1025989 | 1049717 | 142903 | 26 | 1200883 | 1237810 | 149595 | 26 | – | – |
| 27 | 1030374 | 1056265 | 143062 | 27 | 1206412 | 1241863 | 149933 | 27 | – | – |
| 28 | 1037362 | 1061074 | 143222 | 28 | 1211593 | 1247674 | 150280 | 28 | – | – |
| 29 | 1045348 | 1068476 | 143428 | 29 | 1216278 | 1253127 | 150621 | 29 | – | – |
| 30 | 1055047 | 1076880 | 143621 | 30 | 1219455 | 1258093 | 150984 | 30 | – | – |
| 31 | 1064479 | 1086997 | 143834 | 31 | 1221485 | 1261546 | 151335 | 31 | – | – |

Notes: NC denotes the number of cases. \( V_j \) – National statistics \([7, 8]\), \( V_j \) and \( V_j \) – according to COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [10].
estimations for time of spreading infection $\tau_i$ during $i$-th epidemic wave. So, we are interested in increasing the values of parameters $\rho_i$ and decreasing $1/\rho_i$. People and public authorities should work on this and organize immediate isolation of suspicious cases.

Since the derivative $d(S + I + R)/dt$ is equal to zero (it follows from summarizing Eqs. (4)–(6)), the sum

$$N_i = S + I + R$$

must be constant for every wave and is not the volume of population.

To determine the initial conditions for the set of equations (4)–(6), let us suppose that at the beginning of every epidemic wave $t_i^*$:

$$I(t_i^*) = I_i, \quad R(t_i^*) = R_i, \quad S(t_i^*) = N_i - I_i - R_i.$$  

It follows from (4) and (5) that

$$\frac{dI}{dS} = \frac{\nu_i}{S} - 1, \quad \nu_i = \frac{\rho_i}{\alpha_i}.$$  

Integration of (9) with the initial conditions (8) and taking into account (7) yields:

$$I = \nu_i \ln S - S + N_i - R_i - \nu_i \ln(N_i - I_i - R_i).$$  

It follows from (9) that function $I$ has a maximum at $S = \nu_i$ and tends to zero at infinity. The corresponding number of susceptible persons at infinity $S_{\infty} > 0$ can be calculated from the nonlinear equation:

$$S_{\infty} = (N_i - I_i - R_i) e^{-\frac{S_{\infty} - N_i - R_i}{\nu_i}}.$$  

Formula (11) follows from (10) at $I = 0$. In [2–4] the set of differential equations (4)–(6) was solved by introducing the function

$$V(t) = I(t) + R(t),$$  

corresponding to the number of victims or the cumulative confirmed number of cases. For many epidemics (including the COVID-19 pandemic) we cannot observe dependencies $S(t), I(t)$ and $R(t)$ but observations of the accumulated number of cases $V_i$ corresponding to the moments of time $t_i$ provide information for direct assessments of the dependence $V(t)$.

It follows from (5) and (6) that:

$$\frac{dV}{dt} = \alpha_i SI.$$  

Eqs. (7), (10) and (13) yield:

$$\frac{dV}{dt} = \alpha_i (N_i - V) G_i(V),$$  

$$G_i(V) = \nu_i \ln(N_i - V) + V - R_i - \nu_i \ln(N_i - R_i - I_i).$$

Integration of (14) provides an analytical solution for the set of equations (4)–(6):

$$F_i^*(V, N_i, I_i, R_i, \nu_i) = \alpha_i(t - t_i^*),$$

$$F_i^* = \int_{R_i + I_i}^{\nu_i (N_i - U) G_i(U)} \frac{dU}{R_i + I_i}.$$  

Thus, for every set of parameters $N_i, I_i, R_i, \nu_i, \alpha_i$ and a fixed value of $V$, integral (16) can be calculated and the corresponding moment of time can be determined from (15).

The final numbers of victims (final accumulated number of cases corresponding to the $i$-th epidemic wave) can be calculated from:

$$V_{\infty} = N_i - S_{\infty}.$$  

To estimate the final day of the $i$-th epidemic wave, we can use the condition:

$$I(t_{if}) = 1$$

which means that at $t > t_{if}$ less than one person still spreads the infection.

