Forecasting SARS-CoV-2 Epidemic Dynamic in Poland with the pDyn Agent-Based Model

Karol Niedzielewski (kniedzie@icm.edu.pl)
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0001-7405-2207

Rafał P. Bartczuk
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0002-0433-7327

Natalia Bielczyk
Ontology of Value https://orcid.org/0000-0003-1604-9143

Dominik Bogucki
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0009-0000-9473-2242

Filip Dreger
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0003-1812-8216

Grzegorz Dudziuk
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw,

Łukasz Górski
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0003-0871-6575

Magdalena Gruziel-Słomka
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0001-6882-7408

Jędrzej Haman
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw,

Artur Kaczorek
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0002-4752-4480

Jan Kisielewski
Faculty of Physics, University of Bialystok https://orcid.org/0000-0002-1082-4324

Bartosz Krupa
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0002-2958-2046

Antoni Moszyński

Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0003-0128-7384

Jędrzej M. Nowosielski
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0002-9627-3008

Maciej Radwan
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0001-5789-316X

Marcin Semeniuk
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0002-3136-7249

Urszula Tymoszuk
Division of Psychiatry, University College London https://orcid.org/0000-0002-0781-2752

Jakub Zieliński
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0001-8935-8137

Franciszek Rakowski
Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, https://orcid.org/0000-0001-6133-8900

Research Article

Keywords: epidemic dynamics, epidemiology, agent-based model, COVID-19

Posted Date: November 14th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2966996/v2

Forecasting SARS-CoV-2 epidemic dynamic in Poland with the pDyn agent-based model

Karol Niedzielewski[1*], Rafał P. Bartczuk[1,2], Natalia Bielczyk[3], Dominik Bogucki[1], Filip Dreger[1], Grzegorz Dudziuk[1], Łukasz Górski[1], Magdalena Gruziel-Słomka[1], Jędrzej Haman[1], Artur Kaczorek[1], Jan Kisielewski[1,4], Bartosz Krupa[1], Antoni Moszyński[1], Jędrzej M. Nowosielski[1], Maciej Radwan[1], Marcin Semeniuk[1], Urszula Tymoszuk[5], Jakub Zieliński[1], Franciszek Rakowski[1]

***For correspondence:** k.niedzielewski@icm.edu.pl (K)

[1]Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw, Warsaw, Poland; [2]Institute of Psychology, John Paul II Catholic University of Lublin, Lublin, Poland; [3]Ontology of Value, Nijmegen, Netherlands; [4]Faculty of Physics, University of Bialystok, Białystok, Poland; [5]Division of Psychiatry, University College London, London, Great Britain

Abstract We employ pDyn, an agent-based epidemiological model, to forecast the fourth wave of the SARS-CoV-2 epidemic, primarily driven by the Delta variant, in Polish society. The model captures spatiotemporal dynamics of the epidemic spread, predicting disease-related states based on pathogen properties and behavioral factors.

We assess pDyn's validity, encompassing pathogen variant succession, immunization level, and the proportion of vaccinated among confirmed cases. We evaluate its predictive capacity for pandemic dynamics, including wave peak timing, magnitude, and duration for confirmed cases, hospitalizations, ICU admissions, and deaths, nationally and regionally in Poland.

Validation involves comparing pDyn's estimates with real-world data (excluding data used for calibration) to evaluate whether pDyn accurately reproduced the epidemic dynamics up to the simulation time. To assess the accuracy of pDyn's predictions, we compared simulation results with real-world data acquired after the simulation date.

The findings affirm pDyn's accuracy in forecasting and enhancing our understanding of epidemic mechanisms.

Introduction

The first confirmed case of coronavirus disease 2019 (COVID-19) in Poland was identified on March 4, 2020, approximately a month behind Western Europe countries (***Ministerstwo Zdrowia, 2020***) (cf. 1(a)). On March 10, the World Health Organization declared the local transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in Poland (***Pinkas et al., 2020***). Within two days, the country recorded its first COVID-19-related fatality (***Duszyński et al., 2021***). As the epidemic spread, Poland's government declared an epidemic emergency, subsequently introducing mitigation policies (***Pinkas et al., 2020***) (see ***Figure 1***(d)). Critical pharmaceutical and non-pharmaceutical interventions (NPIs) implemented between March 4, 2020, and December 31, 2021, are detailed in ***Table 1*** in the ***Appendix 1***. These measures primarily included isolating infected individuals, quarantining contacts (with basic contact tracing), and SARS-CoV-2 testing. Public social distancing

measures, such as gathering bans and school and workplace closures, began in the second week of March 2020, culminating in a national lockdown on March 24, 2020. Further mandates for indoor and outdoor face coverings followed on March 30 and April 14, 2020. The national COVID-19 vaccination program commenced on December 27, 2020.

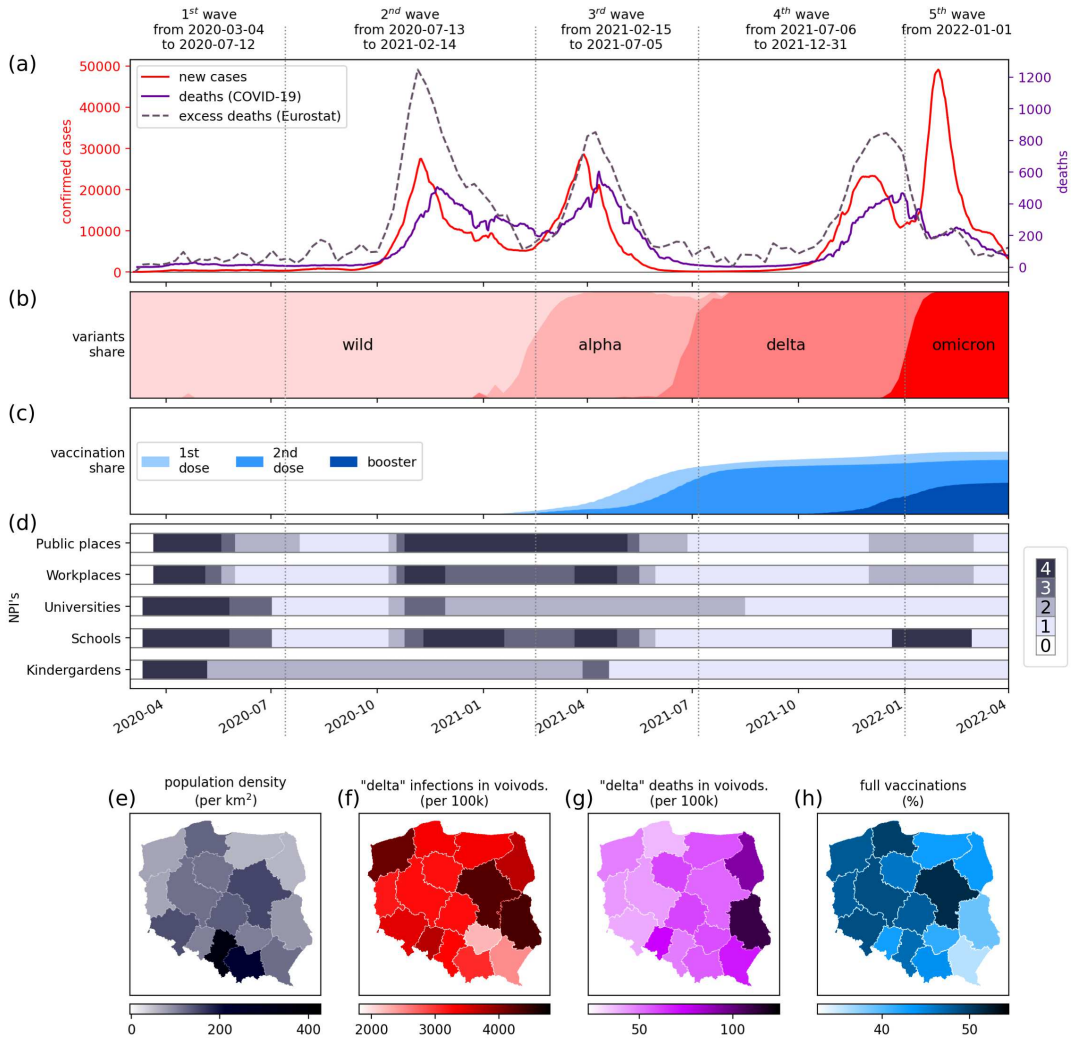


Figure 1. Timeline of SARS-CoV-2 epidemic in Poland. (a): The epidemic curve showing the progression of reported daily new confirmed cases in Poland (red), number of COVID-19-related deaths (purple), and excess mortality (dashed). (b): Proportions of dominating variants. (c): Full vaccination share. (d): Government mitigation measures by implementation areas and ranks of restrictive strength. (e): Map of inhabitants density in voivodships. (f): Map of reported cases during the Delta wave in voivodships. (g): Map of deaths during the Delta wave in voivodships. (h): Map of vaccinations per 100,000 inhabitants in voivodships up to May 2022.

Data sources: Daily cases & COVID-19-related deaths: Ministry of Health https://gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2. Vaccinations: . Excess mortality: Eurostat (***Eurostat, 2023b***). SARS-CoV-2 variants: GISAID study (***Khare et al., 2021***) . Mitigation measures: own elaboration based on governmental information please see ***Table 2*** in the ***Appendix 1*** .

The evolving nature of the epidemic, with factors such as new virus variants, seasonal transmission fluctuations, regional outbreaks, and the introduction of vaccinations, necessitated a dynamic approach to epidemic mitigation. This approach involved localized and reactive strategies. For in-

stance, a reactive policy was initiated on August 7, 2020, with school closures and remote work mandates triggered by defined case thresholds per 10,000 inhabitants in administrative units. Compliance with prevention measures, including face mask use, exhibited temporal and demographic variations. Older adults, urban areas, and different epidemic stages demonstrated varying levels of adherence (***Haischer et al., 2020***; ***Delussu et al., 2022***).

Moreover, ongoing research on SARS-CoV-2 pathogen properties, such as transmission modes, asymptomatic case infectivity, naturally induced immunity, its duration, and reinfection risks, added to the complexity of forecasting SARS-CoV-2 spread. Consequently, the demand for accurate forecasts, encompassing new infections, hospitalizations (general and intensive care units [ICUs] admissions), and COVID-19-related fatalities, intensified in response to the imperative of managing SARS-CoV-2 transmission.

Agent-based models (ABMs) have been a robust method for modeling infectious disease spread for over three decades (***Fox et al., 1971***; ***Elveback et al., 1976***). They offer a direct representation of dynamic social networks of agents and their heterogeneous interactions across georeferenced locations (***Dilaver and Gilbert, 2023***; ***Epstein, 1999***; ***Millington et al., 2012***). These models often rely on synthetic societies that mirror the demographic structure of specific territories (***Banks and Hooten, 2021***). They usually incorporate dynamic microsimulation methodology with elements of agent-based modeling. However, due to the convergence of these concepts, particularly as microsimulation becomes more and more intricate (***Railsback and Grimm, 2019***; ***Richiardi, 2014***; ***Vincenot, 2018***), we employ the term "agent-based model" as an umbrella term in this article. ABMs require input from a georeferenced network of setting where agents can operate and interact, like households, schools, workplaces, and public spaces, referred to as contexts. ABMs, as generative models, excel in replicating complex outbreak phenomena, accounting for regional disparities, demographic structure, behavioral responses, and parameter calibration at finer spatial scales. In contrast, data-based, phenomenological models, such as uniform mixing compartment models, lack implicit interactions among crucial factors like virus variants and social networks (***Silverman et al., 2021***).

The ABM developed by the Ferguson group (***Ferguson et al., 2005***) stands as a textbook ABM approach for modeling infectious disease processes. Designed initially for simulating influenza spread and assessing the effectiveness of targeted antiviral prophylaxis in Southeast Asia, this model classifies individuals into households with distinct generational layers. In 2020, it underwent adaptation to predict SARS-CoV-2 transmission dynamics by adjusting disease parameters to align with the virus's characteristics ***Ferguson et al.*** (***2020***). These forecasts informed the intermittent lockdown strategy in the UK, known as "The Hammer and the Dance" (***Pueyo, 2020***).

Agent-based models (ABMs) have proven effective in modeling and predicting epidemics. They function as virtual laboratories that enable the formalization and testing of epidemic dynamics (***Priesemann et al., 2021***). Unlike models that rely on general factors and aggregate variables, ABMs focus on modeling individual agents and their interactions, allowing for the development of agent-level theories, identification of fundamental principles and assumptions, and uncovering research gaps and inconsistencies in theoretical systems (***Dilaver and Gilbert, 2023***; ***Epstein, 1999***; ***Frias-Martinez et al., 2011***). Consequently, the prediction accuracy of ABMs depends on accurately representing elementary epidemic processes and supporting hypotheses regarding their impact on real-world data (***Dilaver and Gilbert, 2023***; ***Epstein, 1999***; ***Millington et al., 2012***). However, implementing complex epidemic processes and adhering to real-world rules come at the cost of numerous parameters and high computational expenses. Additionally, the calibration process poses a significant challenge, demanding substantial resources to achieve reliable calibration (***Millington et al., 2012***; ***Macal, 2016***; ***Epstein, 1999***).

Our initial model, known as pDyn (derived from "pandemics dynamics"), was developed in 2008 to depict influenza spread scenarios in Poland (***Rakowski et al., 2010a***,b), drawing inspiration from the Ferguson group model (***Ferguson et al., 2005***). In response to the COVID-19 pandemic, we adapted this simulation platform to meet the specific requirements of decision-makers based on

the pandemic's unique characteristics (***Niedzielewski et al., 2022***). The model can simulate and forecast various SARS-CoV-2 transmission scenarios. The pDyn model has received official endorsement from the government, alongside the MOCOS model (***Adamik et al., 2020***) and the Ministry of Health Department of Analysis and Strategy model, as one of the primary tools for providing scientific insights and epidemic forecasts to policymakers and medical advisory councils on a long-term basis.

In Poland, several ABM models have been developed. Compared to the MOCOS model (***Adamik et al., 2020***), pDyn distinguishes itself with a detailed and georeferenced structure of various contexts, while MOCOS incorporates advanced contact-tracing analytical methods. Other models, such as those developed as conceptual models (***Pałka et al., 2022***; ***Regulski et al., 2021***) offered valuable methodological insights but were primarily employed locally and did not transition into operational use.

During the initial year of the pandemic, the pDyn model stood out as one of the few robust models subjected to validation against real-world data. A systematic review of 126 SARS-CoV-2 ABMs highlighted that only 17% underwent validation against real-world data, 3% were compared with other models, and 2% underwent systematic testing (***Lorig et al., 2021***). Furthermore, pDyn has continuously undergone external validation with real-world data as part of the German and Polish COVID-19 Forecast Hub since November 2020 (***Bracher et al., 2021***, ***2022***). Both ABMs, pDyn and MOCOS, have demonstrated significant performance improvements in long-term case forecasting in Poland, thanks to their tailored approaches adapted to the specific circumstances of the country (***Bracher et al., 2022***).

As for other single-country ABM models across European states, numerous models are dedicated to Austria (***Bicher et al., 2018***, ***2023***), Germany (***Müller et al., 2021***; ***MONID - MOdeling Network for severe Infectious Diseases, 2023***), Spain (***Singh et al., 2022***; ***Merino et al., 2023***), France (***Hoertel et al., 2020***), UK (***Ferguson et al., 2020***), Italy (***Bouchnita and Jebrane, 2020***; ***Giacopelli, 2021***; ***Lombardo et al., 2022***; ***Fazio et al., 2022***) and Ireland (***Novakovic and Marshall, 2022***). However, these models are tailored to countries other than Poland (or their respective regions) and have not undergone validation within the European COVID-19 Forecast Hub (***Sherratt et al., 2023***). Therefore, comparing the performance and validity of these models with pDyn in a meaningful manner would be challenging, if not impossible. Nonetheless, considering the population size of European nations, pDyn ranks among the top 10 in terms of simulated populations.

This report utilizes the ABM pDyn to forecast the spatiotemporal dynamics of the COVID-19 epidemic in Poland. Our methodology encompasses disease transmission, disease progression, and epidemic course (see ***Figure 2***). Disease transmission considers multi-variant pathogens, partial immunity, and social contacts. The disease progression component includes a detailed representation of disease-related states, age-dependency, and dark figure estimation. Lastly, the epidemic course encompass changes in risk exposure due to NPIs or other shifts in behavior, vaccination policies, cross-immunity, and immunity waning. We validate the dynamics implemented in the model by inspecting their consistency with real-world data not used for calibration. Many of these features are model enhancements related to COVID-19 epidemics (indicated by asterisks $*$ in ***Figure 2***).

This investigation spans from the onset of the epidemic to the end of 2021, covering four SARS-CoV-2 waves in Poland. The first and second waves (March 4, 2020–July 12, 2020, and July 13, 2020–February 14, 2021) were driven by the wild-type virus variant, followed by the third wave with the Alpha variant (February 15, 2021–July 5, 2021) and the fourth wave with the Delta variant (July 6, 2021–December 31, 2021). During this period, Poland reported 4,106,914 SARS-CoV-2 cases, 96,967 COVID-19-related deaths, and 173,376 total excess deaths (***Ritchie et al., 2020***) (see ***Figure 1***(a) and (b)).

The forecast, formulated on October 28, 2021, using pDyn (***Niedzielewski et al., 2022***), targets the fourth (Delta) wave of the epidemic in Poland. This wave is noteworthy as it subsided spontaneously, without the imposition of restrictions or contact limitations, signifying the attainment of

Figure 2. Features of the current version of the pDyn model. The model encompasses three main classes of features: (1) *Disease transmission*, which incorporates airborne transmission dynamics, multi-variant pathogens, vaccine characteristics, partial immunity, and social contacts structure; (2) *Disease progression*, which models disease-related states, their durations and transition probabilities, age-dependency, and estimates of underreporting; and (3) *Epidemic course*, modeling changes in risk exposure, vaccination policies, cross-immunity, and immunity waning. Many of these components represent enhancements related to COVID-19 epidemics (indicated by asterisks [*]) compared to the original model version.

herd immunity. Subsequent waves in 2022 represented reinfections and conveyed reduced risks of severe illness and death due to decreased susceptibility to new variants. The forecast did not include the emerging Omicron wave in January 2022 due to a lack of information on this variant at that time. To be precise, Omicron variant was not introduced to the forecast of interest, formulated on October 28, 2021. Therefore, any comparison between the forecast and real-world data should only consider the period until December 31, 2021, as the Omicron variant emerged in early 2022.

This study aimed to achieve three specific objectives: (1) assessing the validity of the dynamics embedded in the pDyn model, (2) evaluating its capacity to predict the dynamics of disease-related states at the national level, and (3) gauging its ability to predict epidemic dynamics in Poland's highest administrative units (voivodships) using nationally reported data. We compared real-world data on SARS-CoV-2 variants, immunization dynamics, and the ratio of vaccinated individuals among confirmed cases with our model's estimates to assess its validity. Additionally, we compared simulation results with real-world data obtained after the simulation date to evaluate pDyn's predic-

tive accuracy. We also assessed regional forecasts made using nationally reported data, taking into account the synthetic society's reflection of geographical variations in the social-demographic structure of the Polish population.

As demonstrated, the generative, epidemiology-driven dynamic approach of pDyn achieved high predictive accuracy when modeling the spread of COVID-19 epidemics.

