



OPEN ACCESS

EDITED BY

Dmytro Chumachenko,
National Aerospace University – Kharkiv
Aviation Institute, Ukraine

REVIEWED BY

Tetyana Chumachenko,
Kharkiv National Medical University, Ukraine
Sergiy Yakovlev,
Lodz University of Technology, Poland

*CORRESPONDENCE

Igor Nesteruk
✉ inesteruk@yahoo.com

RECEIVED 13 January 2025

ACCEPTED 31 January 2025

PUBLISHED 24 February 2025

CITATION

Nesteruk I (2025) General SIR model for
visible and hidden epidemic dynamics.
Front. Artif. Intell. 8:1559880.
doi: 10.3389/frai.2025.1559880

COPYRIGHT

© 2025 Nesteruk. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

General SIR model for visible and hidden epidemic dynamics

Igor Nesteruk*

Institute of Hydromechanics, National Academy of Sciences of Ukraine, Kyiv, Ukraine

To simulate hidden epidemic dynamics connected with asymptomatic and unregistered patients, a new general SIR model was proposed. For some cases, the analytical solutions of the set of 5 differential equations were found, which allow simplifying the parameter identification procedure. Two waves of the pertussis epidemic in England in 2023 and 2024 were simulated with the assumption of zero hidden cases. The accumulated and daily numbers of cases and the duration of the second wave were predicted with rather high accuracy. If the trend will not change, the monthly figure of 9 new pertussis cases (as it was in January–February 2023) can be achieved only in May 2025. The proposed approach can be recommended for both simulations and predictions of different epidemics.

KEYWORDS

mathematical modeling of infection diseases, SIR model, parameter identification, pertussis epidemic in England, hidden epidemic dynamics

1 Introduction

Asymptomatic and unregistered cases are characteristic of almost all infectious diseases, in particular, SARS-CoV-2 ([Mass coronavirus testing in Slovakia-a, 2020](#); [Mass coronavirus testing in Slovakia-b, 2020](#); [An experiment with mass testing for COVID-19 was conducted in Khmelnytsky, n.d.](#); [Schreiber et al., 2023](#); [Fowlkes et al., 2022](#); [Shang et al., 2022](#)) and pertussis ([Craig et al., 2020](#)) are no exception. The percentage of asymptomatic patients can be age dependent and lead to huge differences in registered numbers of cases for countries with young and old population ([Davies et al., 2020](#); [Nesteruk, 2024b](#); [Nesteruk and Keeling, 2023](#)). Some theoretical estimations of the visibility coefficient β —the ratio of real infections to the registered ones can be found in [Nesteruk \(2024b\)](#), [Nesteruk \(2021e\)](#), [Nesteruk \(2021c\)](#), and [Nesteruk \(2021d\)](#). In this study we will use the concepts of the classical SIR (susceptible-infectious-removed) model ([Kermack and McKendrick, 1927](#); [Weiss, 2013](#); [Daley and Gani, 2005](#); [Keeling and Rohani, 2008](#); [Cherniha, 2020](#); [Mohammadi et al., 2021](#); [Nesteruk, 2021a](#)), its generalization for simulations of different epidemic waves ([Nesteruk, 2021a](#); [Nesteruk, 2021b](#); [Nesteruk, 2023b](#)) and procedures of parameter identification ([Nesteruk, 2023b](#); [Nesteruk, 2017](#)). Numerous improvements of SIR model (see, e.g., [Hethcote, 2000](#); [Nakamura et al., 2020](#); [Britton, 2004](#); [Pesco et al., 2014](#); [Nesteruk, 2023a](#)) do not take into account the visibility coefficient.

The obtained theoretical results will be applied for simulations of the pertussis (whooping cough) epidemic in England in 2023 and 2024 ([Confirmed cases of pertussis in England by month, n.d.](#)). This disease increases the risk of infant fatality and became a serious problem in many countries including the developed ones ([Pesco et al., 2014](#); [Confirmed cases of pertussis in England by month, n.d.](#)). Numerical differentiation of the monthly numbers of new cases revealed two waves of the epidemic in England (before and after November 2023) ([Nesteruk, 2024a](#)). Due to the absence of necessary amount of observations, SIR simulations were performed in [Nesteruk \(2024a\)](#) only for the first wave. In this study we will use the new approach for simulation of both waves of the pertussis epidemic and compare the predictions with the recent statistical data.

