Heliyon 11 (2025) e43171

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Spatial age-stratified epidemiological model with applications to South African COVID-19 pandemic

Raeesa Manjoo-Docrat [a,*], Nada Abdelatif [b], Jenny Holloway [c], Nontembeko Dudeni-Tlhone [c], Claudia Dresselhaus [d], Elona Mbayise [a], Charl Janse van Rensburg [b], Inger Fabris-Rotelli [d], Pravesh Debba [a,c], Sibusisiwe Makhanya [e]

[a] Department of Statistics and Actuarial Sciences, University of the Witwatersrand, South Africa
[b] Biostatistics Research Unit, South African Medical Research Council, South Africa
[c] Council for Scientific and Industrial Research, South Africa
[d] Department of Statistics, University of Pretoria, South Africa
[e] IBM, South Africa

A B S T R A C T

Objectives: This study aims to address the heterogeneity among different age groups in terms of their susceptibility to and transmissibility of infectious diseases. It also seeks to understand how spatial disparities affect disease spread and local population responses to emerging and re-emerging infectious diseases (EIDs and REIDs).

Design/Methods: We developed a spatial age-stratified SEIR model using COVID-19 hospitalisation data from South Africa, focusing on the first two waves of the pandemic. This model incorporates contact matrices and demographic data to capture age-dependent and spatial variations in disease dynamics.

Results: The spatial age-stratified model produced more biologically plausible and accurate predictions compared to non-stratified models investigated. It highlighted significant differences in COVID-19 risk and transmission across different age groups and regions, offering insights into targeted intervention strategies.

Conclusions: The proposed model demonstrates the importance of considering both age and spatial heterogeneity in mathematical models for infectious disease prediction. It provides a valuable tool for governments and public health officials, particularly in resource-limited settings, to develop more effective and targeted interventions. This model can be adapted for other EIDs and REIDs with similar dynamics, enhancing preparedness and response strategies.

1. Introduction

Infectious diseases such as smallpox, tuberculosis, HIV/AIDS and cholera have had significant impacts on humanity for centuries, resulting in millions of deaths [1]. Emerging and re-emerging infectious diseases (EIDs and REIDs) such as the bubonic plague, avian influenza, SARS-CoV and MERS-CoV and most recently, COVID-19 have resulted in major epidemics and pandemics due to increased global trade and travel [2]. Although infectious disease mortality has decreased over the years, the threat of infectious diseases is ever-present and disastrous. COVID-19 highlighted the importance of reliable surveillance programs to timeously detect and effectively respond to these EIDs and REIDs [2,3].

* Corresponding author. 211 MSB, University of the Witwatersrand, Johannesburg, South Africa.
E-mail address: Raeesa.docrat@wits.ac.za (R. Manjoo-Docrat).

https://doi.org/10.1016/j.heliyon.2025.e43171
Received 5 August 2024; Received in revised form 15 February 2025; Accepted 7 March 2025
Available online 7 June 2025

Pathogenic virulence of infectious diseases is dependent on many intrinsic and extrinsic factors, including age [4]. Immunity to infection at the extreme ends of the age spectrum are due to a weak innate immune system and consequently higher susceptibility to infectious diseases, whereas regulation of the inflammatory response results in age differences seen in infection-related mortality. Diseases such as the Spanish influenza, Ebola, yellow fever and tuberculosis are more severe in younger children and adults, whereas SARS-CoV, COVID-19 and MERS-CoV are more severe from 40 years and older [5]. School-aged children between 5 and 14 years of age seem to have peak immune function in terms of severity of infectious diseases [6]. In addition to virulence, transmission through contact patterns differ across age-groups. Early in the COVID-19 pandemic, studies showed that adults aged between 20 and 49 years old maintained transmission and introducing mass vaccination in those age-groups will bring resurgent epidemics under control [7].

South Africa, as with many developing countries, has a young population, where 63 % of the population is below the age of 35 years old [8]. In addition, there are substantial spatial disparities in the country concerning the quality of and access to basic and critical healthcare services, running water, and sanitation (discussed in detail in Ref. [9]). These disparities lead to significant variations in how the diseases spread and the resources available to mitigate their effects at both individual and community levels. In the early stages of the COVID-19 pandemic, many governments required the introduction of interventions such as 'lockdowns' to restrict mobility and spread of the disease. However, the capacity of the majority of the South African population to self-isolate, as well as the detrimental effect on an already significant wealth gap, means that surveillance programs to monitor and respond to EIDs and REIDs need not only consider epidemiological factors but country-specific limitations.

Mathematical models have been used widely in epidemiology to understand disease transmission dynamics as well as to predict the impact of the disease and inform how governments will respond. The work presented here has applied a spatial age-stratified SEIR model to the COVID-19 pandemic in South Africa. Although it uses COVID-19 as the focus disease, the aim of this study is to provide a general predictive framework that can be used for future EIDs and REIDs that have similar disease dynamics and where heterogeneity exists between age-groups and spatial areas.