**Parameter identification procedure**

In the case of a new epidemic, the values of its parameters are unknown and must be identified with the use of limited data sets. For the first wave of an epidemic starting with one infected person, the number of unknown parameters is only four, since $I_i = 1$ and $R_i = 0$. The corresponding statistical approach was used in [1, 4] to simulate the first COVID-19 pandemic wave in Ukraine and many other countries.

For the next epidemic waves ($i > 1$), the moments of time $t_i^*$ corresponding to their beginning are known. Therefore the exact solution (15)–(17) depends only on five parameters $- N_i, I_i, R_i, \nu_i, \alpha_i$.

$$dV = \alpha_i SI.$$  

$$\frac{dV}{dt} = \alpha_i (N_i - V) G_i(V),$$  

$$G_i(V) = \nu_i \ln(N_i - V) + V - R_i - \nu_i \ln(N_i - R_i - I_i).$$

Integration of (14) provides an analytical solution for the set of equations (4)–(6):

$$F_i^*(V, N_i, I_i, R_i, \nu_i) = \alpha_i(t - t_i^*),$$

$$F_i^* = \int_{R_i + I_i}^{\nu_i (N_i - U) G_i(U)} \frac{dU}{R_i + I_i}.$$  

Thus, for every set of parameters $N_i, I_i, R_i, \nu_i, \alpha_i$ and a fixed value of $V$, integral (16) can be calculated and the corresponding moment of time can be determined from (15). Then functions $R(t)$ and $R(t)$ can be easily calculated with the use of formulas (10) and:

$$S = N_i - V, \quad R = V - I.$$  

The final numbers of victims (final accumulated number of cases corresponding to the $i$-th epidemic wave) can be calculated from:

$$V_{\infty} = N_i - S_{\infty}.$$  

To estimate the final day of the $i$-th epidemic wave, we can use the condition:

$$I(t_{if}) = 1$$

which means that at $t > t_{if}$ less than one person still spreads the infection.
Then the registered number of victims \( V_i \) corresponding to the moments of time \( t_j \) can be used in eq. (16) in order to calculate
\[
F_{i,j} = F_i(V_j, N_j, v_j, I_j, R_i)
\]
for every fixed values of \( N_j, v_j, I_j, R_i \) and then to check how the registered points fit the straight line (15).

Eq. (15) can be rewritten as follows:
\[
y = F_i^* (V_j, N_j, I_j, R_i, v_j) = \alpha_i t - \alpha_i t_j^*.
\]

Assuming
\[
\gamma = \alpha_i, \quad \beta = -\alpha_i t_j^*
\]
we can estimate the values of parameters \( \gamma \) and \( \beta \), by treating the values \( y_j = F_i^* (V_j, N_j, I_j, R_i, v_j) \) and corresponding time moments \( t_j \) as random variables. Then we can use the observations of the accumulated number of cases and the linear regression in order to calculate the coefficients \( \gamma \) and \( \beta \) of the regression line
\[
\hat{y} = \gamma t + \hat{\beta}
\]
using the standard formulas from, e.g., [13]. Values \( \gamma \) and \( \beta \) can be treated as statistics-based estimations of parameters \( \gamma \) and \( \beta \) from relationships (21).

The reliability of the method can be checked by calculating the correlation coefficients \( r_i \) (see e.g., [13]) for every epidemic wave and checking how close its value is to unity. We can use also the \( F \)-test for the null hypothesis that says that the proposed linear relationship (20) fits the data set. The experimental values of the Fisher function can be calculated for every epidemic wave with the use of the formula
\[
F_i = \frac{r_i^2 (n_i - m)}{(1 - r_i^2)(m - 1)}
\]
where \( n_i \) is the number of observations for the \( i \)-th epidemic wave, \( m = 2 \) is the number of parameters in the regression equation. The corresponding experimental value \( F_i \) has to be compared with the critical value \( F_c(k_1, k_2) \) of the Fisher function at a desired significance or confidence level \( (k_1 = m - 1, \quad k_2 = n_i - m) \). When the values \( n_i \) and \( m \) are fixed, the maximum of the Fisher function coincides with the maximum of the correlation coefficient. Therefore, to find the optimal values of parameters \( N_j, v_j, I_j, R_i \), we have to find the maximum of the correlation coefficient for the linear dependence (20). To compare the reliability of different predictions (with different values of \( n_i \)) it is useful to use the ratio \( F_i / F_c(1, n_i - 2) \) at fixed significance level. We will use the level 0.001; corresponding values of \( F_c(1, n_i - 2) \) can be taken from [14]. The most reliable prediction yields the highest \( F_i / F_c(1, n_i - 2) \) ratio.