2 Differential equations and initial conditions

For every epidemic wave i , let us divide the compartment of infectious persons $I(t)$ (t is time) into visible (registered) and hidden (invisible/asymptomatic and unregistered) parts $I = I^{(v)} + I^{(h)}$ and suppose that these persons are appearing according to the visibility coefficient $\beta_i \geq 1$ and removing with rates $\rho_i^{(v)} I^{(v)}$ and $\rho_i^{(h)} I^{(h)}$. Then the general SIR model (Nesteruk, 2021a; Nesteruk, 2021b; Nesteruk, 2023b) takes the following form:

$$\frac{dS}{dt} = -\alpha_i S (I^{(v)} + I^{(h)}) \quad (1)$$

$$\frac{dI^{(v)}}{dt} = \frac{\alpha_i}{\beta_i} S (I^{(v)} + I^{(h)}) - \rho_i^{(v)} I^{(v)} \quad (2)$$

$$\frac{dI^{(h)}}{dt} = \frac{(\beta_i - 1)\alpha_i}{\beta_i} S (I^{(v)} + I^{(h)}) - \rho_i^{(h)} I^{(h)} \quad (3)$$

$$\frac{dR^{(v)}}{dt} = \rho_i^{(v)} I^{(v)} \quad (4)$$

$$\frac{dR^{(h)}}{dt} = \rho_i^{(h)} I^{(h)} \quad (5)$$

The compartment of removed persons $R(t)$ is also divided into visible (registered) and hidden parts $R = R^{(v)} + R^{(h)}$. Infection and removal rates (α_i , $\rho_i^{(v)}$, $\rho_i^{(h)}$) and the visibility coefficient β_i are supposed to be constant for every epidemic wave, i.e., for the time periods: $t_i^* \leq t \leq t_{i+1}^*$, $i = 1, 2, 3, \dots$. Summarizing Equations 1–5 yields zero value of the derivative $d(S + I^{(v)} + I^{(h)} + R^{(v)} + R^{(h)})/dt$.

Then the sum:

$$N_i = S + I^{(v)} + I^{(h)} + R^{(v)} + R^{(h)} \quad (6)$$

must be constant for every epidemic wave. We will consider the value N_i to be an unknown parameter of the model corresponding to the i -th wave, which is not equal to the known volume of population and must be estimated by observations. There is no need to assume that before the outbreak all people are susceptible, since many of them are protected by their immunity, distance, lockdowns, etc. Thus, we will not reduce the problem to a 4-dimensional one. It means that the solution can be obtained by numerical integration of the set of 5 differential Equations 1–5. Nevertheless, there are some separate cases, when analytical solutions are possible (see next Section).

Taking into account Equation 6, the initial conditions for the set of Equations 1–5 at the beginning of every epidemic wave t_i^* can be written as follows:

$$\begin{aligned} I^{(v)}(t_i^*) &= I_{vi}, I^{(h)}(t_i^*) = I_{hi}, R^{(v)}(t_i^*) = R_{vi}, R^{(h)}(t_i^*) = \\ R_{hi} S(t_i^*) &= N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi} \end{aligned} \quad (7)$$

If at moment t_i^* all previously infected persons are removed, we can take into account only cases starting to appear during i -th wave and use the initial conditions:

$$I_{vi} = 1, I_{hi} = \beta_i - 1, R_{vi} = 0, R_{hi} = 0 \quad (8)$$

3 Examples of analytical solutions

Let us introduce the functions corresponding to the accumulated numbers of visible and hidden cases:

$$V^{(v)} = I^{(v)} + R^{(v)}, V^{(h)} = I^{(h)} + R^{(h)} \quad (9)$$

Then it follows from Equations 2–5 that

$$\frac{dV^{(v)}}{dt} = \frac{\alpha_i}{\beta_i} S (I^{(v)} + I^{(h)}), \frac{dV^{(h)}}{dt} = \frac{(\beta_i - 1)\alpha_i}{\beta_i} S (I^{(v)} + I^{(h)}) \quad (10)$$

Dividing Equation 10 by Equation 1 yields:

$$\frac{dV^{(v)}}{dS} = -\frac{1}{\beta_i}, \frac{dV^{(h)}}{dS} = -\frac{(\beta_i - 1)}{\beta_i} \quad (11)$$

and simple linear solutions taking into account initial Equation 7:

$$V^{(v)} = -\frac{S}{\beta_i} + \frac{N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi}}{\beta_i} + I_{vi} + R_{vi}, \quad (12)$$

$$V^{(h)} = -\frac{(\beta_i - 1)S}{\beta_i} + \frac{(\beta_i - 1)(N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi})}{\beta_i} + I_{hi} + R_{hi}. \quad (13)$$

Equations 12, 13 allow obtaining simple linear relationship:

$$V^{(h)} = (\beta_i - 1)V^{(v)} - (\beta_i - 1)(I_{vi} + R_{vi}) + I_{hi} + R_{hi}$$

which demonstrates that the ratio of total accumulated cases $V = V^{(v)} + V^{(h)}$ to the registered ones:

$$\frac{V}{V^{(v)}} = \beta_i - \frac{(\beta_i - 1)(I_{vi} + R_{vi}) - I_{hi} - R_{hi}}{V^{(v)}} \quad (14)$$

is not constant and equals β_i only approximately at large $V^{(v)}$ numbers. Equation 14 limits the accuracy of the approach used in Nesteruk (2021e), Nesteruk (2021c), and Nesteruk (2021d).