Results

Evaluation of the model validity

In this section, we assess the validity of the model's dynamics by comparing its outputs with real-world data pertaining to the dominant variant of concern, immunization levels, and the fraction of vaccinated detected cases.

Dominating variant of pathogen

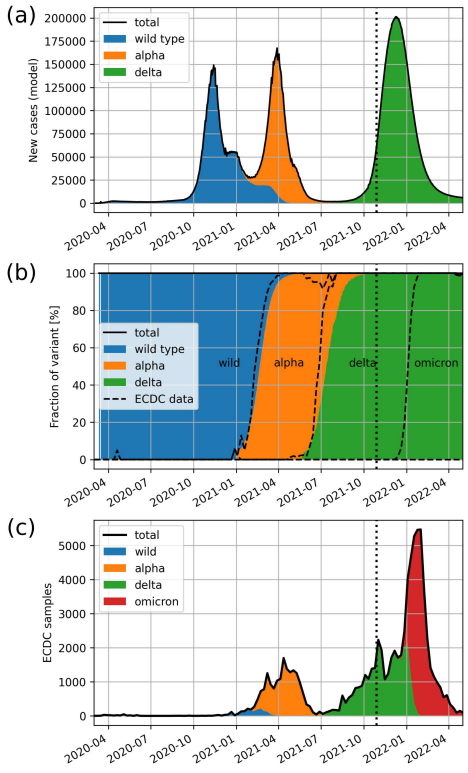


Figure 3. Comparison of the dynamics of a succession of SARS-CoV-2 variants obtained from the pDyn model (colors) with the dataset obtained from GISAID study (dashed lines). All the data is aggregated in weekly intervals. The vertical dotted line marks the simulation date.

Data source: GISAID study (***Khare et al., 2021***)

Before the simulation date, three predominant variants had been identified in Poland: the wild type, Alpha, and Delta. In ***Figure 3***(a), the distribution of these variants (wild type [blue], Alpha [orange], and Delta [green]) among infected agents is depicted. Panel (b) compares the model's variant succession dynamics with real-world data from GISAID (***Khare et al., 2021***). This assessment excludes the Omicron variant, which was not part of our October 2021 forecast.

Given that GISAID data relies on samples of varying sizes and considering potential biases in the data due to relatively small samples for Poland (as shown in ***Figure 3***(c)), we primarily compared the relative prevalence of variants, expressed as percentages. To validate our findings, we compared the timing of variant succession at the 25%, 50%, and 75% percentile thresholds.

The pDyn model reached 25% prevalence of the Alpha variant one week after the reference GISAID data, while it reached 50% and 75% prevalence two weeks after the reference GISAID data.

Regarding the Delta variant, our model reached 25% prevalence two weeks after the reference GISAID data, 50% prevalence three weeks after, and 75% prevalence four weeks after the reference GISAID data. This transition from the Alpha to Delta variant occurred during a period of relatively low newly detected cases, supporting the realism of our model's predictions.

Immunization level

The immunization dynamics during the epidemic originating from the pDyn model, categorized as disease-induced, vaccination-induced, and total immunization, are presented in ***Figure 4***. This

model output is systematically compared with data from the OBSER-CO nationwide seroprevalence study conducted by the National Institute for Public Health in Poland (***National Institute of Public Health, 2021***).

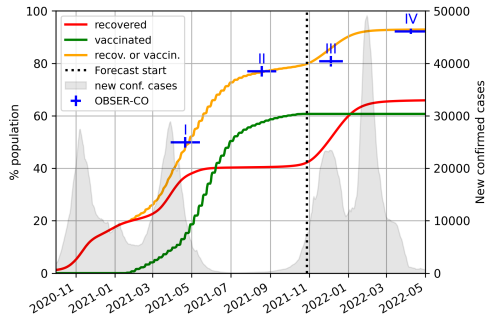


Figure 4. Comparison of immunization dynamics between the model output and the OBSER-CO study. The lines show the cumulative percentage of the agents (left axis) that are recovered (red), vaccinated (green), or recovered or vaccinated (yellow). The blue markers indicate the estimated fraction of the population with SARS-CoV-2-specific antibodies from four rounds of the OBSER-CO. The horizontal marker line denotes the duration of each round, the vertical one represents the 95% confidence interval of the estimate. A vertical dotted line indicates the simulation date. The gray shape represents the number of real-world confirmed cases (right axis).

Data source: OBSER-CO study (***National Institute of Public Health, 2021***)

Figure 4 reveals a close alignment between the cumulative sum of recovered and vaccinated individuals predicted by the model and the estimates derived from the seroprevalence study at all four study rounds. The percentage of the entire population represented by these estimates is as follows: 48.1 (model) vs. 49.9 (study, 95% CI [47.9; 51.9]) in April/May 2021, 76.8 (model) vs. 77.0 (study, 95% CI [75.0; 79.0]) in September 2021, 85.8 (model) vs. 80.8 (study, 95% CI [78.8; 82.8]) in December 2021, and 92.9 (model) vs. 92.2 (study, 95% CI [91.2; 93.2]). Notably, rounds I and II of the OBSER-CO study fell within the calibration stage of the simulation. In contrast, the model results for rounds III and IV are purely prognostic values.

It is important to acknowledge that the OBSER-CO study primarily focuses on seroprevalence, which relies on antibody levels in trial groups, differing somewhat from the indicator of the sum of recovered and vaccinated cases obtained from the model. Nevertheless, the significant alignment between the model's approximation of societal immunity and OBSER-CO data underscores the model's reliability in forecasting future epidemic waves despite these variations in indicators.

Fraction of vaccinated detected cases

The third validation involves assessing the fraction of vaccinated detected cases, which refers to the number of vaccinated individuals among all infected and detected individuals. The model adopted a vaccination strategy based on government data, which included the number of vaccinated agents at specific ages, times, and locations. However, the dynamics of the fraction of vaccinated detected cases emerged from the model and could be compared to real-world data (obtained under a non-disclosure agreement). The comparison between the model's outcomes and epidemiological data regarding the fraction of vaccinated detected cases is presented in ***Figure 5***.

Generally, the dynamics obtained from the pDyn model align closely with the epidemiological data. The mean absolute error from January 1, 2021, to October 28, 2021 (forecast

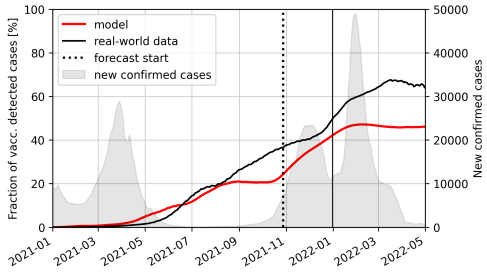


Figure 5. Comparison between the fraction of vaccinated detected cases generated from the pDyn (red line) and epidemiological data (black line). The vertical dotted line indicates the simulation date, and the solid vertical line — the end of the estimation period for the model-to-real-data fit indices. The gray shape in the background represents the number of real-data new cases.

date) equals 3.44%; from October 29, 2021, to December 31, 2021, equals 7.23%; and from January 1, 2022, equals 16.34%. The maximal error from January 1, 2021, to October 28, 2021, equals 14.01%; from October 29, 2021, to December 31, 2021, equals 12.20%; and from January 1, 2022, equals 21.69%. The larger maximal error before the forecast date may have been due to data variability when the number of cases was still low, but vaccination uptake had reached its saturation point (***Figure 4***, green line).

The quantitative indices used to validate the Delta wave forecast were estimated until December 31, 2021, when the Omicron wave officially began. Considering the whole period (from January 1, 2021, until May 1, 2022), the maximal error occurred on March 17, 2022, after the Delta domination period. Given that the Omicron variant was not considered in the forecast, the most considerable discrepancy between our simulation and real-world data was expected to occur after December 31, 2021.

Prediction of the epidemic dynamics during the Delta wave on the national level

In this section, we evaluate the accuracy of the pDyn forecast by comparing its results with real-world data for new confirmed cases, COVID-19-related deaths, hospitalized patients, and ICU patients published by the Ministry of Health. Visualizations of the forecast for the Delta-wave (***Figure 6***) and the entire epidemic (***Figure 6—figure Supplement 1***) are presented. To assess the model's performance, we conducted a comparative analysis with real-world data, emphasizing the accuracy of peak timing, magnitude, and duration, with summarized results in ***Table 1***.

Table 1. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Poland. All data is reported daily. Real-world numbers of reported deaths and excess deaths are compared to the same number of COVID-19-related deaths from the simulation.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	25770	2021-12-05	68
	Real-world	24120	2021-11-29	45
	Difference	1659	6	23
	Relative difference	6.84%		51.11%
Hospitalized	Simulation	41315	2021-12-12	66
	Real-world	23520	2021-12-10	61
	Difference	17795	2	5
	Relative difference	75.66%		8.20%
ICU patients	Simulation	5311	2021-12-21	66
	Real-world	2115	2021-12-14	67
	Difference	3196	7	-1
	Relative difference	151.11%		-1.49%
Reported deaths	Simulation	889	2021-12-21	68
	Real-world	443	2021-12-17	62
	Difference	446	4	6
	Relative difference	100.68%		9.68%
Excess deaths	Simulation	889	2021-12-21	68
	Real-world	845	2021-12-10	66
	Difference	44	11	2
	Relative difference	5.21%		3.03%

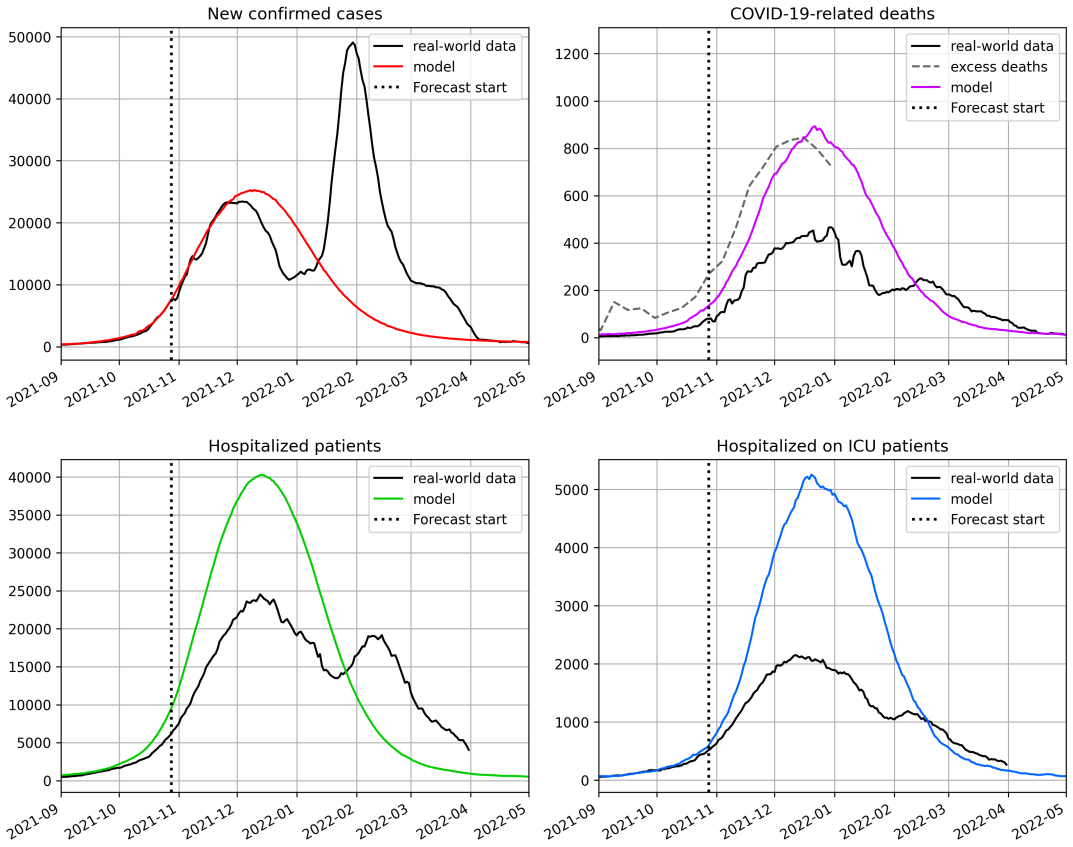


Figure 6. Comparison between the output generated from the model (colored lines) and the COVID-19 data from the Polish Ministry of Health (black) and Eurostat (dashed gray) for the Delta wave of the COVID-19 epidemics in Poland. Top left: new detected cases. Top right: deaths. Bottom left: hospitalized patients. Bottom right: ICU patients. The vertical dotted line marks the simulation date.

Data source: Eurostat (***Eurostat, 2023a***) **Figure 6—figure supplement 1.** Comparison between the pDyn model-generated output and the epidemiological data for the entire course of the COVID-19 epidemics in Poland.

The forecasted peak values tended to be overestimated, with the most accurate prediction for new cases (a relative difference of ~7%) compared to other metrics. As shown in ***Figure 6***, the predicted number of hospitalized patients, ICU patients, and COVID-19-related deaths exceeded the official data provided by the Ministry of Health: hospitalizations by approximately 76%, ICU admissions by around 151%, and COVID-19-related deaths by roughly 101%. Concerning the timing of peaks, our predictions were most accurate for hospitalized patients (with a 2-day difference), followed by reported deaths (4 days), new confirmed cases (6 days), and ICU patients (7 days). The forecasted wave length, as measured by the Full-Width Half-Maximum (FWHM), was the most accurate for ICU patients (approximately 1%) and hospitalized patients (around 8%), followed by reported deaths (approximately 10%) and new cases (about 51%). Notably, the relative difference in FWHM between the modeled and observed new confirmed cases was likely due to the holiday period in late December 2021, leading to a lower testing rate and detection ratio than before the holidays. Given that our model assumes a constant detection ratio, the real-life decrease in reporting likely contributed to the observed discrepancy in the confirmed cases' wave length.

It is important to emphasize that the model's primary aim was not to predict the actual number of hospitalized and ICU patients but to estimate their expected numbers. Therefore, this distinction should be kept in mind when interpreting the results, as it may explain the significant differ-

ences between the model's peak value predictions for hospital and ICU beds and the real-world data. Nonetheless, the predictions regarding the peak timing of hospitalized and ICU patients demonstrated that our forecast accurately captured the dynamics of the Delta wave. We relied on occupied beds, rather than hospital admissions, to assess hospitalizations since the Ministry of Health only provided data on occupied beds. A similar limitation affected our assessment of ICU hospitalizations.

Moreover, we found that excess deaths were a more reliable parameter than officially registered COVID-19 deaths. Consequently, we present the forecast of COVID-19-related deaths in comparison to estimates of excess deaths as defined by Eurostat ***Eurostat*** (***2023b***): "Excess mortality is the rate of additional deaths in a month compared to the average number of deaths in the same month over a baseline period (2016-2019)." A positive value indicates an increase in deaths compared to the baseline, while a negative value signifies fewer deaths compared to the baseline period. For a more accurate quantitative comparison, we provided data in weekly resolution computed using Eurostat weekly excess deaths data. The quantitative estimates of peak timing, peak value, and wave length are presented in the lower panel of ***Table 1***. Notably, the predicted number of deaths more closely followed excess deaths then reported deaths in terms of the peak value (a relative difference of approximately 5% vs. approximately 101%) and wave length (around 3% vs. approximately 10%).

Prediction of the epidemic dynamics during Delta wave on regional level

Here, we demonstrate the model's capability to forecast epidemic dynamics of disease-related states in voivodships while using national-level epidemic data for calibration. The regional trajectories of pDyn outputs diverged due to synthetic society's spatial structure, vaccination process, regional variation in weight multipliers (accounting for differences in NPIs implemented before the Delta wave), and the locations of initial infections for each introduced variant in the simulation.

For a comprehensive comparison between the data generated by the model and real-world data throughout the entire course of the epidemic in voivodships (comprising the total number of detected cases, COVID-19-related hospitalizations, COVID-19-related ICU occupation, and COVID-19-related deaths), please refer to ***Appendix 2***. Detailed quantitative comparisons of peak timing, peak value, and wave length are also included in Tables in ***Appendix 2***. In this context, ***Figure 7*** primarily presents the summary statistics of model accuracy at the regional level.

The top panel of ***Figure 7*** illustrates the distributions of peak values, peak dates, and FWHM values in voivodships, obtained from both the model (upward distributions) and real-life data (downward distributions). For clarity, the bottom panel shows these data as distributions of absolute differences (for peak date) or relative differences (for peak value and FWHM) between the model and real-life data.

The medians of the difference distributions, indicated by vertical black lines in the bottom panel of ***Figure 7***, broadly align with the differences reported at the national level. Notably, there are a few outliers in the graphs depicting relative differences in peak value and FWHM for newly detected cases, hospitalized patients, and deaths.

Upon inspecting the difference distributions (bottom panel in ***Figure 7***), particularly the relative peak value difference (left plot) concerning detected cases, occupied beds, and deaths, one can observe a clear outlier in each plot, which corresponds to Podkarpackie voivodeship. The substantial relative differences observed across voivodeships may be partially attributed to regional behavioral factors, such as varying levels of willingness to undergo COVID-19 testing or seek hospital treatment for COVID-19, compared to other regions in the country.

On average, the pDyn model demonstrates convergence with real-world data and predicts the number of newly detected cases at the individual voivodeship level with lower accuracy than the predictions made at the national level.

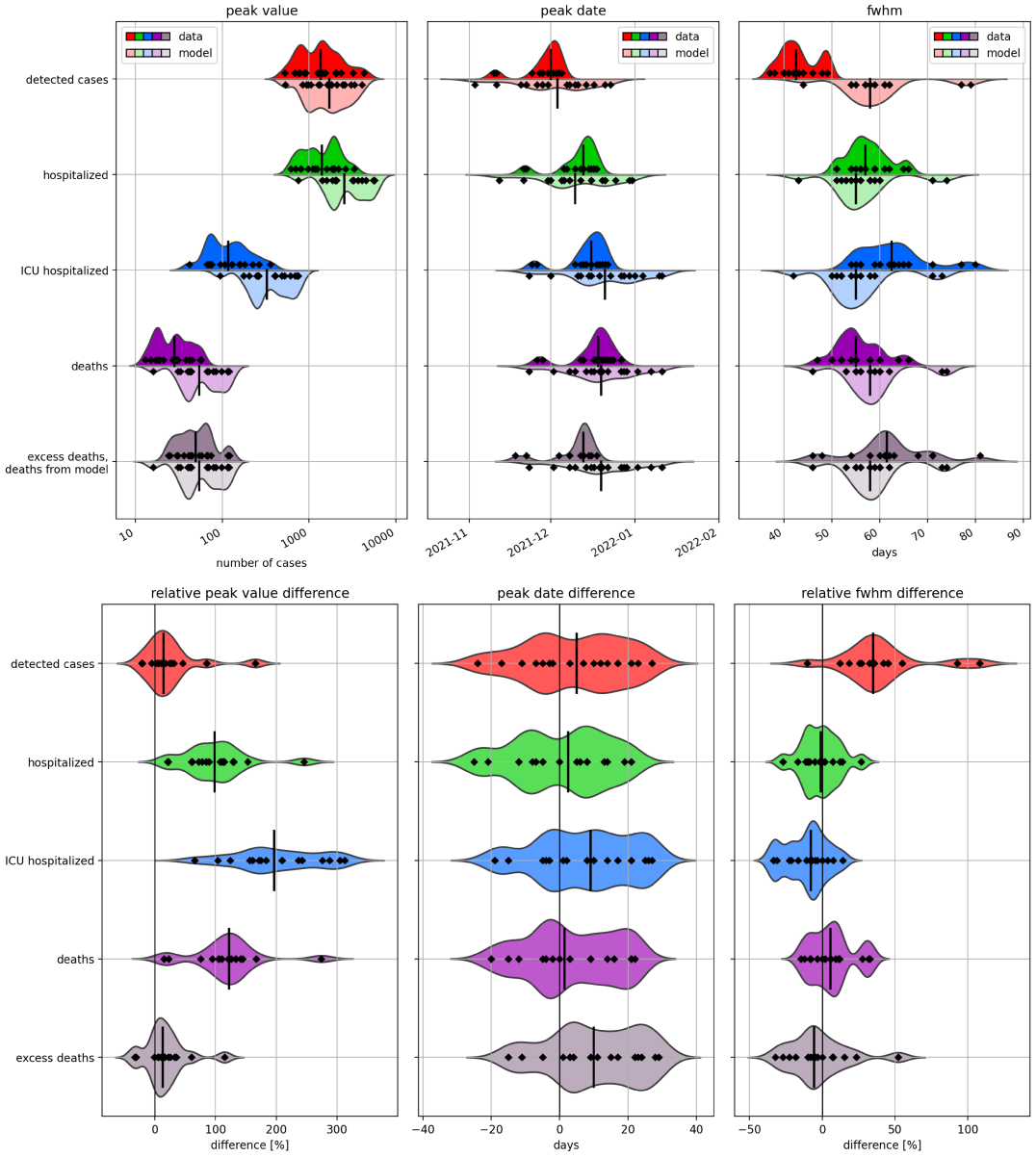


Figure 7. Summary results for the model forecast on a regional level. Each data point refers to one voivodeship. Upper panels present smoothed distributions of peak value, peak date, and FWHM separately for the Ministry of Health data (above horizontal reference lines) and the model forecast data (below the reference lines) for daily new detected cases, occupied hospital beds, occupied ICU beds, and COVID-19 deaths (compared also to excess deaths based on the Eurostat data (***Eurostat, 2023a***)). Lower panels present smoothed distributions of relative peak value difference, peak date difference, and relative FWHM difference between the model forecast and the official data. The data points are accompanied by median values (vertical black segments).