The exact solution (15)–(17) allows avoiding numerical solutions of differential equations (4)–(6) and significantly reduce the time spent on calculations. In the case of sequential calculation of epidemic waves \( i = 1, 2, 3..., \) it is possible to avoid determining the four optimal unknown parameters \( N_j, v_j, I_j, R_i \), thereby reducing the amount of calculations and difficulties in isolation a maximum of the correlation coefficient. For parameters \( I_j, R_i \) it is possible to use the numbers of \( I \) and \( R \) calculated for the previous wave of epidemic at the moment of time when the following wave began.

Then we need to calculate values \( F_i^* (V_j, N_j, v_j) \), linear regression coefficients (22), correlation coefficient \( r_i \), \( F_i / F_c(1, n_i - 2) \) and to isolate the values of parameters \( N_j, v_j, I_j, R_i \) corresponding to the maximum of \( r_i \). Knowing the optimal values of five parameters \( N_j, I_j, R_i, v_j, \alpha_j \), the SIR curves and other characteristics of the corresponding epidemic wave can be calculated with the use of formulas (10)–(17). This approach has been successfully used in [2, 4]. In particular, six waves of the Covid-19 epidemic in Ukraine and four pandemic waves in the world were calculated.

Segmentation of epidemic waves and their sequential SIR simulations need a lot of efforts. To avoid this, a new method of obtaining the optimal values of SIR parameters was proposed in [3, 4]. First of all we can use the relationship
\[
V_i = I_i + R_i
\]
which follows from (12). To estimate the value \( V_i \), we can use the smoothed accumulated number of cases (e.g., formula (1)). Then
\[
V_i \approx \frac{1}{7} \sum_{j=i-3}^{j=i+3} V_j
\]
where \( i \) corresponds to the moment of time \( t_j^* \). To obtain one more relationship, let us use (7) and (13).
\[ I_i = \frac{1}{\alpha_i (N_i - V_j)} \frac{dV}{dt} \bigg|_{t=t_i}. \]  

(26)

To estimate the average number of new cases \( dV/dt \) at the moment of time \( t_i \), we can use (2). Thus the dependences \((24)-(26)\) allow us to have only two independent parameters \( N_i \) and \( v_j \). To calculate the value of parameter \( \alpha_i \), some iterations can be used (see details in [4]).

**Results**

The COVID-19 pandemic characteristics for Ukraine and Qatar after December 1, 2020 are shown in Figs. 1 and 2. "Circles", "triangles", and "stars" correspond to the accumulated numbers during period of time taken for SIR simulations \( T_c \); December 11–24, before \( T_c \), and after \( T_c \), respectively. The derivatives of the smoothed number of cases (see eq. (1)) are represented by "crosses" (the first derivative, eq. (2)) and "dots" (the second derivative, eq. (3)). In Fig. 1, the red color corresponds to the Ukrainian national statistics [7, 8]; black — to JHU data [10]. There is significant differences in two data sets visible in Fig. 1. The numbers of cases reported by JHU in January 2021 are 30,000–40,000 higher than those presented by the national statistics [7, 8]. However, both data sets are probably incomplete. We will discuss this issue later.

"Dots" in Fig. 2 illustrate some increases in the second derivatives in late December and after the middle of January, which can be treated as the new epidemic waves in Qatar. The increase in the average number of new cases ("crosses" in Fig. 2) confirm this conclusion. Severe jumps of the second derivative on February 11–12 are probably caused by data irregularities.