Introducing

$$I = I^{(v)} + I^{(h)}, I^* = \rho_i^{(v)} I^{(v)} + \rho_i^{(h)} I^{(h)} \quad (15)$$

summarizing Equations 2, 3 and dividing by Equation 1 yield the following differential equation:

$$\frac{dI}{dS} = -1 + \frac{I^*}{\alpha_i I S} \quad (16)$$

In 3 separate cases:

- I $\rho_i^{(v)} = \rho_i^{(h)} = \rho_i$
- II $I^{(v)} \gg I^{(h)}, I^{(v)} \approx I, \rho_i = \rho_i^{(v)}$
- III $I^{(v)} \ll I^{(h)}, I^{(h)} \approx I, \rho_i = \rho_i^{(h)}$

Equation 16 simplifies and has an analytical solution taking into account the initial Equation 7:

$$\frac{dI}{dS} \cong -1 + \frac{\rho_i}{\alpha_i S}, \quad (17)$$

$$I \cong -S + N_i - R_{vi} - R_{hi} + \frac{\rho_i}{\alpha_i} \ln \frac{S}{N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi}} \quad (18)$$

Equations 17, 18 exact in the case (I) and approximate in cases (II) and (III).

Putting Equation 18 into Equation 1 and integration yield:

$$F(S) \cong \alpha_i (t - t_i^*), \quad (19)$$

$$F(S) \cong \int_{N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi}}^S \frac{dU}{U \left(-\frac{\rho_i}{\alpha_i} \ln \frac{U}{N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi}} \right)} \quad (20)$$

It follows from Equations 1, 2, 15 that:

$$\frac{dI^{(v)}}{dS} = -\frac{1}{\beta_i} + \frac{\rho_i^{(v)} I^{(v)}}{\alpha_i S I}, \quad (21)$$

Taking into account that

$$\frac{dF}{dS} \cong -\frac{1}{SI}$$

(see Equations 18, 20), the solution of the non-homogenous linear Equation 21 satisfying the first initial Equation 7 can be written as follows:

$$I^{(v)}(S) \cong I_{vi} \exp \left(-\frac{\rho_i^{(v)} F(S)}{\alpha_i} \right) - \frac{1}{\beta_i} \exp \left(-\frac{\rho_i^{(v)} F(S)}{\alpha_i} \right) \times \int_{N_i - I_{vi} - I_{hi} - R_{vi} - R_{hi}}^S \exp \left(\frac{\rho_i^{(v)} F(U)}{\alpha_i} \right) dU \quad (22)$$

With the use of Equations 9, 15 it is possible to express other functions as follows:

$$\begin{aligned} I^{(h)}(S) &\cong I(S) - I^{(v)}(S); \\ R^{(v)}(S) &\cong V^{(v)}(S) - I^{(v)}(S); \\ R^{(h)}(S) &\cong V^{(h)}(S) - I^{(h)}(S) \end{aligned} \quad (23)$$

Then for every value of S, all unknown functions can be calculated with the use of Equations 12, 13, 18, 22, 23. Corresponding moments of time can be found with the use of Equations 19, 20. Thus, Equations 12, 13, 18–20, 22, 23 yield an approximate analytical solution of the set of differential Equations 1–5 with the initial Equation 7. In the case (I) and when $I^{(v)} = I$, this solution is exact. For $I^{(v)} = I$, there is no need in Equations 22, 23 and corresponding formulas obtained in Nesteruk (2021a), Nesteruk (2021b), Nesteruk (2023b) are also valid.

4 Examples of parameter identifications and predictions

The analytical solution simplifies the procedure of identification of unknown parameters, since there is no need in numerical integration of differential Equations 1–5. In particular, having the set of accumulated cases $V_j^{(v)}$ registered at moments t_j , we can calculate corresponding values S_j with the use of the linear Equation 12 for any values of unknown constant parameters appearing in Equations 1–7. Then Equation 20 allows calculating values $F_j = F(S_j)$. Due to the linear relationship (Equation 19 shows that there is a linear dependence between time and the function F (Equation 20), which depends on the accumulated numbers of cases), standard linear regression formulas (Draper and Smith, 1998) can be used to calculate the correlation coefficient r and values of parameters α_i and t_i^* . The optimal values of model parameters (providing the best fitting between the theoretical $V^{(v)}(t)$ curves and the results of observations $V_j^{(v)}$) correspond to the maximum value of the correlation coefficient r . Thus, the parameter identification problem can be reduced to the problem of searching the maximum of complicated but analytical function r . For $I^{(v)} = I$ (a completely visible epidemic), such approach was successfully used to simulate and predict the dynamics of mysterious children disease (Nesteruk, 2017), COVID-19 pandemic (Nesteruk, 2021a; Nesteruk, 2021b; Nesteruk, 2023b) and the pertussis epidemic in England (Nesteruk, 2024a).