Discussion

Mathematical epidemic models play a crucial role in understanding and informing effective mitigation strategies for disease outbreaks (***Brauer, 2008***; ***James et al., 2021***; ***Marshall, 2017***; ***Ferguson et al., 2020***, ***2005***). This manuscript focuses on validating the epidemic dynamics and assessing the forecasting accuracy of pDyn, an agent-based model specifically designed to capture and predict the dynamics of COVID-19 in Poland.

The pDyn possesses several key strengths for modeling epidemic dynamics. Firstly, it excels in

capturing intricate social networks and contact patterns among individuals, factors with a substantial impact on disease transmission. Consequently, it provides valuable insights into the individual-based and network-based mechanisms governing epidemic spread. Secondly, the model's versatility allows it to simulate epidemics at different spatial scales, thanks to its incorporation of geospatial data such as population demographics and transportation networks. These features enable the simulation of various intervention strategies, such as quarantine and social distancing, and their impact on epidemic spread. Additionally, pDyn models multiple pathogen variants and cross-immunity, shedding light on the role of variant and vaccine diversity in epidemic dynamics. It also integrates a model for immunity acquisition and waning, enabling the simulation of the effects of vaccination and natural infection.

In the study presented in this manuscript, we aimed to achieve three objectives:

1. We first assessed the model's validity in simulating the dynamics of pathogen variants succession, immunization processes, and the proportion of vaccinated individuals among confirmed cases.
2. We then assessed the model's predictive capabilities by examining its performance in forecasting the dynamics of confirmed cases, hospitalizations, ICU admissions, and deaths during the epidemic wave, focusing on critical metrics like peak timing, peak magnitude, and wave duration.
3. Lastly, we explored the utility of pDyn in forecasting disease-related dynamics within Poland's highest administrative units using national-level data. This was made possible through the use of a virtual population representing the social and demographic structure of Poland.

We summarize our findings and discuss them below.

The first aspect we examined to validate the model was the progression of variants. This dynamic depends on various factors, including variants' properties like cross-immunity and infectivity, as well as the spatiotemporal distribution of initial infections for each variant. It is important to note that pDyn considers cross-immunity, seasonal fluctuations, and regulatory changes based on official data but does not incorporate emerging behavioral changes that could influence the model.

Our study revealed that the Delta variant reached prevalence milestones of 25%, 50%, and 75% two, three, and four weeks later, respectively, compared to the GISAID genomic data. Our validation aligns with prior research such as ***Eales et al.*** (***2022***) and ***Dong et al.*** (***2022***), which assessed variant succession at the 50% prevalence point. These studies reported prediction errors within one to two weeks, indicating a similar level of accuracy to pDyn, albeit slightly better. However, the superior performance of other models compared to pDyn may be partially attributed to their calibration and validation using the same datasets, whereas pDyn underwent calibration and validation using separate datasets. Two other studies solely offered visual comparisons (***Coutinho et al., 2021***; ***Campbell et al., 2021***). It should also be mentioned that potential selection bias in GISAID estimates for Poland could contribute to observed differences. Nevertheless, it can be concluded that our findings demonstrate that pDyn effectively replicates variant succession.

Next, we compared our modeling outcomes with the OBSER-CO seroprevalence survey conducted by the National Institute of Public Health. Our model's cumulative count of recovered and vaccinated individuals closely aligned with the seroprevalence study's estimates at all four study rounds. However, the estimate for December 2021 was slightly elevated, falling three percentage points outside the 95% confidence interval. While models akin to ours have been calibrated against seroprevalence data, they have not, to our knowledge, undergone validation against such data (***Kemp et al., 2021***; ***Jentsch et al., 2021***).

Some data issues can contribute to the uncertainty of our validation. OBSER-CO results estimates might be influenced by instability in detecting cases related to the testing during rising and falling epidemic waves (***Rippinger et al., 2021***). Furthermore, the study rounds extended over time,

with unevenly distributed testing within each round, while seroprevalence was estimated at specific central time points within each round interval, which could have influenced estimate accuracy. Moreover, the sum of recovered and vaccinated cases derived from the model does not align perfectly with OBSER-CO seroprevalence statistics based on antibody levels in the trial groups. Despite these reservations, pDyn's representation of immunity in society closely mirrors empirical OBSER-CO studies. Importantly, pDyn is, to our knowledge, the first model to faithfully reflect empirical seroprevalence, rendering it a valuable tool for exploring epidemic dynamics.

The proportion of vaccinated detected cases, considered a simulation variable, was compared with surveillance data from the Ministry of Health, using mean absolute error (MAE) as the validation metric. The MAE was smaller for the period before the forecast date (from January 1, 2021, to October 28, 2021) than for the forecast period (from October 29, 2021, to December 31, 2021,), reflecting the inherent uncertainty in predictions. Notably, the maximum absolute error (in the period from January 1, 2021, to December 31, 2021) occurred on October 18, 2021, reaching 14.01 percentage points. This peak coincided with low case numbers and high vaccination coverage, contributing to the observed variation. It is noteworthy that while the proportion of fully vaccinated individuals among detected cases has been used to assess vaccine effectiveness in empirical studies (***Arashiro et al., 2022***), it has not been commonly employed in epidemic modeling validation.

In summary, the pDyn model generated dynamics that generally aligned with epidemiological data, affirming the validity of the model's dynamics of variant succession, immunization, and the emergence of vaccinated individuals among confirmed cases.

Our proposed approach to handling uncertainty in generative models, like pDyn, offers added value to the epidemiological modeling domain. ABMs often involve numerous parameters requiring calibration, and the available data are insufficient for calibrating each parameter individually. In such cases, part of the model validation process may involve comparing variables that are not direct model outputs but can be derived from the model and compared to existing data before making forecasts—examples include the dynamics of pathogen variant succession, immunization, and the emergence of vaccinated individuals among confirmed cases. This approach aids in testing the validity of processes implemented in the model.

In the second phase of our study, we assessed the forecast accuracy for the Delta variant wave at the national level. This assessment encompassed new cases, hospitalizations, ICU admissions, and COVID-19-related deaths, focusing on peak value, peak date, and wave length.

The number of new COVID-19 cases, our primary output and reference for model calibration, provided precise forecasts for the peak value, with only a slight deviation of 6.84%. However, other aspects of the forecast exhibited overestimations. Notably, the predicted peak timing experienced a delay of six days, and the prediction of wave length (measured by FWHM) needed to be more accurate for new cases. This discrepancy can be attributed, in part, to the constant detection ratio implemented in our model. However, during the holiday season in December 2021, testing rates and detection ratios likely decreased, resulting in fewer confirmed cases and contributing to lower forecast accuracy.

The model predicted a significantly higher number of COVID-19-related general and ICU hospitalizations than the Ministry of Health reported, with 76% more hospitalizations and 151% more ICU patients. When building the model, we focused on required instead of occupied beds, assuming that even if some patients needing hospitalization stayed home, the health system should be prepared. This assumption is a primary source of the discrepancy between the model and data. Additionally, the model assumes constant hospitalization durations of 10 days for general and seven days for ICU admissions instead of using distributions, and poor data quality related to hospitalization durations adds uncertainty to estimations.

Our analysis revealed dynamically changing case-to-hospitalization and case-to-death ratios throughout the waves. These changes can be attributed to social reluctance towards testing and hospitalization due to difficult hospital conditions. Factors such as lack of family contact, admission challenges, and long queues at testing sites could contribute to this hesitance (***Kołodziej and Pecka,***

2021; *Grove et al., 2023*; *Rewerska-Juśko and Rejdak, 2022*; *Tran et al., 2020*; *Wong et al., 2022*; *Zheng et al., 2021*). However, the model did not account for psychological and healthcare system overload behavioral effects, which also affected the accuracy of our predictions.

Despite these caveats, peak time forecasts for hospitalizations were the most accurate among predicted disease-related states (2 days delay), and the relative wave length difference was best for ICU patients (-1.5%). In summary, the discrepancy between hospitalization and ICU patient data and model results refers to wave magnitude rather than peak or length. This result emphasizes the need to consider the forecast objectives and factors influencing access and utilization of health services when interpreting modeling outcomes, along with data collection challenges, to improve hospitalization nowcasting (*Wu et al., 2021*; *Wolffram et al., 2023*).

In this study, pDyn projected more COVID-19-related deaths than officially reported, aligning better with excess deaths, which capture undetected infections. For instance, *Walkowiak and Walkowiak* (*2022*) found that combined COVID-19-related deaths accounted for 95% of excess deaths among Polish adults over 40. Death forecasts closely matched excess deaths in peak value (5% relative difference) and wave length (3% relative difference). However, the peak date for deaths was the least accurate among all forecasted states, with an 11-day delay. The alignment of death modeling results with actual data is notably influenced by data collection issues, primarily attributing deaths to COVID-19, which is less reliable in Poland than other epidemiological data collected during the pandemic. The model's alignment with excess mortality data in our study supports COVID-19's substantial contribution to overall excess mortality during the pandemic (*Msemburi et al., 2023*; *Wang et al., 2022*; *Woolf et al., 2020*), particularly in Poland.

The pDyn model, grounded in a synthetic society reflecting regional socio-demographic data, explicitly considers regional variations in vaccination, pre-Delta wave NPIs, and the initial regional spread of wild-type infections in Poland. However, its calibration relies on national epidemic data. Our study aimed to evaluate the precision of regional forecasts generated by this model on the voivodeships level.

Our findings indicate that, on average, regional forecasts align with national-level ones, with median differences resembling those at the national level. However, prediction quality varies among voivodeships, with Podkarpackie voivodeship emerging as an outlier regarding relative peak value differences for detected cases, occupied hospital beds, and deaths. These substantial differences likely stem from localized variations in social responses to the epidemic and restrictions, underscoring the need to consider regional social attitudes for better regional forecasting.

To improve regional forecasting, incorporating agent features related to behavioral traits, such as trust in medicine and willingness to adhere to NPIs and vaccination, is advisable. Additionally, continual monitoring of local conditions and model adaptation to regional specifics enables more accurate predictions. Local models with adaptable parameterization, focusing on the short or medium-term, generally outperform global and long-term models (*Bracher et al., 2022*).

Despite the promising results, our study has limitations common to complex ABMs like pDyn, including the challenge of calibrating numerous parameters with limited data. This parameter calibration issue significantly contributes to forecast uncertainty. We attempted to address this by validating the model's parameterization by comparing model dynamics with real-world data, but these challenges persist, introducing inherent uncertainty into our forecasts.

Like all epidemiological models, pDyn encounters challenges in predicting unpredictable phenomena that can arise during an epidemic, such as pathogen mutations or shifts in social contact patterns, which can substantially influence the epidemic's trajectory. Our model does not include long-term predictions of pathogen evolution or the modeling of socio-behavioral dynamics. Instead, parameters related to these aspects are introduced post-hoc, often with delays, adding to the overall uncertainty of the model's predictions.

To enhance forecast accuracy, developing new features in the future may be necessary. Currently, the model assumes constant durations for health-related states (e.g., hospitalization), while using parameter distributions could improve realism. Simplified modeling of transportation and

commuting could be expanded for better representation of local and long-distance transmission. Agent behavior could be refined by introducing behavioral attributes, and contact change calibration would benefit from using external data, like mobility. However, these extensions increase the number of parameters to calibrate, computational complexity and load, as well as introduce inherent uncertainties (e.g., mobility change only proxies contact pattern change).

Data availability remains a fundamental limitation of the pDyn model. Several crucial datasets were unavailable at the time of our forecasting, including contact tracing data, the influx of new cases, the number of households in quarantine, and estimated transmissions between household members. Also, much of the data available during the data epidemic was biased and would require modeling (like nowcasting of hospitalization data). In particular, there needed to be more effort in obtaining current and reliable data quickly. These issues underscore the importance of robust and timely epidemic surveillance systems for mathematical modeling of epidemics.

Nonetheless, despite limited data availability, the pDyn model provided valuable insights into epidemic processes and demonstrated remarkable forecasting efficiency. It can aid in understanding epidemic mechanisms and inform epidemic policy design by enabling the comparison of multiple scenarios.

In summary, our study highlights the pDyn model's robust capabilities and potential for providing reliable and insightful forecasts across various aspects of the COVID-19 pandemic. Key findings can be summarized as follows:

1. *Model Validation:* The generative ABM pDyn employs intricate internal states to incorporate extensive data, allowing for the representation of mechanisms beyond the scope of phenomenological models. We validated the model's accuracy in simulating pathogen variant succession, immunization processes, and the proportion of vaccinated individuals among confirmed cases, revealing close alignment with real-world data. Additionally, we introduced an innovative approach to address uncertainty in generative models. This approach involves comparing model-generated variables, which were not targeted initially as outputs, with real-world data, thereby enhancing our ability to analyze patterns.
2. *Predictive Capabilities:* The meticulous generative description of epidemic spread in pDyn results in impressive predictive performance, encompassing new cases, hospitalizations, ICU admissions, and deaths. Evaluations within the German and Polish COVID-19 Forecast Hub and the European COVID-19 Forecast Hub confirmed these capabilities. In our assessment of predictive capabilities, we focused on peak timing, peak magnitude, and wave duration for confirmed cases, hospitalizations, ICU admissions, and deaths. While peak values were often overestimated, the model consistently captured the dynamics of the Delta wave. Our findings underscore the importance of aligning forecasting interpretation with the challenges related to data collection during epidemics. This highlights the role of informed nowcasting, particularly for data related to infection-related hospitalizations and deaths.
3. *Regional Forecasting:* pDyn enables detailed epidemic simulations at both national and regional levels, providing a granular perspective on disease dynamics. However, forecasting at the regional level using national data has inherent limitations. Our examination of regional forecasting within Poland's administrative units revealed alignment with real-world data, although variations were observed, likely influenced by regional behavioral factors.

In conclusion, the pDyn model possesses numerous strengths, including its capacity to model complex social networks, simulate epidemics across different spatial scales, and account for pathogen variants and immunity dynamics. Our comprehensive evaluation underscores its reliability in modeling COVID-19 dynamics in Poland, providing valuable insights for informing public health decision-making and mitigation strategies.

Finally, we propose recommendations for epidemiological ABMs:

- *Extend Validation:* ABMs should regularly validate their models by comparing internal variables with empirical data. This approach facilitates the validation of emergent epidemiolog-

ical dynamics without the need for individual parameter validation, especially in situations where parameter validation is challenging. Additionally, conducting step-by-step validation for specific phenomena, such as reinfections and vaccine efficacy, can provide a deeper understanding of the model's characteristics and increase confidence in the accuracy and robustness of its results.

- *Monitor Local Changes:* Monitoring local changes in epidemics, including the presence of variants of concern and shifts in seroprevalence, along with behavioral effects of mitigation strategies like vaccination campaigns, lockdowns, and testing, is essential. This practice allows for the customization of models and parameters to specific country or regional situations, leading to improved short and medium-term forecasting accuracy.
- *Enhance Monitoring Systems:* There should be a concerted effort to enhance monitoring systems in two critical dimensions — data quality and data coverage. Institutions responsible for data collection and monitoring should gain a deep understanding of the empirical data requirements for complex models like pDyn. Leveraging the fastest and most accessible data streams can significantly inform and improve modeling efforts.

These recommendations aim to strengthen the reliability and effectiveness of epidemiological ABMs, ultimately aiding in better preparedness and decision-making during disease outbreaks.

Materials and methods

The pDyn model

Our research utilizes pDyn, the detailed epidemiological ABM developed at the Interdisciplinary Center for Mathematical and Computational Modelling at the University of Warsaw, Poland (ICM) (***Niedzielew et al., 2022***). The simulator was optimized for High-Performance Computing environment and runs in the ICM supercomputing facility.

The simulator originated as the influenza epidemic model (***Rakowski et al., 2010a***) with following features implemented: airborne transmission, pathogen characteristics (i.e. transmissibility), self-isolation, social contacts settings (i.e. households, workplaces, schools, universities, public places, long distance travels), SIR states. Subsequently, during the COVID-19 pandemic, it has been expanded with features tailored to represent characteristics of the SARS-CoV-2 infection, to facilitate the Polish government's infection prevention and control the decision-making process. The following new components have been implemented: partial immunity, variants of pathogen/vaccines, quarantine, partial school closure (i.e. age dependent), reactive NPIs, regional NPIs, changing contact rates, vaccination, immunity waning, cross-immunity, dark figure, times and transition probability table (i.e. of the disease-related states), age-dependency of time and transition to disease states, new social contact settings (i.e. kindergardens), new disease-related states (i.e. asymptomatic, symptomatic, hospitalized pre-ICU, at ICU, not at ICU).

To better illustrate the pDyn's scale and complexity, we present a mind map in the ***Figure 2*** that organizes the model elements in a transparent, modular way. It explicitly depicts the version of the model used in the study. Functions developed by adapting the original version of the simulator to the COVID-19 are marked in the figure by asterisk (*). The detailed description of the pDyn model following the Overview, the Design concepts, and the Details protocol (ODD, (***Grimm et al., 2020***)) is publicly available (***Niedzielewski et al., 2022***).

The overall *purpose* of the pDyn model is to describe and explain the spatial and temporal dynamics of SARS-CoV-2 spread across Polish society. The model predicts the dynamics of the number and locations of disease-related states of agents in response to specific changes in the properties of the pathogen and the social structure and behaviour.