The results of SIR simulations are shown in Table 2 and Figs. 1 and 2. Since eight epidemic wave was already calculated for Ukraine [1-4], we took the period \( T_c \) December 11–24, 2020 to calculate the ninth epidemic wave in Ukraine with the use of two datasets (presented in Table 1). We have used the same period \( T_c \) and the JHU data set (see Table 1) to calculate the optimal parameters of SIR model and other epidemic characteristics for Qatar. Since previous epidemic waves in this country were not simulated before, we use the name “second” for this wave. The number of observations taken for calculations \( n_9 \) was 14 in all the cases.

It can be seen that two data sets yield rather different values of the optimal parameters for the ninth epidemic wave in Ukraine (especially for \( N_9, S_{9x} \), and \( v_9 \)), nevertheless the final sizes of this wave \( V_{9x} \) and \( \rho_9 \) are rather close; the duration based on the national statistics is one month longer in comparison with the calculations based on JHU data set. Both simulations for the ninth epidemic wave in Ukraine yield slightly higher final sizes in comparison with the eighth wave calculated in [3].

**Figure 1:** Pandemic dynamics in Ukraine (markers) and SIR simulations (lines) calculated with the use of data sets from Tables 1 and 2: national statistics [7, 8] (red) and data set reported by JHU [10] (black). Number of victims \( V(t) = R(t) + R(t) \) — solid lines; numbers of infected and spreading \( I(t) \) multiplied by 10 — dashed; derivatives \( dV/dt \) (eq. (13)) multiplied by 10 — dotted. Markers show accumulated numbers of cases \( V_j \) and \( V_j \) from Table 1 and derivatives. "Circles" correspond to the accumulated numbers of cases taken for calculations (during period of time \( T_c \)); "triangles" — numbers of cases before \( T_c \); "stars" — number of cases after \( T_c \). "Crosses" show the first derivatives (eq. (2)) multiplied by 10, "dots" — the second derivative (eq. (3)) multiplied by 1000.
Figure 2: Pandemic dynamics in Qatar (markers) and SIR simulations (lines) calculated with the use of data sets from Tables 1 and 2. Numbers of victims \( V(t) = I(t) + R(t) \) – solid lines; numbers of infected and spreading \( I(t) \) multiplied by 100 – dashed; derivatives \( dV/dt \) (eq. (13)) multiplied by 100 – dotted. Markers show accumulated numbers of cases \( V_j \) from Table 1 and derivatives.

"Circles" correspond to the accumulated numbers of cases taken for calculations (during period of time \( T_c \)); "triangles" – numbers of cases before \( T_c \); "stars" – number of cases after \( T_c \). "Crosses" show the first derivatives (eq. (2)) multiplied by 100, "dots" – the second derivative (eq. (3)) multiplied by 1000.

Table 2: Calculated optimal values of SIR parameters and other characteristics of the COVID-19 pandemic waves in Ukraine and Qatar.

| Characteristics                                                                 | Ukraine, 9th wave, \( i = 9 \) (JHU), \( V_{i9} \) | Ukraine, 9th wave, \( i = 9 \) (National statistics), \( V_{i9} \) | Qatar, 2nd wave, \( i = 2 \) (JHU), \( V_{i2} \) |
|---------------------------------------------------------------------------------|---------------------------------------------------|---------------------------------------------------|---------------------------------------------------|
| \( I_i \)                                                                        | 62,384.6901000672                                  | 59,686.0031221196                                  | 581.685205791480                                   |
| \( R_i \)                                                                        | 830,900.452757076                                  | 821,069.282592166                                  | 203,678.101450066                                  |
| \( N_i \)                                                                        | 1,524,200.0384                                     | 2,037,235.2                                       | 343,800                                           |
| \( v_i \)                                                                        | 580,121.592967068                                  | 1,141,207.02807484                                 | 203,678.101450066                                  |
| \( \alpha_i \)                                                                   | 2.91326081249e-07                                  | 1.5593061618721e-07                                | 1.26938469796243e-06                               |
| \( \rho_i \)                                                                     | 0.169004550327192                                  | 0.177949115084894                                 | 0.258345863290753                                  |
| \( 1/\rho_i \)                                                                   | 5.91700044799982                                   | 5.6195840003078                                   | 3.86778569781199                                   |
| \( r_i \)                                                                        | 9.98082602967150                                   | 0.998062592012312                                 | 0.99994364528809                                   |
| \( F_j, eq. (23) \)                                                              | 3120.24523044173                                   | 3087.9241796763                                   | 105,931.16920481                                  |
| \( F_j/F_{C1}, eq. (11) \)                                                       | 167.75119916222                                    | 166.0174172815464                                 | 5695.2241507965                                    |
| \( S_{\infty}, eq. (18) \)                                                       | 347,782                                          | 810,539                                           | 188,661                                           |
| \( V_{\infty}, eq. (19) \)                                                       | 1,167,418                                        | 1,226,696                                         | 155,139                                           |