Let us illustrate the parameter identification procedure for two waves of the pertussis epidemic in England in 2023 and 2024 discussed in (Nesteruk, 2024a). The accumulated confirmed numbers of cases ($V_j^{(v)}$) and corresponding moments of time t_j are

TABLE 1 Accumulated numbers of confirmed pertussis cases in England in 2023 and 2024 and estimations of the average daily numbers of visible cases.

Months, starting with January 2023	Days, starting with 1 January 2023, t_j	Accumulated numbers of visible cases, $V_j^{(v)}$, (Confirmed cases of pertussis in England by month, n.d.)	Calculated daily numbers of new visible cases at moments t_j dV/dt , equation 1 in Nesteruk (2024a)
1	31	9	–
2	59	18	0.35426
3	90	30	0.52688
4	120	50	0.86559
5	151	83	1.41559
6	181	136	2.04462
7	212	208	2.66129
8	243	301	3.20
9	273	403	3.34516
10	304	505	3.47849
11	334	615	5.75269
12	365	858	12.87096
13	396	1,413	24.79644
14	425	2,332	38.86096
15	456	3,759	58.23280
16	486	5,872	84.44247
17	517	8,924	89.67581
18	547	11,351	67.65968
19	578	13,038	44.27419
20	609	14,096	28.79785
21	639	14,800	19.94301
22	670	15,309	–

listed in Table 1 according to the official site of UK government [(Confirmed cases of pertussis in England by month, n.d.), version available on 10 August 2024, the last 4 figures were taken from 10 December 2024 version]. The values $V_j^{(v)}$ were used to calculate approximate daily numbers of new cases dV/dt at moments t_j according to the Equation 1 from (Nesteruk, 2024a) (see the last column in Table 1).

Since the general problem contains 10 unknown parameters, their identification needs high performance computing and applying AI methods even for the analytical solution Equations 12, 13, 18–20, 22, 23. When this solution is approximate, the optimal values of parameters will contain discrepancies, which can reduce the accuracy of predictions. For our example, let us take the case of exact solution $I^{(v)} = I(\beta = 1)$ and assume that at the beginning of every new epidemic wave all infectious persons from the previous waves are removed. Then we can use initial Equation 8, perform simulations and then add cases accumulated at moments when the monthly numbers of visible cases started to increase. For every wave we will have only four unknown parameters N_b , α_i , $\rho_i^{(v)}$ and t_i^* . Due to linear relation 19 and using linear regression, only two of these parameters are independent.

The first and second epidemic waves were simulated with the use of $V_j^{(v)}$ and t_j corresponding to $j = 3–9$ and $j = 11–18$, respectively. The optimal values of parameters (corresponding to the maximum of correlation coefficients $r = 0.999721131761662$ (0.999822556136920)) are:

$$\begin{aligned} N_i &= 50739.992 \text{ (3657890.47292358)}; \\ \alpha_i &= 1.50488802805402e-05 \text{ (2.11266650706657e-06) [day]}^{-1}; \\ \rho_i^{(v)} &= 0.752503596035381 \text{ (7.7088126606047) [day]}^{-1}; \\ t_i^* &= 89.1566573802040 \text{ (333.429529143536) days.} \end{aligned}$$

(figures in brackets correspond to the second wave). It should be noted that these optimal values are very different for the first and second waves and differ from the figures in Nesteruk (2024a) for the simulation of the first wave using observations with $j = 1–10$.

Using these optimal values in the analytical solution Equations 10, 12, 18–20, 22, 23 yielded the theoretical curves shown in Figure 1 (solid and dashed for the first and the second wave, respectively). The predicted values (see blue and black “crosses” for $j = 19–22$, July–October 2024) are in good agreement with the theoretical blue and black curves. To estimate the accuracy of 4-month prediction, let us take the accumulated number of visible cases $V_{22}^{(v)} = 15,309$ corresponding to the end of October 2024 (see Table 1) and compare with the theoretical value $V^{(v)}|_{t=670}$

$= 17,430$ corresponding to the blue dashed line in Figure 1. After adding 505 cases accumulated at t_{10} and extracted for the simulation of the second wave (compare blue “crosses” and “circles” in Figure 1), we obtain the accuracy $(17,935-15,309)/15,309$ around 17%. Since the final number of visible cases decreases with