Two types of *entities* are included in the model: agents and contexts. Agents represent members of the society. Contexts capture interactions between agents; they represent locations at which the agents come in contact, such as households, workplaces, kindergartens, schools, universities or public places. Their geo-localized representations are included in the synthetic society as

model input (***Rakowski et al., 2010b***). The synthetic society is based on data provided by the Statistics Poland (***Statistics Poland, 2019***) and reflects the state at the beginning of 2019. The *spatial resolution* of the contexts is a grid of 1×1 km^2 (for Poland, it requires 800×800 grid cells). Additionally, pDyn models the mobility of agents via random long-distance travels (i.e. when an agent leaves its household for more than a day). Each agent is assigned to one or more contexts (household at least) that it visits daily.

Both agents and contexts are characterized by *state variables*. The agent's state variables are as follows: age, list of contexts to which it is assigned (including primary household), disease-related state, presence of symptoms, being on quarantine, travel status, transmission location, and history of immunization events. The context's state variables are as follows: spatial coordinates of a given context, transmission rate in this context, the number of agents in this context, the number of symptomatic infectious agents in this context, and the number of non-symptomatic infectious agents in this context. The *time resolution* in the simulation is one day.

The most important *process* of the model is airborne transmission.

For a given susceptible agent, for a given day of the simulation, and a given variant of the virus, the probability of becoming infected by that variant on the following day is computed based on three factors: (1) the infectivity parameter specific to the variant, (2) the infectivity of the contexts visited during the day which we define as the fraction of infectious agents infected with the considered variant in that context, (3) the weights of the daily visited contexts which represent the contact rate of the agents in that context.

Thus, the probability of each susceptible agent getting infected on the next day of the simulation is a function of the *disease-related states* of all agents with whom it has been in contact in contexts during the current day (***Niedzielewski et al., 2022***). Immediately after the recovery or after taking a vaccine, the agent is immune to the infection variant of the pathogen, but the level of immunity wanes over time. The level of immunity calculated on a given day of the simulation modifies the probability of infection with the variant. In addition, recovery from infection with a particular variant of the virus generates a certain level of cross-immunity to other variants. Furthermore, the context weights are adjusted using *multipliers* in time to represent the changes in the contact rates (i.e., the number of contacts made divided by the number of contact opportunities) due to behavioural reactions to the epidemic, both spontaneous or in response to the control measures.

The model of possible *disease-related states* in pDyn expands the SEIR (Susceptible, Exposed, Infected, Recovered) compartmental model (***Li and Muldowney, 1995***). An agent can find itself in one of the following disease-related states: susceptible, latent, asymptomatic, symptomatic, hospitalized outside the ICU, hospitalized before ICU, hospitalized at ICU, dead, or recovered state. These disease-related states form an ordered graph that defines possible courses of infection (***Figure 8***). At each branching, probability parameters have been introduced to control the likelihood of the specific transitions between states (specific to the pathogen variant).

In addition, the duration of each state is defined. The transition probabilities and the duration of states depend on agent's age. It is assumed that both the asymptomatic and symptomatic states are infectious and that infectivity in the symptomatic state is higher than in the asymptomatic one (***Sayampanathan et al., 2021***; ***Han et al., 2020***; ***Zhao et al., 2020***). On the other hand, there is a possibility for an agent in a symptomatic state to undergo self-isolation or quarantine, meaning the agent withdraws from all contexts except for the household. The probability and duration parameters were selected based on several studies (***Gold et al., 2020***; ***Carrillo-Vega et al., 2020***; ***Petrilli et al., 2020***; ***Ko et al., 2021***; ***Twohig et al., 2022***) and their values implemented in the present simulation are presented in ***Appendix 3***.

In the pDyn, the number of infected agents includes both detected and undetected cases. Undetected cases impact various aspects of pandemic dynamic such as the true disease spread, the number of immunised individuals, numbers of hospitalized cases and deaths. The model introduces a *dark figure* representing the number of undetected cases, generating outputs for both real

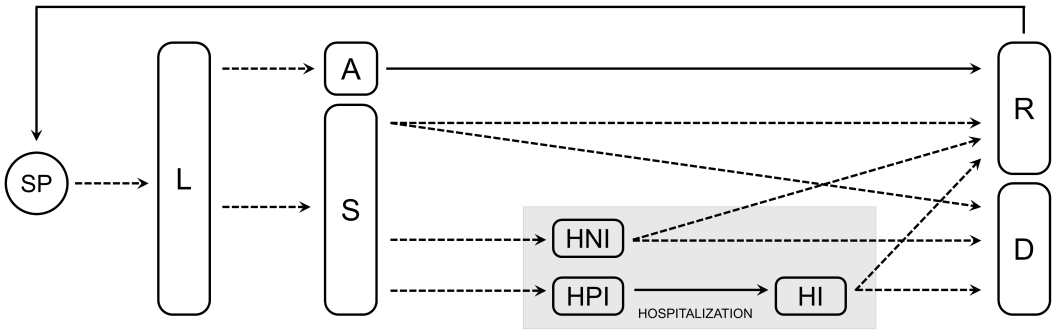


Figure 8. The possible paths through disease-related states in the pDyn model. The state abbreviations stand for SP — susceptible, L — latent, A — asymptomatic, S — symptomatic, HPI — hospitalized, pre-ICU, HI — hospitalized, at ICU, HNI — hospitalized, not at ICU, D — dead, R — recovered. In addition, there are three surefire paths (with transition probability equal to 1; marked with solid arrows) regardless of the agent's age.

cases (detected and undetected) and detected cases. The dark figure changes over time and is estimated by considering factors such as the ratio of non-symptomatic to symptomatic cases, testing strategies, test types, test numbers, contact tracing, public trust, and seroprevalence screening studies (***National Institute of Public Health, 2023***).

The pDyn model simulate vaccination programs, considering factors like geographical distribution, agent's age, and the number of vaccines administered. However, the presented simulation is agnostic to the type of vaccination, treating boost vaccinations the same as first doses, and not differentiating between various vaccines. The model offers fine-tuned control of vaccination, allowing for region-specific and age-based vaccination strategies with limited supply considerations based on data provided by Polish government under the license defined in a Non-Disclosure Agreement.

The pDyn explicitly addresses the cross-immunity phenomenon. The model assumes that the agent is immune to the infection variant immediately after recovery or after taking a vaccine, albeit the immunity level is waning over time. The decline in the immunity level is described in the function of elapsed time since recovery and can take values between 0 and 1 (***Figure 1*** in ***Appendix 3***). The immunity level of an agent computed on a given day of the simulation modifies the probability of infection with the variant subject to immunity. Moreover, we model the phenomenon of cross-immunity by assuming that recovery from an infection with a specific virus variant generates some immunity level to other variants. The parameters related to (cross-)immunity were estimated from (***Scobie et al., 2021***) and presented in ***Appendix 3***.

The pDyn allows to model risk exposure changes, whether seasonal (e.g. school closure during holidays) or behavioural (e.g. in response to NPIs, e.g., online schooling), by switching off or tuning contexts, using context weight *multipliers*. To our best knowledge, no systematic studies of contact rates changes were carried out during the COVID-19 epidemic in Poland. Instead, the models use intermediate (e.g., estimates based on measurements of the use of mobile networks) or partial (e.g., social mixing surveys) measures. In pDyn, the initial contacting rates were adopted from original influenza model (***Rakowski et al., 2010a***). Changes in contact rates during the outbreak and subsequent restrictions were implemented through multipliers.

In order to model changes in the contact rate for a particular context, we utilized the calibration experiments method, except for educational units, for which these multipliers were estimated based on the proportion of pupils attending them. Multipliers for the households, workplaces, and public places were adjusted with an assessment of the change in contact rates (based on changes in the number of people and their compliance with social distancing measures in a given context). The calibration experiments were executed in the following way: first we established the optimistic and pessimistic contact rate scenarios by assessing the minimum and maximum values of multipliers (such as low vs. high face mask use compliance). For example, on March 12, 2020, the mandate

of remote work and social distancing at the workplace was introduced, therefore we reduced the value of the workplace context multiplier from 1 to 0.5 in the optimistic scenario and to 0.8 in the pessimistic scenario. Then, we tested several multiplier values in the selected range to compare the results with the actual data of the identified cases from 14 days after the introduced change and adjusted the value of multipliers as necessary. In order to determine the best set of multipliers, the Fréchet distance between the number of confirmed cases predicted by the model and the real-world data was minimized. The final list of all context multipliers is presented in the ***Appendix 3***.

Regional diversity of the predicted epidemic dynamic on voivodship level is only due to a spatial structure of the synthetic society, some regional differentiation of weight multipliers motivated by regional NPIs in force before the Delta wave as well as location of infected agents spatially placed at the simulation date.

Input data and calibration

In pDyn, the infection spread is simulated on a synthetic representation of Polish society comprising about 38 million agents representing Poland's population in 2019, simulated based on Statistics Poland data, both publicly (***Statistics Poland, 2019***; ***Rakowski et al., 2010b***) and not publicly available. Non-public data was provided under the license defined in a Non-Disclosure Agreement and can be made available with the permission of the data provider. The spread of epidemics and individual virus variants begins with initial infections, which serve as an initial condition of the simulation. Data on the date and location of the initial infections have provided by the Polish Ministry of Health (please see ***Appendix 4***, containing data sources). Initial parameters are loaded together with the synthetic society at the beginning of the simulation. The initial parameters include pathogen properties (infectivity, probabilities and times of disease-related states per variant), the proportion of undetected cases, quarantine probability, cross-immunity and immunity waning parameters, and context weights, and their multipliers.

Two parameters, namely the basic pathogen variant infectivity (α) and the fraction of not self-isolating symptomatic agents (f) were fitted in the model calibration process using Bayesian optimization (***Shahriari et al., 2016***) to the real-world number of confirmed cases provided initially by Michał Rogalski and then by the Polish Ministry of Health. The remaining parameters were taken from the literature (as indicated in the model description) or estimated based on calibration experiments, such as those described for modeling changes in the contact rates for different contexts (*multipliers*). Similarly to setting optimistic and pessimistic scenarios for multipliers, we dealt with the uncertainty for the remaining model parameters by setting specific prediction intervals based on optimistic and pessimistic scenarios.

In stochastic models, such prediction intervals may arise from several interrelated sources. Firstly, it can be derived from a number of simulations carried out with alternating seeds of the pseudo-random number generator. Secondly, it can be derived from several simulations with alternating input parameter values taken from appropriate distributions. Thirdly, the assumed or prepared initial state of the system, e.g. the immunisation of the population, might strongly affect the outcome values of the simulation. Finally, the result of time-dependent curve prediction intervals for each time point forms a confidence interval.

As a result, broad prediction intervals can be obtained in the simulations of highly non-linear systems, where the small random change of input parameters might result in a significant output change. However, the broad prediction intervals appear when input parameters are delivered with a broad range of possible values or where the system's initial state features are largely unknown. In our case, the nonlinearity of the model is limited, and the main source of the output uncertainty comes from the uncertainty of various parameter values and the system's initial state. In such a situation, apart from computing the confidence corridors resulting from the randomness of the process, the two extreme scenarios have been formulated: the lowest (optimistic) and the highest (pessimistic), regarding possible but still realistic values of parameters and initial states of the system. The two scenarios determine the prediction interval for our forecast. The contrast in

uncertainty coming from different sources (random seed vs two scenarios) is illustrated for the simulation described in***Appendix 5***.

Simulation setup

Hardware

Computations are performed on Cray XC40 (Okeanos) that is part of ICM computing infrastructure. System is composed of 1084 computing nodes. Each node has 24 Intel Xeon E5-2690 v3 CPU cores with a 2-way Hyper Threading (HT) with 2.6 GHz clock frequency. Single simulation on single nodes takes around 2 hours (time depends on parameters configuration).

Model calibration

The simulation used in this study was conducted on October 28, 2021. In order to account for the uncertainty, we have formulated pessimistic and optimistic scenarios differing in the dark figure parameter (see ***Appendix 5***) that was estimated using seroprevalence and registered cases data. The pessimistic scenario proved to yield a more accurate prediction of the Delta-variant wave than the optimistic scenario. Therefore, all presented results come from the pessimistic scenario).

Testing validity of the model dynamics

It should be noted at the beginning that when testing the validity of the model, we compared the real-world data (other than those to which we calibrated the model) to our model estimates to evaluate whether the pDyn reproduced the dynamics of the epidemic accurately up to the time of simulation (i.e., October 28 2021). When testing the accuracy of the pDyn's predictions, we retrospectively compared the results obtained in the simulation with real-world data acquired after the simulation date to evaluate pDyn as a tool for predicting the future epidemic spread.

We tested the validity of the epidemic dynamics implemented in the model by comparing our simulations with real-world data regarding the dominating SARS-CoV-2 variant, immunization level in the population, and the fraction of vaccinated amongst detected cases.

The emergence of the variants of pathogen in the real world is monitored, and data are collected and accessible via Global Initiative on Sharing Avian Influenza Data (GISAID) portal (***Khare et al., 2021***). The distribution of SARS-CoV-2 variants in our model was validated by comparison with the genomic data from the GISAID. Before the day of our simulation, three dominant variants have been detected in Poland (namely, the wild type, Alpha, and Delta). To account for the possible low representativeness of the GISAID samples available for Poland, we assessed whether the curves representing the temporal succession of the wild type, Alpha and Delta variants obtained from our model mirrors the analogous "succession curves" obtained from GISAID by comparing the time convergence of reaching 25%, 50%, and 75% prevalence for each variant.

Similarly, to establish the immunization level (the fraction of agents who have been vaccinated or have undergone disease and are still immune), we compared the model results with the results of a nation-wide seroprevalence survey of adults aged 19 years and older (named OBSER-CO) run by the National Institute for Public Health in Poland (***National Institute of Public Health, 2021***, ***2023***). This data was collected in four rounds (I round: 29 March to 14 May 2021, II round: 27 July to 10 September 2021, III round: 16 November to 23 December 2021, IV round: 14 March to 4 May 2022) alongside with 95% confidence intervals for each estimate. The OBSER-CO seroprevalence estimates were used to approximate the validity of pDyn's predictions of the cumulative sum of recovered and vaccinated agents. As only the adult population was studied in the OBSER-CO study, data of agents younger than 19 years were not included in ***Figure 4***.

Lastly, using the Ministry of Health data on the age, time, and location distribution of vaccinations, pDyn model computed the fraction of vaccinated among the detected cases. We tested the validity of this estimate by comparing it with the Ministry of Health's estimate of the fraction of vaccinated detected cases in the population using mean absolute error (MAE) method.

Testing the forecast accuracy

In order to evaluate the performance of our model and the accuracy of the simulation in reproducing the COVID-19 dynamics, we compared its results to real-world data (from the Ministry of Health [quote]) using three key measures of discrepancy: (1) the difference in peak date, (2) the difference in peak value, and (3) the difference in wave length.

To calculate the differences, we first characterized the peaks of the COVID-19 pandemic by fitting a parameterized analytical function to the data indicating the occurrence of a wave. As the logistic curve is typically used to approximate a cumulative number of infected cases in epidemics (***Lee et al., 2020***; ***Postnikov, 2020***), its derivative, known as the logistic distribution, is a natural choice for a description of daily cases. The logistic distribution is parameterized by three quantities, which can be matched to our measures: (1) the mean (peak data, the central point of the wave peak), (2) the height (peak value), and (3) the width (wave length). The latter was adapted for our analysis as a full width at half-maximum (FWHM) (***Bonifazi et al., 2021***). In ***Appendix 6*** we provided a mathematical formula for calculating FWHM, as well as details and examples of the fitting procedure.

This analysis was applied to the peaks of new confirmed cases, COVID-19-related deaths, hospitalized patients, and ICU patients, both at the national and regional levels, and both for model results and real-world data. Although within the real-world data the Delta wave peaks are usually partially overlapped with arising Omicron wave peaks (not taken into account in the forecast), a sum of two logistic distributions of individual parameters were fitted in this case, and only the first peak of Delta wave was taken for further analysis. The same method was employed to test the accuracy of predictions at the level of voivodships, which are the basic administrative units in Poland where epidemic data is collected and potential NPIs are introduced.

Acknowledgments

This research was carried out with the support of the Interdisciplinary Centre for Mathematical and Computational Modelling University of Warsaw (ICM UW) under computational allocation no GS80-31.

Data and Code Availability

The research presented in this paper is based on both publicly available data and data obtained through an agreement that includes a non-disclosure agreement (NDA). The sources of the data are specified in the Supplementary materials section "S2 Data sources". We have taken all necessary measures to ensure the protection and confidentiality of the data used in this study. We recognize the importance of data sharing for scientific progress and are committed to making our data sets available to other researchers upon request, while adhering to any constraints imposed by the NDA.

All publicly available data used in this article is available in the public repository from the link: https://doi.org/10.18150/8XITKG The code used in this research is available at https://git.icm.edu.pl/covid19/1127 under Apache License 2.0.

References

Adamik B, Bawiec M, Bezborodov V, Bock W, Bodych M, Burgard JP, Götz T, Krueger T, Migalska A, Pabjan B, Ożański T, Rafajłowicz E, Rafajłowicz W, Skubalska-Rafajłowicz E, Ryfczyńska S, Szczurek E, Szymański P, Mitigation and herd immunity strategy for COVID-19 is likely to fail; 2020. doi: 10.1101/2020.03.25.20043109. Preprint at https://www.medrxiv.org/content/10.1101/2020.03.25.20043109v2.

Arashiro T, Arima Y, Muraoka H, Sato A, Oba K, Uehara Y, Arioka H, Yanai H, Kuramochi J, Ihara G, Chubachi K, Yanagisawa N, Nagura Y, Kato Y, Ueda A, Numata A, Kato H, Ishii K, Ooki T, Oka H, et al. Coronavirus Disease 19 (COVID-19) Vaccine Effectiveness Against Symptomatic Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection During Delta-Dominant and Omicron-Dominant Periods in Japan: A Multicenter Prospective Case-control Study (Factors Associated with SARS-CoV-2 Infection and the Effectiveness of COVID-19 Vaccines Study). Clinical Infectious Diseases. 2022; 76(3). doi: 10.1093/cid/ciac635.

Banks DL, Hooten MB. Statistical Challenges in Agent-Based Modeling. The American Statistician. 2021; 75(3):235–242. doi: 10.1080/00031305.2021.1900914.

Bicher M, Rippinger C, Brunmeir D, Urach C, Zechmeister M, Popper N, Agent-Based SARS-CoV-2 simulation model. Model specification. dwh GmbH - TU Wien; 2023. Report at https://www.dwh.at/en/projects/covid-19/Covid19_Model-20230322.pdf.

Bicher M, Urach C, Popper N. Gepoc ABM: A Generic agent-based population model for Austria. In: *2018 Winter Simulation Conference (WSC)*; 2018. p. 2656–2667. https://ieeexplore.ieee.org/abstract/document/8632170, doi: 10.1109/WSC.2018.8632170.

Bonifazi G, Lista L, Menasce D, Mezzetto M, Pedrini D, Spighi R, Zoccoli A. A study on the possible merits of using symptomatic cases to trace the development of the COVID-19 pandemic. The European Physical Journal Plus. 2021; 136(5):481. doi: 10.1140/epjp/s13360-021-01448-2.

Bouchnita A, Jebrane A. A Hybrid Multi-Scale Model of COVID-19 Transmission Dynamics to Assess the Potential of Non-Pharmaceutical Interventions. Chaos, Solitons, & Fractals. 2020; 138:109941. https://www.sciencedirect.com/science/article/pii/S0960077920303404, doi: 10.1016/j.chaos.2020.109941.