Final day of the epidemic wave, eq. (19) June 15, 2021 | July 14, 2021 | January 16, 2022

Lines in Fig. 1 illustrate the results of SIR simulations of the ninth epidemic wave in Ukraine obtained with the use of two data sets: national statistics \([7, 8]\) (red) and data set reported by JHU \([10]\) (black). Numbers of victims \( V(t) = I(t) + R(t) \) (eq. (12)) are represented by solid lines. Numbers of infected and spreading the infection persons \( I(t) \) are shown by dashed lines. Dotted lines represent the derivatives \( dV/dt \) calculated with the use of eq. (13). This derivative yields the estimation of the daily number of new cases and can be compared with the calculations according to the formula (2) (crosses in Figs. 1 and 2). Fig. 1 illustrates that some discrepancies between red dotted line and red "crosses" appeared only after January 10. These deviations from the theoretical estimates can be explained by the New Year and Orthodox Christmas celebrations.

It can be seen that the accuracy of simulations based on the national statistics is rather good (the deviations between red "stars" and the red solid line are small). The use of JHU data sets yields worth accuracy. Nevertheless, the real numbers of cases already exceed the predicted saturations.
levels $V_{9e}$ for both data sets and corresponding simulations. As of February 20, 2021 the national statistics yields the figure 1,304,456 of accumulated cases in Ukraine (see Table 1). Thus the epidemic observations during 58 days (after $T_c$) demonstrated only 6% exceeding of the saturation level in the case of national statistics.

Lines in Fig. 2 illustrate the results of SIR simulations for Qatar. Numbers of victims $V(t) = I(t) + R(t)$ (eq. (12)) are represented by the solid line. Numbers of infected and spreading the infection persons $I(t)$ are shown by the dashed line. The dotted line represents the derivatives $dV/dt$ calculated with the use of eq. (13). This derivative yields the estimation of the daily number of new cases and can be compared with the calculations according to the formula (2) ("crosses"). In comparison with the case of Ukraine the discrepancies between the dotted line and "crosses" appeared already after December 27, and show that new epidemic waves occurred in Qatar in 2021.

The deviations between "stars" and the solid line in Fig. 2 are not very large, but real numbers of cases already exceed the predicted saturations level $V_{2e}$. As of February 20, 2021 the JHU yields the figure 159,967 of accumulated cases in Qatar (see Table 1). The corresponding value on the $V(t)$ curve is 149,512. Thus the epidemic observations during 58 days (after $T_c$) demonstrated only 6.5% exceeding.

Discussion

It must be noted that the data presented in Table 1 does not show all the COVID-19 cases in Ukraine. Many infected persons are not identified, since they have no symptoms. For example, employees of two kindergartens and two schools in the Ukrainian city of Chmelnytskyi were tested for antibodies to COVID-19 [15]. In total 292 people work in the surveyed institutions. Some of the staff had already fallen ill with COVID-19, or were hospitalized. Therefore, they were tested and registered accordingly. In the remaining tested 241 educators, antibodies were detected in 148, or 61.4%.

Many people know that they are ill, since they have similar symptoms as other members of families, but avoid making tests. Unfortunately, one laboratory confirmed case can correspond to several other cases which are not confirmed and displayed in the official statistics. Special simulations are necessary to take into account this data incompleteness.

The accurate estimation achieved in our study using the SIR model for Qatar new COVID-19 cases is mostly explained by the fact that the used data reflect in some extent the true reality of the spread of the virus in the country. Indeed, the massive daily number of tests performed allows detecting and reporting realistically the true number of COVID-19 cases. In addition the speed announcement of the result of tests helps to report these numbers without any delay reducing the errors in daily reporting.