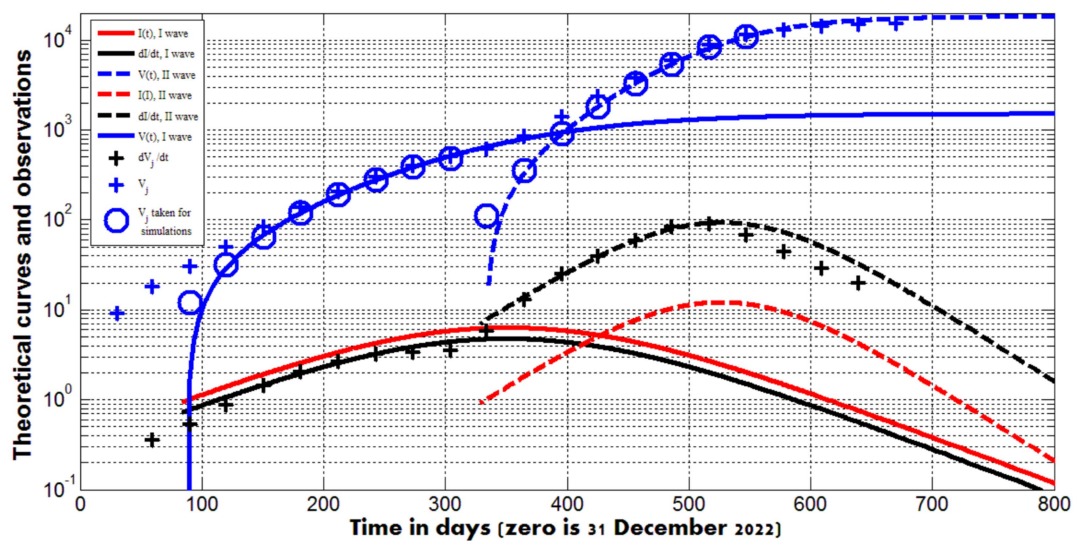


FIGURE 1

Accumulated numbers of visible pertussis cases (blue curves, the first Equation 9); the average daily numbers of new cases (black curves, the first Equation 10); numbers of infectious persons (red curves, Equation 18). "Circles" represent the confirmed numbers of cases $V_j^{(v)}$ taken for identification of parameters of the first ($j = 3-10$) and second ($j = 11-18$) waves; blue "crosses" – all confirmed numbers of cases $V_j^{(v)}$ listed in Table 1; black "crosses" – results of calculations of approximate daily numbers of new cases at moments t_j listed in the last column of Table 1.

the increase of the visibility coefficient (see the next Section), we can expect to obtain a lower theoretical value $V^{(v)}|_{t=670}$ and

better accuracy after the real visibility coefficient will be calculated and taken into account. The accuracy of 17% is comparable with the long-time predictions for the first waves of the COVID-19 in different countries (Nesteruk, 2021a) and even for the case $\beta = 1$ is likely to allow healthcare professionals to develop the right strategy. The average daily numbers of new cases will be less than 1.0 only in March 2025. If trend will not change, the monthly figure of 9 new cases (as it was before starting the first wave, see Table 1) can be achieved only in May 2025.

5 Examples of exact solutions at different values of the visibility coefficient

The use of initial Equation 8 allows reducing the numbers of unknown parameters by 4. Then, Equations 12, 14 yield

$$V^{(v)} = \frac{N_i - S}{\beta_i}, \quad (24)$$

$$V = \beta_i V^{(v)} \quad (25)$$

According to Equation 25 the real accumulated numbers of new cases V are β_i times higher than visible figures $V^{(v)}$ registered during the fixed epidemic wave, if all infectious patients were removed before this wave started.

In the case (I), i.e., equal removing rates for visible and hidden patients $\rho_i^{(v)} = \rho_i^{(h)} = \rho_i$ another simplification can be obtained with the use of Equation 18:

$$I = -S + N_i + \frac{\rho_i}{\alpha_i} \ln \frac{S}{N_i - \beta_i} \quad (26)$$

Assuming that spreading the infection stops when the real number of infectious I (visible and hidden) is less than 1.0, the corresponding final number of susceptible S_f can be obtained as a solution of the non-linear equation:

$$S_f + 1 = N_i + \frac{\rho_i}{\alpha_i} \ln \frac{S_f}{N_i - \beta_i} \quad (27)$$

following from Equation 26 and allowing us to calculate the corresponding final accumulated numbers of visible and total cases with the use of Equations 24, 25.

Figure 2 represents the results of calculations for the optimal values of parameters corresponding to two pertussis waves in England (see previous Section) and the first COVID-19 pandemic waves in Austria and the UK (in brackets), (Nesteruk, 2021a):

$$\begin{aligned} N_i &= 75176.032 \text{ (479782.4);} \\ \alpha_i &= 1.924971386379e-05 \text{ (9.1371956639e-07) [day]}^{-1}; \\ \rho_i^{(v)} &= 1.29635017900866 \text{ (0.330545378991741) [day]}^{-1}. \end{aligned}$$

Dashed lines demonstrate that the final accumulated numbers of all cases (Equations 25, 26) very slightly depend on the visibility coefficient. The final accumulated numbers of visible cases (Equations 24, 26) diminish with the increase of β_i (see solid curves). The values of other parameters are fixed and