Bracher J, Wolffram D, Deuschel J, Görgen K, Ketterer JL, Ullrich A, Abbott S, Barbarossa MV, Bertsimas D, Bhatia S, Bodych M, Bosse NI, Burgard JP, Castro L, Fairchild G, Fuhrmann J, Funk S, Gogolewski K, Gu Q, Heyder S, et al. A pre-registered short-term forecasting study of COVID-19 in Germany and Poland during the second wave. Nature Communications. 2021; 12(1):5173. doi: 10.1038/s41467-021-25207-0.

Bracher J, Wolffram D, Deuschel J, Görgen K, Ketterer JL, Ullrich A, Abbott S, Barbarossa MV, Bertsimas D, Bhatia S, Bodych M, Bosse NI, Burgard JP, Castro L, Fairchild G, Fiedler J, Fuhrmann J, Funk S, Gambin A, Gogolewski K, et al. National and subnational short-term forecasting of COVID-19 in Germany and Poland during early 2021. Communications Medicine. 2022; 2(1):136. doi: 10.1038/s43856-022-00191-8.

Brauer F. Compartmental Models in Epidemiology. In: Brauer F, van den Driessche P, Wu J, editors. *Mathematical Epidemiology*, vol. 1945 Berlin: Springer; 2008.p. 19–79. doi: 10.1007/978-3-540-78911-6_2.

Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, Pavlin B, Vandemaele K, Van Kerkhove MD, Jombart T, Morgan O, le Polain de Waroux O. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. Eurosurveillance. 2021; 26(24). doi: 10.2807/1560-7917.ES.2021.26.24.2100509.

Carrillo-Vega MF, Salinas-Escudero G, García-Peña C, Gutiérrez-Robledo LM, Parra-Rodríguez L. Early estimation of the risk factors for hospitalization and mortality by COVID-19 in Mexico. PLOS ONE. 2020; 15(9). doi: 10.1371/journal.pone.0238905.

Coutinho RM, Marquitti FMD, Ferreira LS, Borges ME, da Silva RLP, Canton O, Portella TP, Poloni S, Franco C, Plucinski MM, Lessa FC, da Silva AAM, Kraenkel RA, de Sousa Mascena Veras MA, Prado PI. Model-based estimation of transmissibility and reinfection of SARS-CoV-2 P.1 variant. Communications Medicine. 2021; 1(1):1–8. doi: 10.1038/s43856-021-00048-6.

Delussu F, Tizzoni M, Gauvin L. Evidence of pandemic fatigue associated with stricter tiered COVID-19 restrictions. PLOS Digital Health. 2022; 1(5). doi: 10.1371/journal.pdig.0000035.

Dilaver O, Gilbert N. Unpacking a Black Box: A Conceptual Anatomy Framework for Agent-Based Social Simulation Models. Journal of Artificial Societies and Social Simulation. 2023; 26(1):4. doi: 10.18564/jasss.4998.

Dong R, Hu T, Zhang Y, Li Y, Zhou XH. Assessing the Transmissibility of the New SARS-CoV-2 Variants: From Delta to Omicron. Vaccines. 2022; 10(4):496. doi: 10.3390/vaccines10040496.

Duszyński J, Afelt A, Kossowska M, Ochab-Marcinek A, Owczuk R, Paczos W, Plater-Zyberk A, Pyrć K, Rosińska M, Rychard A, Smiatacz T. Kroniki pandemii lata 2020-2021 [Chronicles of the 2020-2021 pandemic]. Academia — Magazyn Polskiej Akademii Nauk. 2021; 4(68):1–118. doi: 10.24425/academiaPAN.2021.140621.

Eales O, de Oliveira Martins L, Page AJ, Wang H, Bodinier B, Tang D, Haw D, Jonnerby J, Atchison C, Ashby D, Barclay W, Taylor G, Cooke G, Ward H, Darzi A, Riley S, Elliott P, Donnelly CA, Chadeau-Hyam M. Dynamics of competing SARS-CoV-2 variants during the Omicron epidemic in England. Nature Communications. 2022; 13(1):4375. doi: 10.1038/s41467-022-32096-4.

Elveback LR, Fox JP, Ackerman E, Langworthy A, Boyd M, Gatewood L. An influenza simulation model for immunization studies. American Journal of Epidemiology. 1976; 103(2):152–165. doi: 10.1093/oxfordjournals.aje.a112213.

Epidemiological Model Team – ICM UW, COVID-19 w Polsce. RepOD; 2023. https://doi.org/10.18150/PXSBZM, doi: 10.18150/PXSBZM.

Epstein JM. Agent-based computational models and generative social science. Complexity. 1999; 4(5):41–60. doi: 10.1002/(SICI)1099-0526(199905/06)4:5<41::AID-CPLX9>3.0.CO;2-F.

Eurostat, Deaths by week and sex; 2023. Data at https://ec.europa.eu/eurostat/databrowser/product/view/DEMO_R_MWK_TS?lang=en.

Eurostat, Excess mortality - statistics; 2023. Available from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Excess_mortality_-_statistics.

Fazio M, Pluchino A, Inturri G, Le Pira M, Giuffrida N, Ignaccolo M. Exploring the Impact of Mobility Restrictions on the COVID-19 Spreading through an Agent-Based Approach. Journal of Transport & Health. 2022; 25:101373. https://www.sciencedirect.com/science/article/pii/S2214140522000457, doi: 10.1016/j.jth.2022.101373.

Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunuba Perez Z, Cuomo-Dannenburg G, Dighe A, Dorigatti I, Fu H, Gaythorpe K, Green W, Hamlet A, Hinsley W, Okell L, Van Elsland S, Thompson H, et al., Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Imperial College London; 2020. doi: 10.25561/77482. Report at https://doi.org/10.25561/77482.

Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S, Burke DS. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature. 2005; 437(7056):209–214. doi: 10.1038/nature04017.

Fox JP, Elveback L, Scott W, Gatewood L, Ackerman E. Herd Immunity: Basic concept and relevance to public health immunization practices. American Journal of Epidemiology. 1971; 94(3):179–189. doi: 10.1093/oxfordjournals.aje.a121310.

Frias-Martinez E, Williamson G, Frias-Martinez V. An Agent-Based Model of Epidemic Spread Using Human Mobility and Social Network Information. In: *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing* Boston, MA: IEEE; 2011. p. 57–64. doi: 10.1109/PASSAT/SocialCom.2011.142, preprint at https://doi.org/10.1109/PASSAT/SocialCom.2011.142.

Giacopelli G. A Full-Scale Agent-Based Model to Hypothetically Explore the Impact of Lockdown, Social Distancing, and Vaccination During the COVID-19 Pandemic in Lombardy, Italy: Model Development. JMIRx Med. 2021; 2(3):e24630. https://xmed.jmir.org/2021/3/e24630, doi: 10.2196/24630.

Gold JAW, Wong KK, Szablewski CM, Patel PR, Rossow J, da Silva J, Natarajan P, Morris SB, Fanfair RN, Rogers-Brown J, Bruce BB, Browning SD, Hernandez-Romieu AC, Furukawa NW, Kang M, Evans ME, Oosmanally N, Tobin-D'Angelo M, Drenzek C, Murphy DJ, et al. Characteristics and Clinical Outcomes of Adult Patients Hospitalized with COVID-19 — Georgia, March 2020. MMWR Morbidity and Mortality Weekly Report. 2020; 69(18):545–550. doi: 10.15585/mmwr.mm6918e1.

Grimm V, Railsback SF, Vincenot CE, Berger U, Gallagher C, DeAngelis DL, Edmonds B, Ge J, Giske J, Groeneveld J, Johnston ASA, Milles A, Nabe-Nielsen J, Polhill JG, Radchuk V, Rohwäder MS, Stillman RA, Thiele JC, Ayllón D. The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. Journal of Artificial Societies and Social Simulation. 2020; 23(2):7. doi: 10.18564/jasss.4259.

Grove C, Marinucci A, Montagni I. Australian Youth Resilience and Help-Seeking during COVID-19: A Cross-Sectional Study. Behavioral Sciences. 2023; 13(2):121. doi: 10.3390/bs13020121.

Haischer MH, Beilfuss R, Hart MR, Opielinski L, Wrucke D, Zirgaitis G, Uhrich TD, Hunter SK. Who is wearing a mask? Gender-, age-, and address-related differences during the COVID-19 pandemic. PLOS ONE. 2020; 15(10). doi: 10.1371/journal.pone.0240785.

Han D, Li R, Han Y, Zhang R, Li J. COVID-19: Insight into the asymptomatic SARS-COV-2 infection and transmission. International Journal of Biological Sciences. 2020; 16(15):2803–2811. doi: 10.7150/ijbs.48991.

Hoertel N, Blachier M, Blanco C, Olfson M, Massetti M, Rico MS, Limosin F, Leleu H. A Stochastic Agent-Based Model of the SARS-CoV-2 Epidemic in France. Nature Medicine. 2020; 26(9):1417–1421. https://www.nature.com/articles/s41591-020-1001-6, doi: 10.1038/s41591-020-1001-6.

James LP, Salomon JA, Buckee CO, Menzies NA. The Use and Misuse of Mathematical Modeling for Infectious Disease Policymaking: Lessons for the COVID-19 Pandemic. Medical Decision Making. 2021; 41(4):379–385. doi: 10.1177/0272989X21990391.

Jentsch PC, Anand M, Bauch CT. Prioritising COVID-19 vaccination in changing social and epidemiological landscapes: A mathematical modelling study. The Lancet Infectious Diseases. 2021; 21(8):1097–1106. doi: 10.1016/S1473-3099(21)00057-8.

Kemp F, Proverbio D, Aalto A, Mombaerts L, Fouquier d'Hérouël A, Husch A, Ley C, Gonçalves J, Skupin A, Magni S. Modelling COVID-19 dynamics and potential for herd immunity by vaccination in Austria, Luxembourg and Sweden. Journal of Theoretical Biology. 2021; 530:110874. doi: 10.1016/j.jtbi.2021.110874.

Khare S, Gurry C, Freitas L, B Schultz M, Bach G, Diallo A, Akite N, Ho J, TC Lee R, Yeo W, Core Curation Team G, Maurer-Stroh S, GISAID Global Data Science Initiative (GISAID), Munich, Germany, Bioinformatics Institute, Agency for Science Technology and Research, Singapore, Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro, Brazil, Institut Pasteur de Dakar, Dakar, Senegal, National Institutes of Biotechnology Malaysia, Selangor, Malaysia, Smorodintsev Research Institute of Influenza, St Petersburg, Russia, Genome Institute of Singapore, Agency for Science Technology and Research, Singapore, China National GeneBank, Shenzhen, China, et al. GISAID's Role in Pandemic Response. China CDC Weekly. 2021; 3(49):1049–1051. doi: 10.46234/ccdcw2021.255.

Ko JY, Danielson ML, Town M, Derado G, Greenlund KJ, Kirley PD, Alden NB, Yousey-Hindes K, Anderson EJ, Ryan PA, Kim S, Lynfield R, Torres SM, Barney GR, Bennett NM, Sutton M, Talbot HK, Hill M, Hall AJ, Fry AM, et al. Risk Factors for Coronavirus Disease 2019 (COVID-19)–Associated Hospitalization: COVID-19–Associated Hospitalization Surveillance Network and Behavioral Risk Factor Surveillance System. Clinical Infectious Diseases. 2021; 72(11). doi: 10.1093/cid/ciaa1419.

Kołodziej B, Pecka I, Trudna sytuacja w szpitalach. Wstrzymane przyjęcia, coraz mniej wolnych łóżek [Difficult situation in hospitals. Suspended admissions, fewer and fewer free beds]; 2021. Available from https://www.medonet.pl/koronawirus/koronawirus-w-polsce,koronawirus--brakuje-lozek-i-lekarzy--czwarta-fala-uderzyla-w-szpitale,artykul,58144898.html.

Lee SY, Lei B, Mallick B. Estimation of COVID-19 spread curves integrating global data and borrowing information. PLOS ONE. 2020; 15(7). doi: 10.1371/journal.pone.0236860.

Li MY, Muldowney JS. Global stability for the SEIR model in epidemiology. Mathematical Biosciences. 1995; 125(2):155–164. doi: 10.1016/0025-5564(95)92756-5.

Lombardo G, Pellegrino M, Tomaiuolo M, Cagnoni S, Mordonini M, Giacobini M, Poggi A. Fine-Grained Agent-Based Modeling to Predict Covid-19 Spreading and Effect of Policies in Large-Scale Scenarios. IEEE Journal of Biomedical and Health Informatics. 2022; 26(5):2052–2062. https://ieeexplore.ieee.org/abstract/document/9737382, doi: 10.1109/JBHI.2022.3160243.

Lorig F, Johansson E, Davidsson P. Agent-based social simulation of the Covid-19 pandemic: A systematic review. Journal of Artificial Societies and Social Simulation. 2021; 24(3):5. doi: 10.18564/jasss.4601.

Macal CM. Everything you need to know about agent-based modelling and simulation. Journal of Simulation. 2016; 10(2):144–156. doi: 10.1057/jos.2016.7.

Marshall BD. Agent-based modeling. In: El-Sayed AM, Galea S, editors. *Systems science and population health* New York: Oxford University Press; 2017.p. 87–98. doi: 10.1093/acprof:oso/9780190492397.003.0008.

Merino MG, Marinescu MC, Cascajo A, Carretero J, Singh DE. Evaluating the Spread of Omicron COVID-19 Variant in Spain. Future Generation Computer Systems. 2023; 149:547–561. https://www.sciencedirect.com/science/article/pii/S0167739X23002753, doi: 10.1016/j.future.2023.07.025.

Millington JDA, O'Sullivan D, Perry GLW. Model histories: Narrative explanation in generative simulation modelling. Geoforum. 2012; 43(6):1025–1034. doi: 10.1016/j.geoforum.2012.06.017.

Ministerstwo Zdrowia, Pierwszy przypadek koronawirusa w Polsce [The first coronavirus case in Poland]; 2020. Available from https://www.gov.pl/web/zdrowie/pierwszy-przypadek-koronawirusa-w-polsce.

MONID - MOdeling Network for severe Infectious Diseases, OptimAgent-Eng – Modellierungsnetz Für Schwere Infektionskrankheiten; 2023. Available from https://webszh.uk-halle.de/monid/?page_id=2028&lang=en.

Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO Estimates of Excess Mortality Associated with the COVID-19 Pandemic. Nature. 2023; 613(7942):130–137. https://www.nature.com/articles/s41586-022-05522-2, doi: 10.1038/s41586-022-05522-2.

Müller SA, Balmer M, Charlton W, Ewert R, Neumann A, Rakow C, Schlenther T, Nagel K. Predicting the Effects of COVID-19 Related Interventions in Urban Settings by Combining Activity-Based Modelling, Agent-Based Simulation, and Mobile Phone Data. PLOS ONE. 2021; 16(10):e0259037. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0259037, doi: 10.1371/journal.pone.0259037.

National Institute of Public Health, Ogólnopolskie Badanie Seroepidemiologiczne COVID-19 OBSER-CO Podsumowanie wyników II tury badania [Nationwide Seroepidemiological Study COVID -19 OBSER-CO. Summary of the results of the 2nd round of the study]. Warszawa: National Institute of Public Health; 2021. Report at https://www.pzh.gov.pl/download/28465/.

National Institute of Public Health, Ogólnopolskie Badanie Seroepidemiologiczne COVID-19: OBSER-CO. Raport końcowy z badania [Nationwide Seroepidemiological Study COVID-19: OBSER-CO. Final report.]. Warszawa: National Institute of Public Health; 2023. Report at https://www.pzh.gov.pl/wp-content/uploads/2023/02/OBSERCO-Raport-koncowy-z-badania.pdf.

Niedzielewski K, Nowosielski JM, Bartczuk RP, Dreger F, Górski Ł, Gruziel-Słomka M, Kaczorek A, Kisielewski J, Krupa B, Moszyński A, Radwan M, Semeniuk M, Zieliński J, Rakowski F, The Overview, Design Concepts and Details Protocol of ICM Epidemiological Model (PDYN 1.5). Rochester, NY: Social Science Research Network; 2022. doi: 10.2139/ssrn.4039054. Preprint at https://doi.org/10.2139/ssrn.4039054.

Novakovic A, Marshall AH. The CP-ABM Approach for Modelling COVID-19 Infection Dynamics and Quantifying the Effects of Non-pharmaceutical Interventions. Pattern Recognition. 2022; 130:108790. https://www.sciencedirect.com/science/article/pii/S0031320322002710, doi: 10.1016/j.patcog.2022.108790.

Pałka P, Olszewski R, Kęsik-Brodacka M, Wendland A, Nowak K, Szczepankowska-Bednarek U, Liebers DT. Using multiagent modeling to forecast the spatiotemporal development of the COVID-19 pandemic in Poland. Scientific Reports. 2022; 12(1):11314. doi: 10.1038/s41598-022-15605-9.

Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, Tobin KA, Cerfolio RJ, Francois F, Horwitz LI. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: Prospective cohort study. BMJ. 2020; doi: 10.1136/bmj.m1966.

Pinkas J, Jankowski M, Szumowski Ł, Lusawa A, Zgliczyński WS, Raciborski F, Wierzba W, Gujski M. Public health interventions to mitigate early spread of SARS-CoV-2 in Poland. Medical Science Monitor. 2020; 26. doi: 10.12659/MSM.924730.

Postnikov EB. Estimation of COVID-19 dynamics "on a back-of-envelope": Does the simplest SIR model provide quantitative parameters and predictions? Chaos, Solitons & Fractals. 2020; 135:109841. doi: 10.1016/j.chaos.2020.109841.

Priesemann V, Meyer-Hermann M, Pigeot I, Schöbel A. Der Beitrag von epidemiologischen Modellen zur Beschreibung des Ausbruchsgeschehens der COVID-19-Pandemie. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz. 2021; 64(9):1058–1066. https://link.springer.com/10.1007/s00103-021-03390-1, doi: 10.1007/s00103-021-03390-1.

Pueyo T, Coronavirus: The Hammer and the Dance; 2020. Available from https://tomaspueyo.medium.com/coronavirus-the-hammer-and-the-dance-be9337092b56.

Railsback SF, Grimm V. Agent-Based and Individual-Based Modeling: A Practical Introduction. 2 ed. Princeton University Press; 2019.

Rakowski F, Gruziel M, Bieniasz-Krzywiec Ł, Radomski JP. Influenza epidemic spread simulation for Poland — a large scale, individual based model study. Physica A: Statistical Mechanics and its Applications. 2010; 389(16):3149–3165. doi: 10.1016/j.physa.2010.04.029.

Rakowski F, Gruziel M, Krych M, Radomski JP. Large scale daily contacts and mobility model - an individual-based countrywide simulation study for Poland. Journal of Artificial Societies and Social Simulation. 2010; 13(1):13. doi: 10.18564/jasss.152.

Regulski P, Wendykier P, Kantiem K, Murdzek W. Advanced methods of visual analysis and visualization of various aspects of the COVID-19 outbreak in Poland. Procedia Computer Science. 2021; 192:4194–4199. doi: 10.1016/j.procs.2021.09.195.

Rewerska-Juśko M, Rejdak K. Social Stigma of Patients Suffering from COVID-19: Challenges for Health Care System. Healthcare. 2022; 10(2):292. doi: 10.3390/healthcare10020292.