Conclusions

Overall, presented results of SIR simulations indicate their high accuracy and allow us to make correct medium-term predictions about the expected number of cases and the number of people who spread the infection. The accuracy of long-term forecasts may be limited due to new waves of the epidemic. To improve their accuracy, new simulations must be performed with the use of fresh data sets. It is possible to use the approximate formulas in Chapter 13 of the book [4] to get quick estimates right of the final size and duration of an epidemic wave immediately after its start. We hope that the mass vaccination, which began in Qatar in late December 2020 (in Ukraine, this campaign started two months later) will be able to improve these forecasts. Comparing the results of calculations with the actual number of reported cases can be useful for assessing the effectiveness of vaccinations.

Acknowledgements

The authors are grateful to Oleksii Rodionov for his help in collecting and processing data.

References

[1] Nesteruk I. Simulations and predictions of COVID-19 pandemic with the use of SIR model. Innov Biosyst Bioeng. 2020;4(2):110-21. DOI: 10.20535/ibb.2020.4.2.204274

[2] Nesteruk I. Waves of COVID-19 pandemic. Detection and SIR simulations. medRxiv [Preprint] 2020. DOI: 10.1101/2020.08.03.20167098

[3] Nesteruk I. COVID-19 pandemic dynamics in Ukraine after September 1, 2020. medRxiv [Preprint] 2020. DOI: 10.1101/2020.12.21.20248627
Ключові слова: обережними та солідарними. Можливо, представлені результати можуть бути корисними для оцінки ефективності щеплень.

Висновки. На жаль, наші прогнози не надто оптимістичні: в Україні нові випадки не перестануть з'являтися до червня–липня 2021 року. Очікувана велика тривалість пандемії змушує нас бути висвітленням процесу тестування та ідентифікації параметрів моделі, що були запропоновані раніше.

Результати. Розраховано оптимальні значення параметрів SIR-моделі та представлено деякі прогнози щодо остаточного розміру й тривалості епідемій. Наведено відповідні SIR-криві та порівняння з реальною кількістю випадків.

Мета. Виявити зміни в динаміці пандемії та представити результати SIR-моделювання на основі української національної статистики та даних, повідомлених Університетом Джона Хопкінса (JHU) для України та Катару.

Методика реалізації. Ми використовували метод згладжування для залежності кількості випадків COVID–19 від часу, узагальнює SIR-моделі та представлено деякі прогнози щодо остаточного розміру й тривалості епідемій. Наведено відповідні SIR-криві та порівняння з реальною кількістю випадків.

Висновки. На жаль, наші прогнози не надто оптимістичні: в Україні нові випадки не перестануть з'являтися до червня–липня 2021 року.

Ключові слова: пандемія COVID-19; динаміка епідемії в Україні; динаміка епідемії в Катарі; математичне моделювання інфекційних захворювань; SIR-модель; ідентифікація параметрів; статистичні методи.

І.Г. Нестерук, Н. Бенлага

1 Інститут гідромеханіки НАН України, Київ, Україна
2 КПІ ім. Ігоря Сікорського, Київ, Україна
3 Катарський університет, Доха, Катар

ПРОГНОЗИ ДИНАМИКИ ПАНДЕМИИ COVID-19 В УКРАЇНІ ТА КАТАРЕ НА ОСНОВІ УЗАГАЛЬНЕНОЇ SIR-МОДЕЛІ

Проблематика. Для моделювання збільшення кількості випадків COVID-19 із часом можна використовувати різні набори даних і різні математичні моделі. Оскільки в статистичних даних існують деякі відмінності, то результати моделювання можуть бути різними. Складні математичні моделі містять багато невідомих параметрів, значення яких необхідно визначати, використовуючи різні математичні моделі. Оскільки в статистичних даних існують деякі відмінності, то результати моделювання можуть бути різними.

Мета. Виявити зміни в динаміці пандемії та представити результати SIR-моделювання на основі української національної статистики та даних, повідомлених Університетом Джона Хопкінса (JHU) для України та Катару.