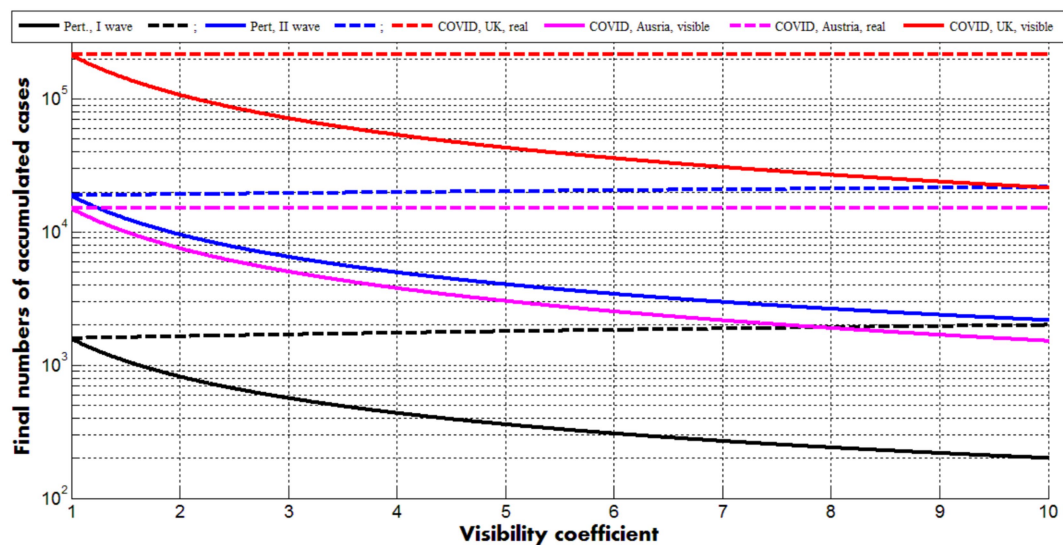


FIGURE 2

Solid curves represent final numbers of visible (registered) cases (Equations 24, 27) dashed ones—final numbers of all cases (registered and unregistered, Equations 25, 27). Back and blue lines correspond to the optimal values of parameters (listed in previous Section) for the first and second pertussis waves in England, respectively. Magenta and red curves show the results for the first COVID-19 waves in Austria and the UK (Nesteruk, 2021a), respectively.

correspond to the case $\beta_i=1$. The values $V_j^{(v)}$ for $j = 20, 21, 22$, which are smaller than theoretical prediction for the second pertussis wave (compare blue “crosses” and the blue dashed line in Figure 1), reflect reducing the final value of $V^{(v)}$ for $\beta_i > 1$. Nevertheless, good estimations of the visibility coefficient can be obtained only with the use of all parameters. Since removing rates can be different for symptomatic and asymptomatic patients, a general parameter identification procedure needs a numerical solution of the set of differential Equations 1–5 and huge numbers of calculations, which can be performed only with the use of high performance computing and AI methods. The full parameter sensitivity analysis will be considered in future research.

With the use of Equations 24, 10, 26 can be rewritten as follows:

$$I = \beta_i V^{(v)} + \frac{\rho_i}{\alpha_i} \ln \frac{N_i - \beta_i V^{(v)}}{N_i - \beta_i}, \quad (28)$$

$$\frac{dV^{(v)}}{dt} = \frac{\alpha_i}{\beta_i} (N_i - \beta_i V^{(v)}) I, \quad \frac{dV}{dt} = \alpha_i (N_i - \beta_i V^{(v)}) I \quad (29)$$

Figure 3 represents the calculations of real number of infectious I (visible and hidden, Equation 28) and visible and real numbers of new daily cases (Equation 29) versus accumulated numbers of visible cases $V^{(v)}$ for different values of the visibility coefficient. Values of other parameters correspond to the second pertussis wave in England (listed in the previous Section).

The maximum values of infectious slightly increase with the increase of the visibility coefficient (compare red lines in Figure 3).

The black and the solid blue curves show the same but more pronounced trend for the real numbers of new daily cases. The positions of the maxima on these lines are very close to ones on the red curves. Corresponding moments of time can be calculated with the use of Equations 19, 20, 24. Thus, the average daily numbers of visible cases reflects trends in real numbers of infectious persons (symptomatic and asymptomatic) and can be used to control epidemics. The numbers of new visible daily cases decrease with the increase of visibility coefficient (for fixed values of other parameters and $\rho_i^{(v)} = \rho_i^{(h)} = \rho_p$ see blue curves in Figure 3). The final numbers of visible cases demonstrate the same trend (see solid curves in Figure 2).