Richiardi M. The Missing Link: AB Models and Dynamic Microsimulation. In: Leitner S, Wall F, editors. *Artificial Economics and Self Organization* Springer; 2014.p. 3–15. http://link.springer.com/10.1007/978-3-319-00912-4_1, doi: 10.1007/978-3-319-00912-4_1.

Rippinger C, Bicher M, Urach C, Brunmeir D, Weibrecht N, Zauner G, Sroczynski G, Jahn B, Mühlberger N, Siebert U, Popper N. Evaluation of undetected cases during the COVID-19 epidemic in Austria. BMC Infectious Diseases. 2021; 21(1):70. doi: 10.1186/s12879-020-05737-6.

Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, Hasell J, Macdonald B, Beltekian D, Roser M, Coronavirus pandemic (COVID-19); 2020. Data at https://ourworldindata.org/coronavirus.

Sayampanathan AA, Heng CS, Pin PH, Pang J, Leong TY, Lee VJ. Infectivity of asymptomatic versus symptomatic COVID-19. The Lancet. 2021; 397(10269):93–94. doi: 10.1016/S0140-6736(20)32651-9.

Scobie HM, Johnson AG, Suthar AB, Severson R, Alden NB, Balter S, Bertolino D, Blythe D, Brady S, Cadwell B, Cheng I, Davidson S, Delgadillo J, Devinney K, Duchin J, Duwell M, Fisher R, Fleischauer A, Grant A, Griffin J, et al. Monitoring Incidence of COVID-19 Cases, Hospitalizations, and Deaths, by Vaccination Status—13 U.S. Jurisdictions, April 4–July 17, 2021. MMWR Morbidity and Mortality Weekly Report. 2021; 70(37):1284–1290. doi: 10.15585/mmwr.mm7037e1.

Shahriari B, Swersky K, Wang Z, Adams RP, de Freitas N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. Proceedings of the IEEE. 2016; 104(1):148–175. doi: 10.1109/JPROC.2015.2494218.

Sherratt K, Gruson H, Grah R, Johnson H, Niehus R, Prasse B, Sandmann F, Deuschel J, Wolffram D, Abbott S, Ullrich A, Gibson G, Ray EL, Reich NG, Sheldon D, Wang Y, Wattanachit N, Wang L, Trnka J, Obozinski G, et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. eLife. 2023; 12. doi: 10.7554/eLife.81916.

Silverman E, Gostoli U, Picascia S, Almagor J, McCann M, Shaw R, Angione C. Situating agent-based modelling in population health research. Emerging Themes in Epidemiology. 2021; 18(1):10. doi: 10.1186/s12982-021-00102-7.

Singh DE, Luceron CO, Sanchez AL, Merino MG, Gonzalez CD, Delgado-Sanz C, Gomez-Barroso D, Carretero J, Marinescu MC. Evaluation of vaccination Strategies for the Metropolitan Area of Madrid via Agent-Based Simulation. BMJ Open. 2022; 12(12):e065937. https://bmjopen.bmj.com/content/12/12/e065937, doi: 10.1136/bmjopen-2022-065937.

Statistics Poland, Statistical Yearbook of the Republic of Poland. Warszawa: Statistics Poland; 2019. Report at https://stat.gov.pl/download/gfx/portalinformacyjny/pl/defaultaktualnosci/5515/2/19/1/rocznik_statystyczny_rzeczypospolitej_polskiej_2019.pdf.

Tran BX, Vu GT, Le HT, Pham HQ, Phan HT, Latkin CA, Ho RC. Understanding health seeking behaviors to inform COVID-19 surveillance and detection in resource-scarce settings. Journal of Global Health. 2020; 10(2):0203106. doi: 10.7189/jogh.10.0203106.

Twohig KA, Nyberg T, Zaidi A, Thelwall S, Sinnathamby MA, Aliabadi S, Seaman SR, Harris RJ, Hope R, Lopez-Bernal J, Gallagher E, Charlett A, De Angelis D, Presanis AM, Dabrera G, Koshy C, Ash A, Wise E, Moore N, Mori M, et al. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: A cohort study. The Lancet Infectious Diseases. 2022; 22(1):35–42. doi: 10.1016/S1473-3099(21)00475-8.

Vincenot CE. How New Concepts Become Universal Scientific Approaches: Insights from Citation Network Analysis of Agent-Based Complex Systems Science. Proceedings of the Royal Society B: Biological Sciences. 2018; 285(1874):20172360. https://royalsocietypublishing.org/doi/10.1098/rspb.2017.2360, doi: 10.1098/rspb.2017.2360.

Walkowiak MP, Walkowiak D. Underestimation in Reporting Excess COVID-19 Death Data in Poland during the First Three Pandemic Waves. International Journal of Environmental Research and Public Health. 2022; 19(6):3692. doi: 10.3390/ijerph19063692.

Wang H, Paulson KR, Pease SA, Watson S, Comfort H, Zheng P, Aravkin AY, Bisignano C, Barber RM, Alam T, Fuller JE, May EA, Jones DP, Frisch ME, Abbafati C, Adolph C, Allorant A, Amlag JO, Bang-Jensen B, Bertolacci GJ, et al. Estimating Excess Mortality Due to the COVID-19 Pandemic: A Systematic Analysis of COVID-19-related Mortality, 2020–21. The Lancet. 2022; 399(10334):1513–1536. https://www.thelancet.com/article/S0140-6736(21)02796-3/fulltext, doi: 10.1016/S0140-6736(21)02796-3.

Wolffram D, Abbott S, an der Heiden M, Funk S, Günther F, Hailer D, Heyder S, Hotz T, van de Kassteele J, Küchenhoff H, Müller-Hansen S, Syliqi D, Ullrich A, Weigert M, Schienle M, Bracher J. Collaborative Nowcasting of COVID-19 Hospitalization Incidences in Germany. PLOS Computational Biology. 2023; 19(8):e1011394. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1011394, doi: 10.1371/journal.pcbi.1011394.

Wong LP, Wu Q, Hao Y, Chen X, Chen Z, Alias H, Shen M, Hu J, Duan S, Zhang J, Han L. The role of institutional trust in preventive practices and treatment-seeking intention during the coronavirus disease 2019 outbreak among residents in Hubei, China. International Health. 2022; 14(2):161–169. doi: 10.1093/inthealth/ihab023.

Woolf SH, Chapman DA, Sabo RT, Weinberger DM, Hill L. Excess Deaths From COVID-19 and Other Causes, March-April 2020. JAMA. 2020; 324(5):510–513. https://doi.org/10.1001/jama.2020.11787, doi: 10.1001/jama.2020.11787.

Wu JT, Leung K, Lam TTY, Ni MY, Wong CKH, Peiris JSM, Leung GM. Nowcasting epidemics of novel pathogens: Lessons from COVID-19. Nature Medicine. 2021; 27(3):388–395. doi: 10.1038/s41591-021-01278-w.

Zhao H, Lu X, Deng Y, Tang Y, Lu J. COVID-19: asymptomatic carrier transmission is an underestimated problem. Epidemiology and Infection. 2020; 148. doi: 10.1017/S0950268820001235.

Zheng W, Kämpfen F, Huang Z. Health-seeking and diagnosis delay and its associated factors: A case study on COVID-19 infections in Shaanxi Province, China. Scientific Reports. 2021; 11(1):17331. doi: 10.1038/s41598-021-96888-2.

Appendix 1

Mitigation measures during the COVID-19 epidemic in Poland

Appendix 1—table 1. Timeline of critical mitigation measures implemented in Poland during the COVID-19 pandemic from March 2020 to December 2022

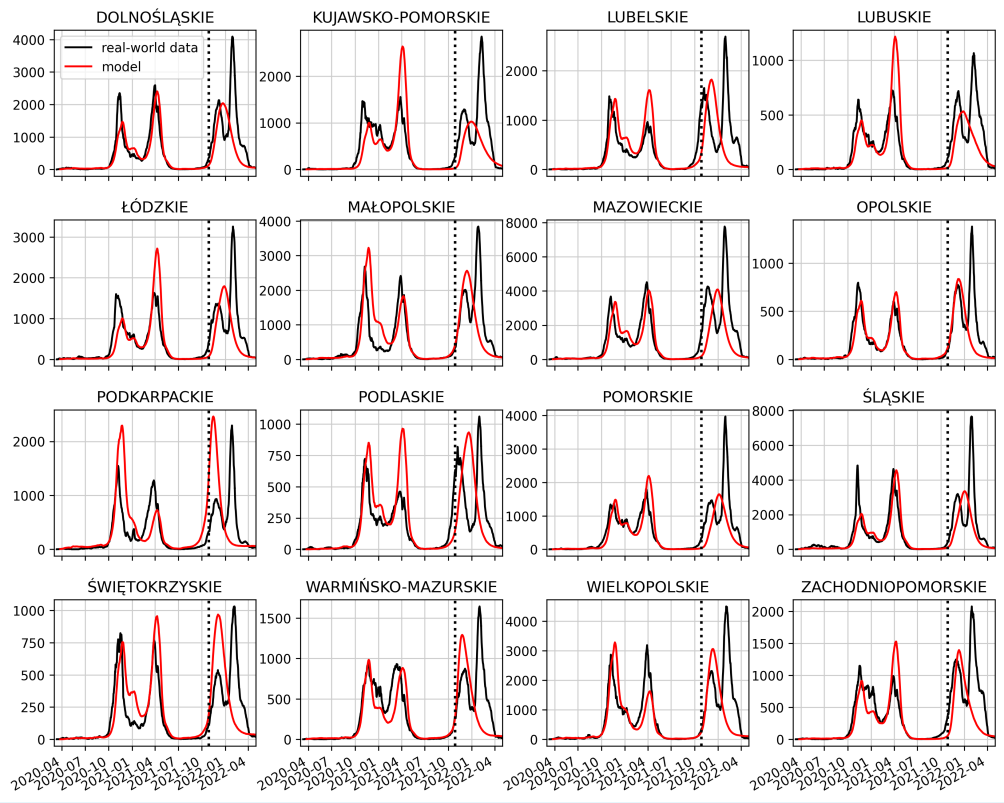
Mitigation measure	Date first introduced
Quarantine for contacts	March 4, 2020
Case detection (testing)	March 4, 2020
Work-from-home order	March 8, 2020
Ban on mass gatherings	March 10, 2020
Online schooling	March 12, 2020
Online studying at universities	March 12, 2020
Ban on entertainment events	March 14, 2020
Closure of sports gyms	March 14, 2020
Closure of hotel accommodations	March 14, 2020
Limits on the number of people in public spaces	March 15, 2020
Closure of public spaces	March 15, 2020
Stay-at-home order	March 24, 2020
Mandatory mask wearing in closed spaces	May 30, 2020
Restrictions on private gatherings	April 2, 2020
Mandatory mask wearing in open spaces	April 14, 2020
Limits in places of worship	April 19, 2020
Limits on sports gyms	June 5, 2020
Limits on hotel accommodations	October 24, 2020
Vaccination programme	December 27, 2020
Availability of booster dose vaccination	November 2, 2021

Appendix 1—table 2. Ranks description for unified restrictions calendar in Poland

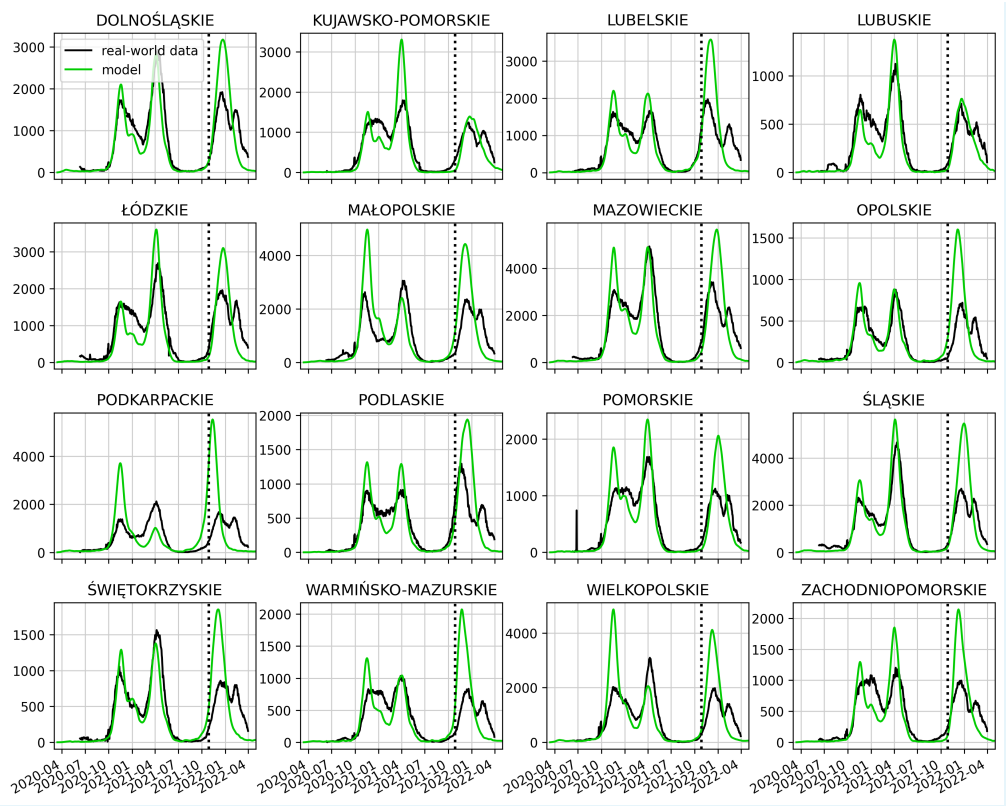
Rank	Type of restriction				
	Public space	Workplaces (services)	Universities	Schools	Kindergartens
0	No restrictions.	No restrictions.	No restrictions.	No restrictions.	No restrictions.
1	Social distancing, personal protective equipment, sanitation stations in buildings are required. Gatherings and some mass events are permitted with limits.	Social distancing, personal protective equipment, sanitation stations in buildings are required. Indoor gyms are available with limits.	Social distancing, personal protective equipment, sanitation stations in buildings are required.	Stationary education with social distancing, personal protective equipment, sanitation stations in buildings are required.	Social distancing, personal protective equipment, sanitation stations in buildings are required.
2	Public transport available with additional safety rules. Medium gatherings (approximately 100 persons) are permitted with limits (e.g., weddings).	Some capacity limits in shopping malls. Hospitality and wellness industry are available with limits. Restaurants are available with limits.	Digital learning/remote lectures are default/highly recommended, but face-to-face courses are available.	Different grades are visiting school alternately or hybrid education.	Partial availability depending on local regulations, additional safety norms, and maximum kids capacity limits.
3	Some public spaces like museums, libraries are available. Public transport limited to approximately 50% available seats. Small gatherings are permitted with limits and additional safety norms (<50 persons).	Capacity limits in shopping malls. Hospitality industry, therapeutic rehabilitation is available with strict limits. Indoor wellness industry, swimming pools are closed or strictly limited. Restaurants are strictly limited or can serve only takeaway food.	Digital learning/remote lectures are default, and face-to-face courses are strongly discouraged.	Face-to-face teaching is available only for certain grades (e.g., I-III), specialized courses (e.g., vocational classes), or final exam candidates (e.g., maturity exam).	Kindergartens are available only for kids of medical service parents.
4	Mobility is restricted to commuting or basic necessities of life. Public gathering is forbidden (limit <5 persons). Public transport limited to 25-50% available seats. Underage are not permitted to walk alone.	Shopping malls are closed or strictly limited. Hospitality and wellness industry are fully suspended. The number of people in shops and service points are strictly limited to the number of till points and surface of the point. Restaurants can serve only takeaway food.	Suspension of face-to-face teaching and transition to digital learning.	Suspension of face-to-face teaching and full transition to digital learning.	Kindergartens are suspended.

Appendix 2

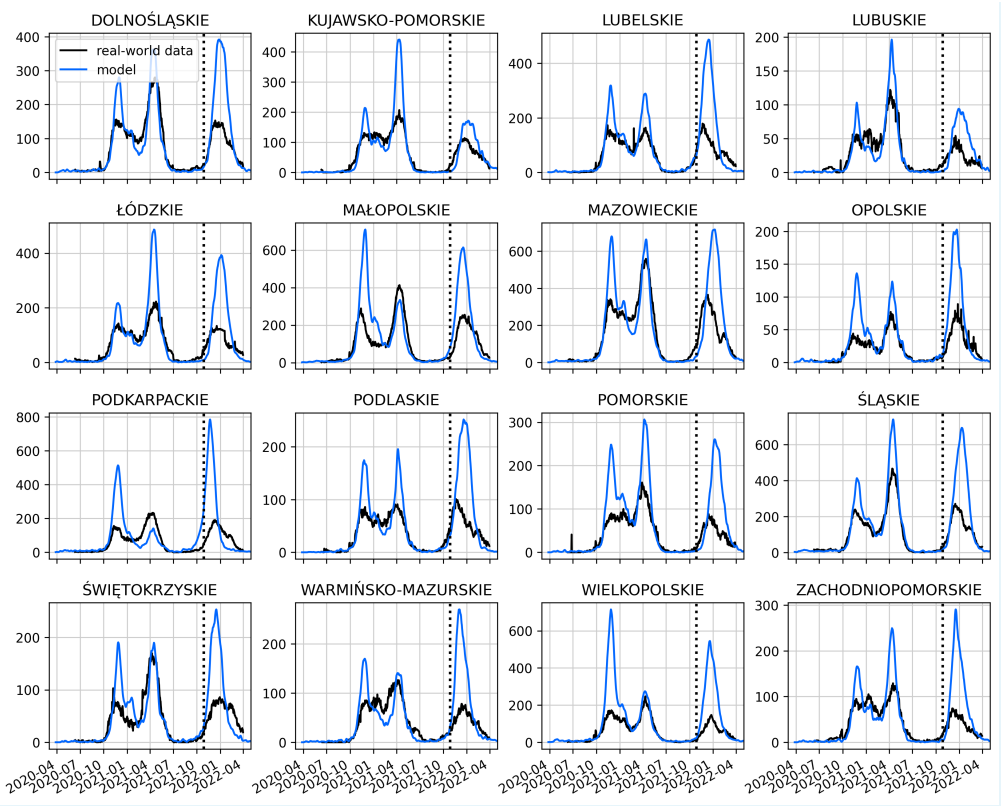
Forecast on regional level



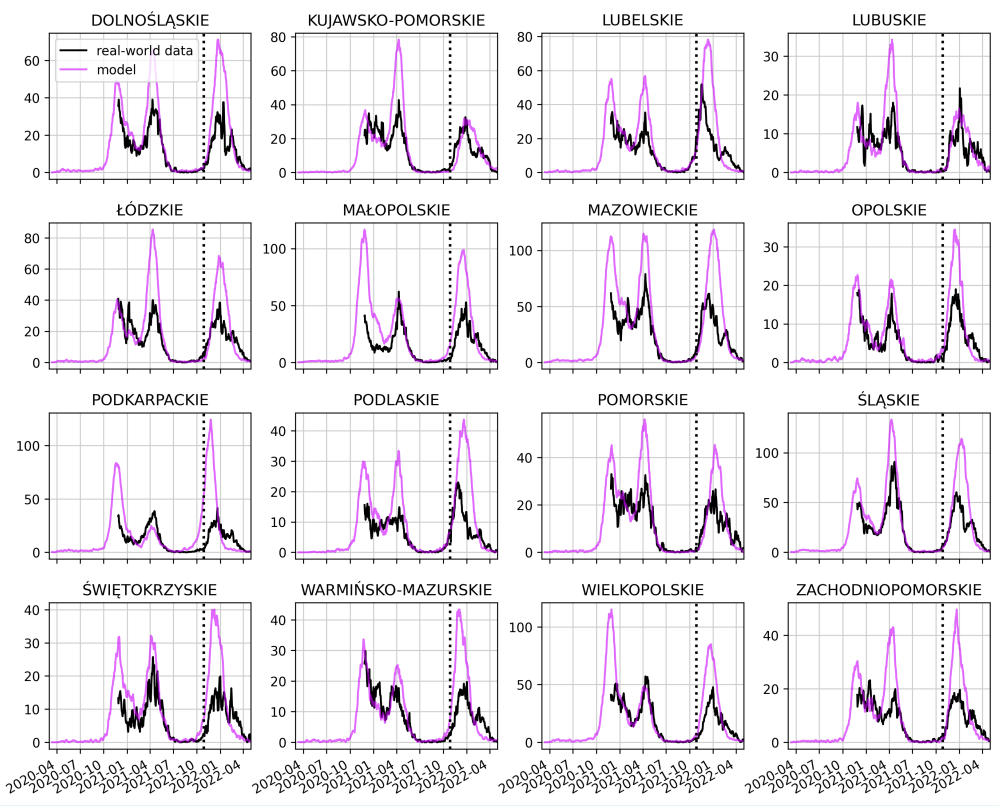
Appendix 2—figure 1. The full course of the COVID-19 detected cases in voivodships: comparison between the dynamics generated from the model (red lines) and the epidemiological data (black lines).