Методика реалізації. Ми використовували метод згладжування для залежності кількості випадків COVID–19, якові епідемічної хвилі, точний розв’язок лінійних диференціальних рівнянь і статистичний підхід для ідентифікації параметрів моделі, що були запропоновані раніше.

Результати. Розраховано оптимальні значення параметрів SIR-моделі та представлено деякі прогнози щодо остаточного розміру й тривалості епідемій. Наведено відповідні SIR-криві та порівняння з реальнію кількістю випадків. Усі випадки можуть з’явитися впродовж усього 2021 року, а в Катарі нові випадки можуть з’являтися впродовж усього 2021 року. Очікувана велика тривалість пандемії змушує нас бути висвітленням процесу тестування та ідентифікації параметрів моделі, що були запропоновані раніше.

Висновки. На жаль, наші прогнози не надто оптимістичні: в Україні нові випадки не перестануть з’являтися до червня–липня 2021 року. Фактична велика тривалість пандемії змушує нас бути обережними та солідарними. Можливо, представлені результати можуть бути корисними для оцінки ефективності щеплень.

Ключові слова: пандемія COVID-19; динаміка епідемії в Україні; динаміка епідемії в Катарі; математичне моделювання інфекційних захворювань; SIR-модель; ідентифікація параметрів; статистичні методи.

І.Г. Нестерук, Н. Бенлага

1 Інститут гідромеханіки НАН України, Київ, Україна
2 КПІ ім. Ігоря Сікорського, Київ, Україна
3 Катарський університет, Доха, Катар

ПРОГНОЗИ ДИНАМИКИ ПАНДЕМИИ COVID-19 В УКРАЇНІ И КАТАРЕ НА ОСНОВЕ ОБОБЩЕННОЙ SIR-МОДЕЛИ

Проблематика. Чтобы смоделировать рост числа случаев COVID-19 со временем, можно использовать различные наборы данных и различные математические модели. Поскольку статистические данные содержат некоторые различия, то результаты...
моделирования могут быть разными. Сложные математические модели содержат множество неизвестных параметров, значения которых необходимо определять, используя ограниченное количество наблюдений за болезнью с течением времени. Даже длительный мониторинг эпидемии может не дать надежных оценок параметров модели из-за постоянного изменения условий тестирования, изоляции инфицированных, условий карантина, мутаций вируса, вакцинации и т.д. Следовательно, необходимы более простые подходы. В частности, предыдущее моделирование динамики эпидемии COVID-19 в Украине основывалось на сглаживании зависимости количества случаев от времени и обобщенной SIR-модели (восприимчивые–инфицированные–удаленные). Эти подходы позволили обнаружить волны пандемии и рассчитать адекватные прогнозы их продолжительности и окончательных размеров. В частности, были исследованы восемь волн пандемии COVID-19 в Украине.

Цель. Выявить изменения в динамике пандемии и представить результаты SIR-моделирования на основе национальной статистики Украины и данных, предоставленных Университетом Джона Хопкинса (JHU) для Украины и Катара.

Методика реализации. Мы используем метод сглаживания зависимостей числа случаев от времени, обобщенную SIR-модель динамики произвольной волны эпидемии, точное решение линейных дифференциальных уравнений и статистический подход для идентификации параметров модели, которые были предложены ранее.

Результаты. Были рассчитаны оптимальные значения параметров SIR-модели и представлены прогнозы относительно окончательных размеров и продолжительности эпидемий. Приведены соответствующие SIR-кривые и сравнение с реальным числом случаев.

Выводы. К сожалению, наши прогнозы не очень оптимистичные: в Украине новые случаи не перестанут появляться до июня–июля 2021 года; в Катаре новые случаи заболевания могут появляться в течение всего 2021 года. Ожидаемая длительность пандемии заставляет нас проявлять осторожность и солидарность. Возможно, представленные результаты могут быть полезны для оценки эффективности вакцинаций.

Ключевые слова: пандемия COVID-19; динамика эпидемии в Украине; динамика эпидемии в Катаре; математическое моделирование инфекционных заболеваний; SIR-модель; идентификация параметров; статистические методы.