6 Conclusion

To simulate hidden epidemic dynamics connected with asymptomatic and unregistered patients, a new general SIR model was proposed containing 5 unknown functions. For some cases, the analytical solutions of the set of 5 differential equations were found which allow simplifying the parameter identification procedure. Two waves of the pertussis epidemic in England in 2023 and 2024 were simulated for the case of zero hidden cases. Observations of accumulated visible numbers of cases during 4 months revealed rather high accuracy of predictions. If trend will not change, the monthly figure of 9 new pertussis cases (as it was in January–February 2023) can be achieved only in May 2025. The proposed approach can be recommended both for preliminary simulations of different epidemics (supposing zero hidden cases) and for further research, using presented analytical solutions or numerical integration of differential equations. The theoretical estimations of numbers of hidden cases will allow healthcare professionals to

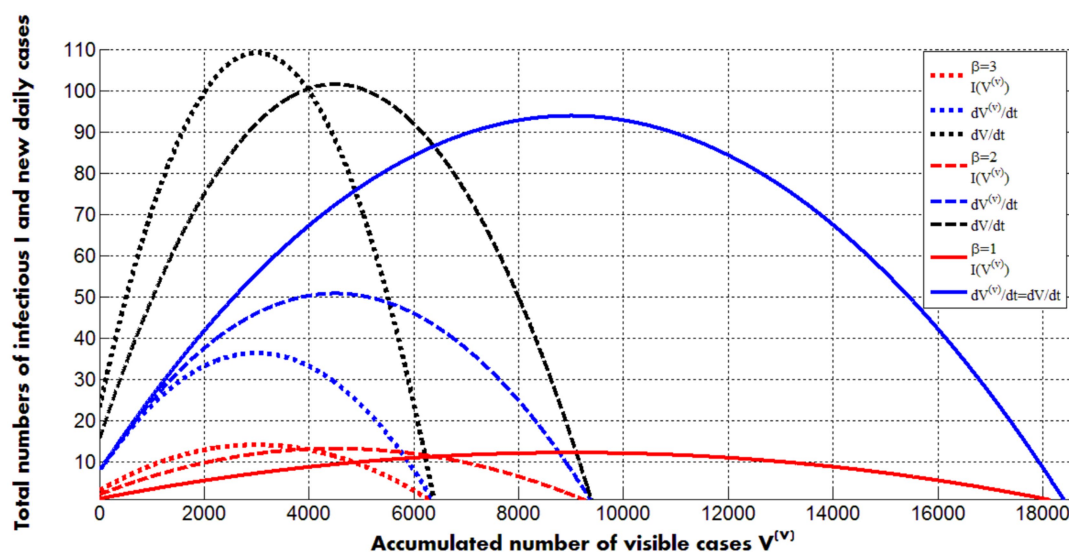


FIGURE 3

Solid and dotted curves represent calculations for values of the visibility coefficient 1; 2 and 3, respectively. Other values of parameters correspond to the second pertussis wave in England in 2023 and 2024 (listed in the previous Section). Red color corresponds to the real number of infectious persons (symptomatic and asymptomatic, Equation 28). Blue lines represent visible numbers of new daily cases; black ones - real (registered and hidden) numbers of new daily cases (Equations 28, 29).

know the real sizes of epidemics and to develop the right strategy without expensive mass testing.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

IN: Conceptualization, Data curation, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing, Validation.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The author is grateful to Ulrike Tillmann, James Robinson, Robin Thompson, Matt Keeling, Paul Brown, and Oleksii Rodionov for their

support and providing very useful information. This paper was written with the support of the INI-LMS Solidarity Programme at the University of Warwick, UK.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

An experiment with mass testing for COVID-19 was conducted in Khmelnytsky. Podillya news. Available at: <https://podillyanews.com/2020/12/17/>

u-shkolah-hmelnytskogo-provely-eksperyment-z-testuvannyam-na-covid-19/ (Accessed November 23, 2024)