Appendix 2—figure 2. The full course of the COVID-19-related hospitalizations in voivodships: comparison between the dynamics generated from the model (green lines) and the epidemiological data (black lines).



Appendix 2—figure 3. The full course of the COVID-19-related ICU occupation in voivodships: comparison between the dynamics generated from the model (blue lines) and the epidemiological data (black lines).



Appendix 2—figure 4. The full course of the COVID-19-related deaths in voivodships: comparison between the dynamics generated from the model (purple lines) and the epidemiological data (black lines).

Appendix 2—table 1. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Dolnośląskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	2092	2021-12-10	62
	Real-world	2076	2021-12-03	40
	Difference	16	7	22
	Relative difference	0.77%		55.0%
Hospitalized	Simulation	3270	2021-12-21	59
	Real-world	1836	2021-12-15	55
	Difference	1434	6	4
	Relative difference	78.1%		7.27%
ICU patients	Simulation	410	2021-12-30	58
	Real-world	160	2021-12-20	60
	Difference	250	10	-2
	Relative difference	156.25%		-3.33%
Reported deaths	Simulation	70	2021-12-27	60
	Real-world	29	2021-12-24	55
	Difference	41	3	5
	Relative difference	141.38%		9.09%
Excess deaths	Simulation	70	2021-12-27	60
	Real-world	66	2021-12-16	60
	Difference	4	11	0
	Relative difference	6.06%		0.0%

Appendix 2—table 2. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Dolnośląskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	2092	2021-12-10	62
	Real-world	2076	2021-12-03	40
	Difference	16	7	22
	Relative difference	0.77%		55.0%
Hospitalized	Simulation	3270	2021-12-21	59
	Real-world	1836	2021-12-15	55
	Difference	1434	6	4
	Relative difference	78.1%		7.27%
ICU patients	Simulation	410	2021-12-30	58
	Real-world	160	2021-12-20	60
	Difference	250	10	-2
	Relative difference	156.25%		-3.33%
Reported deaths	Simulation	70	2021-12-27	60
	Real-world	29	2021-12-24	55
	Difference	41	3	5
	Relative difference	141.38%		9.09%
Excess deaths	Simulation	70	2021-12-27	60
	Real-world	66	2021-12-16	60
	Difference	4	11	0
	Relative difference	6.06%		0.0%

Appendix 2—table 3. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Kujawsko-Pomorskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	1048	2021-12-11	79
	Real-world	1318	2021-12-01	41
	Difference	-270	10	38
	Relative difference	-20.49%		92.68%
Hospitalized	Simulation	1426	2021-12-20	74
	Real-world	1179	2021-12-15	66
	Difference	247	5	8
	Relative difference	20.95%		12.12%
ICU patients	Simulation	177	2021-12-28	73
	Real-world	107	2021-12-18	64
	Difference	70	10	9
	Relative difference	65.42%		14.06%
Reported deaths	Simulation	31	2021-12-28	74
	Real-world	27	2021-12-19	58
	Difference	4	9	16
	Relative difference	14.81%		27.59%
Excess deaths	Simulation	31	2021-12-28	74
	Real-world	45	2021-12-13	60
	Difference	-14	15	14
	Relative difference	-31.11%		23.33%

Appendix 2—table 4. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Lubelskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	1824	2021-11-22	55
	Real-world	1550	2021-11-10	49
	Difference	274	12	6
	Relative difference	17.68%		12.24%
Hospitalized	Simulation	3663	2021-11-30	53
	Real-world	1938	2021-11-22	59
	Difference	1725	8	-6
	Relative difference	89.01%		-10.17%
ICU patients	Simulation	486	2021-12-11	52
	Real-world	157	2021-11-24	56
	Difference	329	17	-4
	Relative difference	209.55%		-7.14%
Reported deaths	Simulation	79	2021-12-10	55
	Real-world	45	2021-11-26	52
	Difference	34	14	3
	Relative difference	75.56%		5.77%
Excess deaths	Simulation	79	2021-12-10	55
	Real-world	68	2021-11-18	61
	Difference	11	22	-6
	Relative difference	16.18%		-9.84%

Appendix 2—table 5. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Lubuskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	535	2021-12-07	77
	Real-world	683	2021-12-04	37
	Difference	-148	3	40
	Relative difference	-21.67%		108.11%
Hospitalized	Simulation	752	2021-12-16	71
	Real-world	616	2021-12-16	56
	Difference	136	0	15
	Relative difference	22.08%		26.79%
ICU patients	Simulation	94	2021-12-27	71
	Real-world	42	2021-12-19	66
	Difference	52	8	5
	Relative difference	123.81%		7.58%
Reported deaths	Simulation	16	2021-12-22	73
	Real-world	13	2021-12-27	55
	Difference	3	-5	18
	Relative difference	23.08%		32.73%
Excess deaths	Simulation	16	2021-12-22	73
	Real-world	24	2021-12-13	48
	Difference	-8	9	25
	Relative difference	-33.33%		52.08%

Appendix 2—table 6. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Łódzkie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	1804	2021-12-14	58
	Real-world	1413	2021-11-30	49
	Difference	391	14	9
	Relative difference	27.67%		18.37%
Hospitalized	Simulation	3153	2021-12-21	55
	Real-world	1968	2021-12-13	62
	Difference	1185	8	-7
	Relative difference	60.21%		-11.29%
ICU patients	Simulation	402	2022-01-01	55
	Real-world	130	2021-12-18	80
	Difference	272	14	-25
	Relative difference	209.23%		-31.25%
Reported deaths	Simulation	67	2021-12-29	58
	Real-world	30	2021-12-20	66
	Difference	37	9	-8
	Relative difference	123.33%		-12.12%
Excess deaths	Simulation	67	2021-12-29	58
	Real-world	58	2021-12-12	62
	Difference	9	17	-4
	Relative difference	15.52%		-6.45%

Appendix 2—table 7. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Małopolskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	2579	2021-11-30	61
	Real-world	2020	2021-12-03	42
	Difference	559	-3	19
	Relative difference	27.67%		45.24%
Hospitalized	Simulation	4499	2021-12-05	58
	Real-world	2165	2021-12-10	51
	Difference	2334	-5	7
	Relative difference	107.81%		13.73%
ICU patients	Simulation	611	2021-12-15	56
	Real-world	225	2021-12-14	54
	Difference	386	1	2
	Relative difference	171.56%		3.7%
Reported deaths	Simulation	98	2021-12-14	60
	Real-world	42	2021-12-18	54
	Difference	56	-4	6
	Relative difference	133.33%		11.11%
Excess deaths	Simulation	98	2021-12-14	60
	Real-world	74	2021-12-13	63
	Difference	24	1	-3
	Relative difference	32.43%		-4.76%

Appendix 2—table 8. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Mazowieckie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	4059	2021-12-17	59
	Real-world	4289	2021-11-24	44
	Difference	-230	23	15
	Relative difference	-5.36%		34.09%
Hospitalized	Simulation	5687	2021-12-25	58
	Real-world	3307	2021-12-06	57
	Difference	2380	19	1
	Relative difference	71.97%		1.75%
ICU patients	Simulation	730	2022-01-04	58
	Real-world	359	2021-12-10	65
	Difference	371	25	-7
	Relative difference	103.34%		-10.77%
Reported deaths	Simulation	120	2022-01-02	60
	Real-world	57	2021-12-12	60
	Difference	63	21	0
	Relative difference	110.53%		0.0%
Excess deaths	Simulation	120	2022-01-02	60
	Real-world	121	2021-12-04	62
	Difference	-1	29	-2
	Relative difference	-0.83%		-3.23%

Appendix 2—table 9. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Opolskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	838	2021-11-28	54
	Real-world	786	2021-12-05	40
	Difference	52	-7	14
	Relative difference	6.62%		35.0%
Hospitalized	Simulation	1595	2021-12-06	51
	Real-world	697	2021-12-18	51
	Difference	898	-12	0
	Relative difference	128.84%		0.0%
ICU patients	Simulation	204	2021-12-16	50
	Real-world	72	2021-12-21	60
	Difference	132	-5	-10
	Relative difference	183.33%		-16.67%
Reported deaths	Simulation	33	2021-12-17	53
	Real-world	17	2021-12-17	52
	Difference	16	0	1
	Relative difference	94.12%		1.92%
Excess deaths	Simulation	33	2021-12-17	53
	Real-world	30	2021-12-13	46
	Difference	3	4	7
	Relative difference	10.0%		15.22%

Appendix 2—table 10. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Podkarpackie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	2428	2021-11-03	44
	Real-world	915	2021-11-27	49
	Difference	1513	-24	-5
	Relative difference	165.36%		-10.2%
Hospitalized	Simulation	5479	2021-11-12	43
	Real-world	1584	2021-12-07	59
	Difference	3895	-25	-16
	Relative difference	245.9%		-27.12%
ICU patients	Simulation	751	2021-11-23	42
	Real-world	182	2021-12-12	63
	Difference	569	-19	-21
	Relative difference	312.64%		-33.33%
Reported deaths	Simulation	116	2021-11-23	46
	Real-world	31	2021-12-13	54
	Difference	85	-20	-8
	Relative difference	274.19%		-14.81%
Excess deaths	Simulation	116	2021-11-23	46
	Real-world	54	2021-12-08	68
	Difference	62	-15	-22
	Relative difference	114.81%		-32.35%

Appendix 2—table 11. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Podlaskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	938	2021-12-08	61
	Real-world	754	2021-11-11	48
	Difference	184	27	13
	Relative difference	24.4%		27.08%
Hospitalized	Simulation	1998	2021-12-12	60
	Real-world	1238	2021-11-21	56
	Difference	760	21	4
	Relative difference	61.39%		7.14%
ICU patients	Simulation	261	2021-12-23	59
	Real-world	95	2021-11-26	62
	Difference	166	27	-3
	Relative difference	174.74%		-4.84%
Reported deaths	Simulation	43	2021-12-20	62
	Real-world	21	2021-11-28	47
	Difference	22	22	15
	Relative difference	104.76%		31.91%
Excess deaths	Simulation	43	2021-12-20	62
	Real-world	41	2021-11-22	58
	Difference	2	28	4
	Relative difference	4.88%		6.9%

Appendix 2—table 12. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Pomorskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	1626	2021-12-23	62
	Real-world	1472	2021-12-02	46
	Difference	154	21	16
	Relative difference	10.46%		34.78%
Hospitalized	Simulation	2038	2021-12-31	56
	Real-world	1105	2021-12-17	59
	Difference	933	14	-3
	Relative difference	84.43%		-5.08%
ICU patients	Simulation	260	2022-01-11	54
	Real-world	76	2021-12-16	59
	Difference	184	26	-5
	Relative difference	242.11%		-8.47%
Reported deaths	Simulation	43	2022-01-11	58
	Real-world	19	2021-12-21	55
	Difference	24	21	3
	Relative difference	126.32%		5.45%
Excess deaths	Simulation	43	2022-01-11	58
	Real-world	43	2021-12-18	71
	Difference	0	24	-13
	Relative difference	0.0%		-18.31%

Appendix 2—table 13. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Śląskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	3340	2021-12-21	58
	Real-world	3207	2021-12-04	40
	Difference	133	17	18
	Relative difference	4.15%		45.0%
Hospitalized	Simulation	5536	2021-12-30	55
	Real-world	2624	2021-12-17	54
	Difference	2912	13	1
	Relative difference	110.98%		1.85%
ICU patients	Simulation	689	2022-01-10	55
	Real-world	264	2021-12-20	55
	Difference	425	21	0
	Relative difference	160.98%		0.0%
Reported deaths	Simulation	115	2022-01-07	58
	Real-world	56	2021-12-22	52
	Difference	59	16	6
	Relative difference	105.36%		11.54%
Excess deaths	Simulation	115	2022-01-07	58
	Real-world	116	2021-12-15	54
	Difference	-1	23	4
	Relative difference	-0.86%		7.41%

Appendix 2—table 14. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Świętokrzyskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	974	2021-11-20	57
	Real-world	526	2021-12-01	42
	Difference	448	-11	15
	Relative difference	85.17%		35.71%
Hospitalized	Simulation	1866	2021-12-01	55
	Real-world	812	2021-12-13	61
	Difference	1054	-12	-6
	Relative difference	129.8%		-9.84%
ICU patients	Simulation	249	2021-12-12	55
	Real-world	74	2021-12-16	71
	Difference	175	-4	-16
	Relative difference	236.49%		-22.54%
Reported deaths	Simulation	40	2021-12-08	59
	Real-world	15	2021-12-20	64
	Difference	25	-12	-5
	Relative difference	166.67%		-7.81%
Excess deaths	Simulation	40	2021-12-08	59
	Real-world	25	2021-12-13	81
	Difference	15	-5	-22
	Relative difference	60.0%		-27.16%

Appendix 2—table 15. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Warmińsko-Mazurskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	1278	2021-11-11	54
	Real-world	877	2021-11-28	43
	Difference	401	-17	11
	Relative difference	45.72%		25.58%
Hospitalized	Simulation	2029	2021-11-21	52
	Real-world	803	2021-12-12	57
	Difference	1226	-21	-5
	Relative difference	152.68%		-8.77%
ICU patients	Simulation	263	2021-12-01	51
	Real-world	70	2021-12-16	77
	Difference	193	-15	-26
	Relative difference	275.71%		-33.77%
Reported deaths	Simulation	42	2021-12-02	56
	Real-world	18	2021-12-17	58
	Difference	24	-15	-2
	Relative difference	133.33%		-3.45%
Excess deaths	Simulation	42	2021-12-02	56
	Real-world	31	2021-12-13	61
	Difference	11	-11	-5
	Relative difference	35.48%		-8.2%

Appendix 2—table 16. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Wielkopolskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	3053	2021-11-28	54
	Real-world	2330	2021-12-03	38
	Difference	723	-5	16
	Relative difference	31.03%		42.11%
Hospitalized	Simulation	4075	2021-12-08	53
	Real-world	1900	2021-12-15	54
	Difference	2175	-7	-1
	Relative difference	114.47%		-1.85%
ICU patients	Simulation	522	2021-12-19	52
	Real-world	129	2021-12-22	56
	Difference	393	-3	-4
	Relative difference	304.65%		-7.14%
Reported deaths	Simulation	84	2021-12-19	56
	Real-world	38	2021-12-23	50
	Difference	46	-4	6
	Relative difference	121.05%		12.0%
Excess deaths	Simulation	84	2021-12-19	56
	Real-world	67	2021-12-15	62
	Difference	17	4	-6
	Relative difference	25.37%		-9.68%

Appendix 2—table 17. The comparison between pDyn simulation results and epidemiological data (see text) for disease-related states regarding the peak value, peak date, and width in terms of the Full-Width Half-Maximum (FWHM) of the Delta wave in Zachodniopomorskie voivodship.

Output	Comparison	Peak value	Peak timing	Width (FWHM)
New confirmed cases	Simulation	1328	2021-11-26	57
	Real-world	1255	2021-11-28	43
	Difference	73	-2	14
	Relative difference	5.82%		32.56%
Hospitalized	Simulation	2068	2021-12-05	54
	Real-world	971	2021-12-13	65
	Difference	1097	-8	-11
	Relative difference	112.98%		-16.92%
ICU patients	Simulation	260	2021-12-15	52
	Real-world	67	2021-12-13	66
	Difference	193	2	-14
	Relative difference	288.06%		-21.21%
Reported deaths	Simulation	44	2021-12-16	55
	Real-world	18	2021-12-18	60
	Difference	26	-2	-5
	Relative difference	144.44%		-8.33%
Excess deaths	Simulation	44	2021-12-16	55
	Real-world	36	2021-12-13	71
	Difference	8	3	-16
	Relative difference	22.22%		-22.54%

Appendix 3

Model parameters

Appendix 3—table 1. General model parameters.

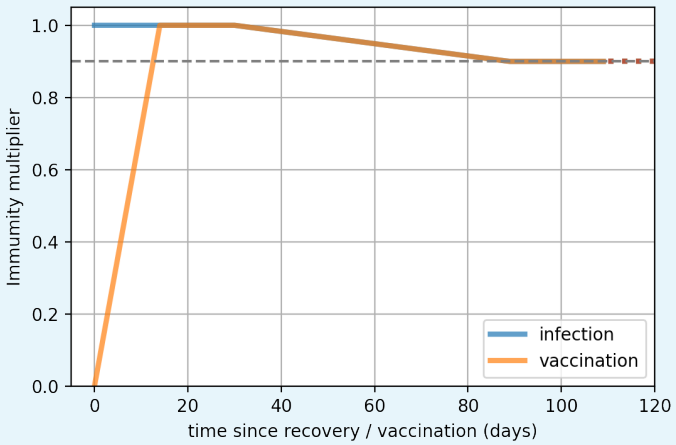
Parameter name	Parameter value
Base virus infectivity (α)	2.047 250
Base fraction of symptomatic agents leaving home (f)	0.403 245
Household contact rate	2.5
School contact rate	1.66
Preschool contact rate	1.66
Workplace contact rate	1.66
University contact rate	1.66
Travel contact rate	1.66
Street contact rate	0.83
Traveller creation rate	0.0005
Asymptomatic agents infectivity multiplier	0.1
Share of asymptomatic agents	0.8

Appendix 3—table 2. Cross-immunity matrix. Cross-immunity matrix C of size $(N + M) \times N$ is used to represent a cross-immunity phenomenon, where N is the number of variants and M is the number of vaccine types. Level of immunity against a new infection (columns), generated by infection recovery or a vaccination event (rows), is different for each variant.

Variant	Wild type	Alpha	Delta
Wild type	1	1	0.975
Alpha	1	1	0.975
Delta	0.975	0.975	1
Vaccine	1	1	0.975

Appendix 3—table 3. Parameters of new virus variants introduction.