- Britton, N. F. (2004). *Essential mathematical biology*. India: Springer (India) Pvt. Limited, 352.
- Cherniha, V. (2020). Davydovych, a mathematical model for the COVID-19 outbreak and its applications. *Symmetry* 12:990. doi: 10.3390/sym12060990
- Confirmed cases of pertussis in England by month. GOV.UK. Available at: www.gov.uk. (Accessed November 23, 2024)
- Craig, R., Kunkel, E., Crowcroft, N. S., Fitzpatrick, M. C., de Melker, H., Althouse, B. M., et al. (2020). Asymptomatic infection and transmission of pertussis in households: a systematic review. *Clin. Infect. Dis.* 70, 152–161. doi: 10.1093/cid/ciz531
- Daley, D. J., and Gani, J. (2005). *Epidemic modeling: An introduction*. eds. C. Cannings, F. C. Hoppensteadt, and L. A. Segel (New York: Cambridge University Press).
- Davies, N. G., Klepac, P., Liu, Y., Prem, K., and Jit, M. (2020). MCMID COVID-19 working group; Eggo RM. (2020). Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* 26, 1205–1211. doi: 10.1038/s41591-020-0962-9
- Draper, N. R., and Smith, H. (1998). *Applied regression analysis*. 3rd Edn: John Wiley.
- Fowlkes, A. L., Yoon, S. K., Lutrick, K., Gwynn, L., Burns, J., Grant, L., et al. (2022). Effectiveness of 2-dose BNT162b2 (Pfizer BioNTech) mRNA vaccine in preventing SARS-CoV-2 infection among children aged 5–11 years and adolescents aged 12–15 years -PROTECT cohort, July 2021–February 2022. *MMWR Morb. Mortal Wkly. Rep.* 71, 422–428. doi: 10.15585/mmwr.mm7111e1
- Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Rev.* 42, 599–653. doi: 10.1137/S0036144500371907
- Keeling, M., and Rohani, P. (2008). *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton University Press.
- Kermack, W. O., and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *J. Royal Stat. Soc. Ser. A.* 115, 700–721.
- Mass coronavirus testing in Slovakia-a. (2020). Available at: <https://edition.cnn.com/2020/11/02/europe/slovakia-mass-coronavirus-test-intl/index.html> (Accessed November 23, 2024)
- Mass coronavirus testing in Slovakia-b. (2020). Available at: <https://www.voanews.com/covid-19-pandemic/slovakias-second-round-coronavirus-tests-draws-large-crowds> (Accessed November 23, 2024)
- Mohammadi, A., Meniailov, I., Bazilevych, K., Yakovlev, S., and Chumachenko, D. (2021). Comparative study of linear regression and SIR models of COVID-19 propagation in Ukraine before vaccination. *Radioelectron. Computer Syst.* 3, 5–18. doi: 10.32620/reks.2021.3.01
- Nakamura, G. M., Cardoso, G. C., and Martinez, A. S. (2020). Improved susceptible-infectious-susceptible epidemic equations based on uncertainties and autocorrelation functions. *R. Soc. Open Sci.* 7:191504. doi: 10.1098/rsos.191504
- Nesteruk, I. (2017). Statistics based models for the dynamics of Chernivtsi children disease. *Naukovi Visti NTUU KPI.* 5, 26–34. doi: 10.20535/1810-0546.2017.5.108577
- Nesteruk, I. (2021e). Visible and real sizes of new COVID-19 pandemic waves in Ukraine. *Innov Biosyst Bioeng.* 5, 85–96. doi: 10.20535/ibb.2021.5.2.230487
- Nesteruk, I. (2021c). Influence of possible natural and artificial collective immunity on new COVID-19 pandemic waves in Ukraine and Israel. *Exploratory Res. Hypothesis Med.* doi: 10.14218/ERHM.2021.00044
- Nesteruk, I. (2021d). The real COVID-19 pandemic dynamics in Qatar in 2021: simulations, predictions and verifications of the SIR model. *Semina: Ciências Exatas Tecnol.* 42, 55–62. doi: 10.5433/1679-0375.2021v42n1Suplp55
- Nesteruk, I. (2021a). *COVID19 pandemic dynamics*. Singapore: Springer Nature.
- Nesteruk, I. (2021b). Detections and SIR simulations of the COVID-19 pandemic waves in Ukraine. *Comput. Math. Biophys.* 9, 46–65. doi: 10.1515/cmb-2020-0117
- Nesteruk, I. (2023b). Improvement of the software for modeling the dynamics of epidemics and developing a user-friendly interface. *Infectious Disease Modelling* 8, 806–821. doi: 10.1016/j.idm.2023.06.003
- Nesteruk, I. (2023a). Endemic characteristics of SARS-CoV-2 infection. *Sci. Rep.* 13:14841. doi: 10.1038/s41598-023-41841-8
- Nesteruk, I. (2024b). Trends of the COVID-19 dynamics in 2022 and 2023 vs. the population age, testing and vaccination levels. *Front. Big Data* 6:1355080. doi: 10.3389/fdata.2023.1355080
- Nesteruk, I. (2024a). Mathematical simulations of the pertussis epidemic in England. International workshop Profit AI 2024. Cambridge, MA, USA.
- Nesteruk, I., and Keeling, M. (2023). Population age as a key factor in the COVID-19 pandemic dynamics. *Research Square* 30:2023. doi: 10.21203/rs.3.rs-3682693/v1
- Pesco, P., Bergero, P., Fabricius, G., and Hozbor, D. (2014). Modelling the effect of changes in vaccine effectiveness and transmission contact rates on pertussis epidemiology. *Epidemics* 7, 13–21. doi: 10.1016/j.epidem.2014.04.001
- Schreiber, P. W., Scheier, T., Wolfensberger, A., Saleschus, D., Vazquez, M., Kouyos, R., et al. (2023). Parallel dynamics in the yield of universal SARS-CoV-2 admission screening and population incidence. *Sci. Rep.* 13:7296. doi: 10.1038/s41598-023-33824-6
- Shang, W., Kang, L., Cao, G., Wang, Y., Gao, P., Liu, J., et al. (2022). Percentage of asymptomatic infections among SARS-CoV-2 omicron variant-positive individuals: a systematic review and Meta-analysis. *Vaccines* 10:1049. doi: 10.3390/vaccines10071049
- Weiss, H. (2013). The SIR model and the foundations of public health. *MatMat* 3, 1–17.