Variant	Introduction date	Number of introduced cases
Wild	06.03.2020	1260
Alpha	25.12.2020	20000
Delta	15.05.2021	5400



Appendix 3—figure 1. Immunity multiplier function. $S(t)$ is an immunity multiplier function, representing immunity decline in time. Immunity is acquired at the moment of recovery or vaccination. Immunity multiplier for vaccines rises from 0 to 1.0 during first 14 days and is equal to 1.0 until day 30. For infections, it is changed to 1.0 immediately after the recovery. In both cases, immunity multiplier decreases linearly from 1.0 on day 30 to 0.9 on day 90 (0.0017 per day).

Appendix 3—table 4. Disease-related states duration.

State name	State duration in days
Latent	4
Asymptomatic	7
Symptomatic	5
Hospitalized, pre-ICU	13
Hospitalized, at ICU	7
Hospitalized, not at ICU	10
Recovered	1

Appendix 3—table 5. State transitions probabilities in different age groups.

	Age range (from inclusive, to exclusive)						
State transition	0–20	20–30	30–40	40–50	50–60	60-70	70+
Latent → Asymptomatic	0.92	0.92	0.84	0.84	0.68	0.63	0.23
Latent → Symptomatic	0.08	0.08	0.16	0.16	0.32	0.37	0.77
Asymptomatic → Recovered	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Symptomatic → Hospitalized, not at ICU	0.02	0.024	0.036	0.07	0.14	0.4	0.5
Symptomatic → Hospitalized, pre-ICU	0.002	0.004	0.006	0.01	0.02	0.1	0.2
Symptomatic → Dead	0.001	0.001	0.002	0.002	0.005	0.02	0.03
Symptomatic → Recovered	0.977	0.971	0.956	0.918	0.835	0.48	0.27
Hospitalized, not at ICU → Dead	0.08	0.08	0.08	0.08	0.08	0.08	0.08
Hospitalized, not at ICU → Recovered	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Hospitalized, pre-ICU → Hospitalized, at ICU	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Hospitalized, at ICU → Dead	0.75	0.75	0.75	0.75	0.75	0.75	0.75
Hospitalized, at ICU → Recovered	0.25	0.25	0.25	0.25	0.25	0.25	0.25

Appendix 3—table 6. Contexts multipliers.

Date	Household multiplier	Kindergarten multiplier	School multiplier	Workplace multiplier	University multiplier	Big university multiplier	Travel multiplier	Street multiplier	Fraction of symptomatic agents leaving home	Travellers creation multiplier	Max. travel duration	Max. travel package	Min. school age	Max. school age
06/03/2020	1	1	1	1	1	1	1	1	1	1	7	40	0	20
12/03/2020	1.01	0.2	0.2	0.8	0	0	1	0.7	1	1	7	40	0	20
14/03/2020	1.02	0.01	0.01	0.5	0	0	1	0.55	0.2	1	7	40	0	20
24/03/2020	1.02	0.01	0.01	0.35	0	0	0.25	0.4	0.1	0.75	7	20	0	20
01/04/2020	1.04	0.01	0.01	0.2	0	0	0.25	0.25	0.03	0.5	7	20	0	20
06/04/2020	1.04	0	0	0.2	0	0	0.25	0.25	0.03	0.5	7	20	0	20
11/04/2020	1.04	0	0	0.2	0	0	0.25	0.15	0.03	0.5	7	20	0	20
16/04/2020	1.04	0	0	0.2	0	0	0.25	0.1	0.03	0.5	7	20	0	20
20/04/2020	1.03	0	0	0.2	0	0	0.27	0.12	0.03	0.5	7	20	0	20
06/05/2020	1.02	0.01	0	0.2	0	0	0.27	0.12	0.03	0.5	7	20	0	20
18/05/2020	1.01	0.01	0.01	0.2	0	0	0.27	0.12	0.03	0.55	7	25	0	20
30/05/2020	1	0.01	0.01	0.2	0	0	0.27	0.12	0.03	0.6	7	25	0	20
26/06/2020	0.95	0.01	0	0.25	0	0	0.27	0.15	0.04	0.6	14	30	0	20
10/07/2020	0.95	0.1	0	0.25	0	0	0.3	0.7	0.05	0.75	14	35	0	20
10/08/2020	0.95	0.1	0	0.25	0	0	0.27	0.4	0.05	0.55	10	35	0	20
03/09/2020	1	0.25	0.25	0.35	0	0	0.27	0.55	0.05	0.55	7	35	0	20
15/09/2020	1	0.35	0.35	0.45	0	0	0.27	0.65	0.05	0.55	7	35	0	20
01/10/2020	1	0.35	0.35	0.54	0.2	0.2	0.27	0.64	0.05	0.55	7	35	0	20
10/10/2020	1	0.3	0.28	0.45	0.2	0.2	0.27	0.45	0.04	0.55	7	30	0	20
17/10/2020	1	0.3	0.26	0.31	0.2	0.2	0.25	0.36	0.03	0.5	7	30	0	20
26/10/2020	1.025	0.3	0.3	0.27	0.1	0.1	0.25	0.3	0.03	0.5	7	25	0	9
31/10/2020	1.03	0.3	0.3	0.2	0.08	0.08	0.25	0.2	0.03	0.5	7	22	0	9
07/11/2020	1.03	0.3	0.02	0.06	0	0	0.25	0.09	0.03	0.5	7	20	0	20
28/11/2020	1.04	0.3	0.02	0.23	0	0	0.25	0.28	0.03	0.5	7	20	0	20
06/12/2020	1.04	0.3	0.02	0.26	0	0	0.25	0.3	0.03	0.5	7	20	0	20
24/12/2020	1.05	0.05	0	0.1	0	0	0.35	0.65	0.03	1	10	25	0	20
28/12/2020	1.04	0.3	0.05	0.15	0	0	0.25	0.19	0.03	0.5	7	20	0	20
13/01/2021	1.04	0.3	0.05	0.2	0	0	0.25	0.2	0.03	0.5	7	20	0	20
18/01/2021	1.03	0.3	0.35	0.22	0	0	0.25	0.23	0.03	0.5	7	20	0	9
01/02/2021	1.03	0.3	0.35	0.28	0.05	0.05	0.25	0.28	0.03	1	7	25	0	9
12/02/2021	1.03	0.3	0.35	0.29	0.05	0.05	0.25	0.29	0.03	2	7	25	0	9
27/02/2021[1]	1.03	0.3	0.15	0.2	0.01	0.01	0.25	0.2	0.03	1	7	20	0	9
08/03/2021	1.03	0.3	0.35	0.18	0.05	0.05	0.25	0.18	0.03	1	7	25	0	9
09/03/2021[1]	1.03	0.3	0.15	0.16	0.01	0.01	0.25	0.16	0.03	0.5	7	20	0	9
15/03/2021[2]	1.03	0.3	0.15	0.15	0.01	0.01	0.25	0.15	0.03	0.5	7	20	0	9
20/03/2021	1.04	0.3	0.05	0.12	0.01	0.01	0.25	0.12	0.03	0.5	7	20	0	20
29/03/2021	1.04	0.02	0.02	0.12	0.01	0.01	0.25	0.12	0.03	0.5	7	20	0	20
19/04/2021	1.03	0.3	0.02	0.19	0.01	0.01	0.25	0.2	0.03	0.5	7	20	0	20
25/04/2021	1.03	0.3	0.02	0.2	0.01	0.01	0.25	0.2	0.03	0.5	7	20	0	9
26/04/2021[3]	1.03	0.3	0.15	0.18	0.01	0.01	0.25	0.18	0.03	0.5	7	20	0	9
01/05/2021	1.02	0.3	0	0.17	0.05	0.05	0.25	0.2	0.03	1.5	7	20	0	9
04/05/2021	1.02	0.3	0.1	0.12	0.05	0.05	0.25	0.12	0.03	1	7	20	0	9
08/05/2021	1.01	0.3	0.1	0.1	0.05	0.05	0.27	0.1	0.04	1	7	20	0	9
15/05/2021	1	0.3	0.05	0.05	0.05	0.05	0.27	0.06	0.04	1	7	20	0	20
21/05/2021	1	0.3	0.05	0.04	0.05	0.05	0.27	0.05	0.04	1	7	20	0	20
29/05/2021	1	0.3	0.1	0.04	0.05	0.05	0.27	0.05	0.04	1	7	20	0	20
06/06/2021	1	0.3	0.1	0.05	0.05	0.05	0.27	0.06	0.04	1	7	20	0	20
13/06/2021	1	0.3	0.07	0.08	0.03	0.03	0.27	0.08	0.04	1	7	20	0	20
26/06/2021	0.98	0.3	0	0.2	0	0	0.27	0.25	0.04	2	14	35	0	20
05/08/2021	0.98	0.3	0	0.22	0	0	0.27	0.26	0.04	2	14	35	0	20
15/08/2021	0.98	0.3	0	0.26	0	0	0.27	0.33	0.04	2	14	35	0	20
01/09/2021	1	0.3	0.2	0.28	0.05	0.05	0.25	0.33	0.03	1	7	25	0	20
01/10/2021	1.02	0.3	0.2	0.35	0.8	0.8	0.25	0.4	0.03	1	7	25	0	20
01/11/2021	1.03	0.3	0.2	0.35	0.8	0.8	0.25	0.4	0.03	1	7	25	0	20

[1] in Warmińsko-Mazurskie Voivodeship,

[2] in Lubuskie, Mazowieckie and Pomorskie Voivodeships,

[3] in Kujawsko-Pomorskie, Lubelskie, Lubuskie, Małopolskie, Mazowieckie, Podkarpackie, Podlaskie, Pomorskie, Świętokrzyskie, Warmińsko-Mazurskie and Zachodniopomorskie Voivodeships

Appendix 4

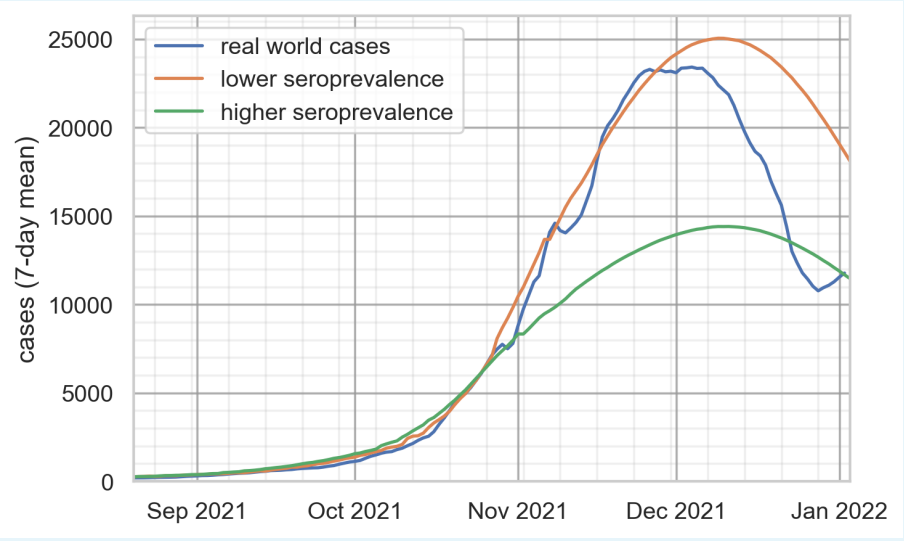
Data sources

Appendix 4—table 1. Input data sources in detail

Data type	Provider	Publicly available	Other
Household structure in Poland	Statistics Poland	No	Under NDA
Age structure in Poland	Statistics Poland	Yes	https://stat.gov.pl
Workplaces in Poland	Statistics Poland	Yes	https://stat.gov.pl
Schools in Poland	Statistics Poland	Yes	https://stat.gov.pl
Universities in Poland	Statistics Poland	Yes	https://stat.gov.pl
COVID-19 classified deaths	Michał Rogalski, Polish Ministry of Health	Yes	***Epidemiological Model Team – ICM UW* (*2023*)**, https://gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2
COVID-19 detected cases	Michał Rogalski, Polish Ministry of Health	Yes	***Epidemiological Model Team – ICM UW* (*2023*)**, https://gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2
COVID-19 hospitalized patients	Michał Rogalski, Polish Ministry of Health	Yes	***Epidemiological Model Team – ICM UW* (*2023*)**, https://twitter.com/MZ_GOV_PL
COVID-19 severeness of illness (ICU demand)	Michał Rogalski, Polish Ministry of Health	Yes	***Epidemiological Model Team – ICM UW* (*2023*)**, https://twitter.com/MZ_GOV_PL
COVID-19 time to onset of symptoms	Publications	Yes	?
COVID-19 time of sickness	The National Institute of Public Health	No	Under NDA
COVID-19 time of hospitalization	The National Institute of Public Health	No	Under NDA
Geographically spanned information about COVID-19 detected cases	Polish Ministry of Health	Yes	https://gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2
Number of people in quarantine	Polish Ministry of Health	Yes	https://gov.pl/web/koronawirus/wykaz-zarazen-koronawirusem-sars-cov-2
Non-pharmaceutical interventions	Polish Ministry of Health	Yes	https://gov.pl/web/koronawirus
Contact tracing data	The National Institute of Public Health	No	Under NDA
COVID-19 seroprevalence in Poland	The National Institute of Publish Health	Yes	https://pzh.gov.pl/projekty-i-programy/obserco/raporty
Initial contacting rates	citation	Yes	-
COVID-19 cross-immunity parameters estimation	***Scobie et al.* (*2021*)**	Yes	-

Appendix 5

Model calibration



Appendix 5—figure 1. Optimistic and pessimistic forecast scenarios. Confidence interval produced by running simulation with multiple seeds is too small to be visible.

Appendix 6

Determination of peak parameters

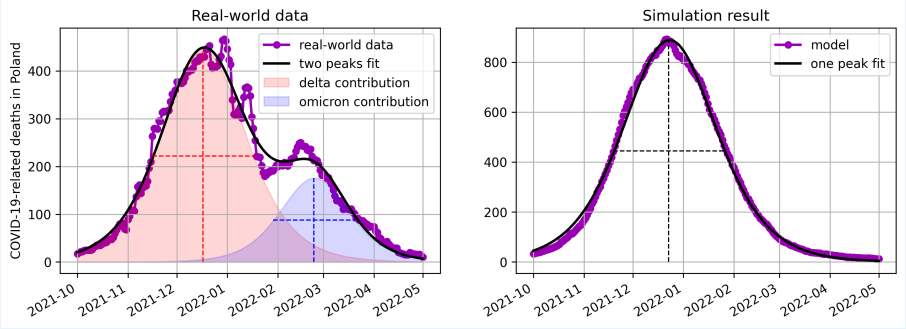
Logistic distribution was used to fit the peak data, in order to determine the peak position, the peak value, and the peak width. Its mathematical formula reads as follows:

$$f(t, t_0, h, w) = \frac{h}{\cosh^2\left(\operatorname{arccosh}(\sqrt{2}) \cdot \frac{t-t_0}{w}\right)}, \tag{1}$$

where t is time, t_0 is peak position, h is peak value, and w is peak width. Because a factor $\operatorname{arccosh}(\sqrt{2}) \approx 0.8814$ is used, the peak width appears as a full-width at half maximum (FWHM) quantity.

The fitting was done using the non-linear least squares method, provided by `curve_fit` tool from the `scipy.optimize` package, yielding the values of t_0, h, and w, which fit the best for the given data. In case of two-peaks fitting, a sum $f(t, t_1, h_1, w_1) + f(t, t_2, h_2, w_2)$ was used instead, returning best values of 6 parameters.

The examples of two-peaks and one-peak fitting to real-world and simulation result, respectively, for exemplary data of COVID-19-related deaths in Poland, are presented in Figure 1.



Appendix 6—figure 1. Example of fitting the peaks with the logistic distribution, for *delta* (and *omicron*) wave(s) of COVID-19-related deaths in Poland: (a) two-peaks fit to real-world data, (b) one-peak fit to the simulation result. Filled red and blue area in (a) show two contributing peaks. Dashed lines in both panels represent the determined parameters of the peaks: the location of the vertical line for the peak position, its length for the peak value, the length of the horizontal line for the peak width.

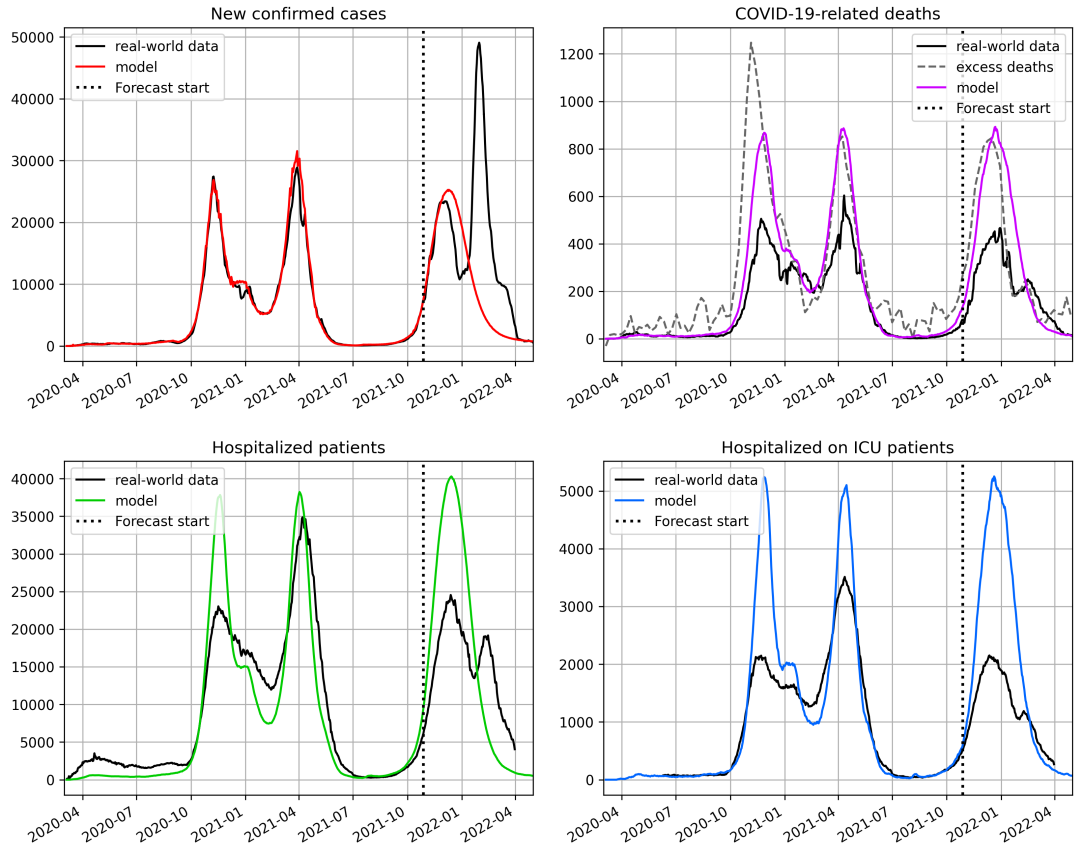


Figure 6—figure supplement 1. Comparison between the pDyn model-generated output (colored lines) and the epidemiological data published by the Polish Ministry of Health (black) and Eurostat (***Eurostat, 2023a***) (dashed grey) for the entire course of the COVID-19 epidemics in Poland. Top left: new confirmed cases. Top right: COVID-19-related deaths. Bottom left: hospitalized patients. Bottom right: ICU patients. The vertical dotted line indicates the simulation date.