1.1. Overview

This section details the development of an advanced SEIR (Susceptible-Exposed-Infected-Recovered) model designed to simulate the spread of infectious diseases, incorporating both spatial and age-stratified components. The basic SEIR model, which is well-documented in epidemiological literature, serves as the foundation [10]. The basic SEIR model is expanded by introducing additional compartments to capture the varying severity of the disease and potential outcomes. The compartments in this extended model include: Susceptible (S), Exposed (E), Asymptomatic (A), mild Infections (I), General Ward hospitalisations (GW), Intensive Care hospitalisations (ICU), Recovered (R), and Deceased (D). This structure allows for a more nuanced simulation of disease progression, accommodating different clinical outcomes and severities of infection, which are relevant when modelling the hospitalisations. The SEIR model is extended to include spatial elements and is further refined by stratifying the population into different age-groups. This stratification accounts for age-specific variations in susceptibility, contact rates, and disease severity. The final model integrates these extensions into a comprehensive framework. The model uses a system of differential equations to simulate the transitions between compartments for each age-group and spatial unit. This advanced SEIR model allows for the simulation of disease dynamics across different geographical areas and age-groups, providing a detailed understanding of disease spread and helping to inform public health interventions.

1.2. The spatial component

The spread of infectious diseases relies on the movement and interactions of infected individuals with susceptible people, therefore assuming uniform interaction [11] across a population in SEIR models is unrealistic. Similarly, treating each area independently and then aggregating results ignores interactions between regions. To address these issues, the SEIR model presented here is enhanced with a spatial mobility matrix [12] that captures the interactions between different spatial units, allowing a proportion of exposed individuals in one area to infect susceptible individuals in another. This matrix, defined as W, incorporates the spatial influence that each unit exerts on others, ensuring realistic modelling of disease spread across geographical boundaries. Consequently, each SEIR compartment is replicated for all spatial units, modifying the differential equations to account for these spatial interactions.

1.3. Age-stratification

For many infectious diseases, age-specific variations have been found to exist both for susceptibility to infection, and for the severity of the disease [13]. With respect to severity, one needs to consider both the probability of the infection being symptomatic [14], also referred to as the clinical fraction by Ref. [14], and the probability of hospitalisation or death of symptomatic cases. For both these probabilities, age is an important factor [15]. Infectious diseases are also spread through contact and subsequently higher contact rates of infected individuals lead to higher transmission rates. It is well known that contact rates vary substantially by age-group, with children tending to have far more contacts than older people [16]. As a result, age-varying contact matrices, together with age-specific susceptibility and age-specific severity parameters, are all important in understanding the dynamics of the spread of an infectious disease. Adding age-stratification to the model is therefore important for better accuracy in estimating the transmission rates of the disease.

1.4. Contact matrices

When adding age-stratifications in models, the rate of mixing or frequency of contact between age-groups needs to be included [17]. Contact matrices show the average number of interactions between two age-groups per day. They are necessary in modelling the spread of contact-transmissible infectious diseases as they show the transmissibility of a disease across various age-groups. Let c_{ij} denote the number of contacts that an individual in age-group i has with an individual in age-group j in a day, where $i,j \in \{1,2,3,4\}$, where 1 = age group 1, 2 = age-group 2, 3 = age-group 3 and 4 = age-group 4. These contact rates were scaled to obtain c^*_{ij}, which were used to appropriately scale the R0 for each geographical area:

$$c^*_{ij} = \left(\frac{c_{ij} - \overline{c}}{\sigma_c}\right) + 1.$$

In this equation, c_{ij} represents the original contact rate between age group i and j, $\overline{c}$ is the mean of all contact rates, and σ_c is the standard deviation of the contact rates. By subtracting the mean and dividing by the standard deviation, we standardise the data, and then adding 1 ensures that the resulting scaled contact rates have a mean of 1.

This approach allows us to maintain the relative proportions and variations in the contact rates while averaging out to the original R_0 value when applied to the basic reproduction number (R_0).

1.5. Age-specific susceptibility

Older adults, particularly those aged 65 and above, are more susceptible to severe illness and mortality from infectious diseases compared to younger individuals. Immunosenescence, the age-related decline in immune function, is a key factor contributing to increased susceptibility in older adults [18]. Age-related changes in both innate and adaptive immune responses, such as reduced T-cell function and impaired antibody production, may compromise the ability to mount an effective immune defense against pathogens [19]. Conversely, children and younger individuals generally exhibit lower susceptibility to severe illness and mortality from infectious diseases. This phenomenon has been observed in various viral and bacterial infections. Factors such as a more robust innate immune response, higher levels of mucosal antibodies, and differences in receptor expression or distribution may contribute to their relative protection [20].

Let $s_i, i \in \{1,2,3,4\}$ denote the age-specific susceptibility. Since the age-specific susceptibility is based on the risk-ratio, we have that $0 \leq s_i \leq 1$.

1.6. Age-specific transmission rates

An important set of parameters are the age-dependent transmission rates, β_{ij}, which indicate the rate of disease transmission from age group j to age group i. These infection rates factor in the age differences from both the contact matrix and the relative susceptibility (s_i for age group i), and can be defined using the following relationship:

$$\beta_{ij} = \frac{\zeta R_0 c^*_{ij} s_i}{d_i} \tag{3}$$

where: R_0 is the overall reproductive number of the infectious disease, defined as the average number of secondary infections caused by a single infected individual in a completely susceptible population [14], ζ is the scaling factor for R_0 to account for any

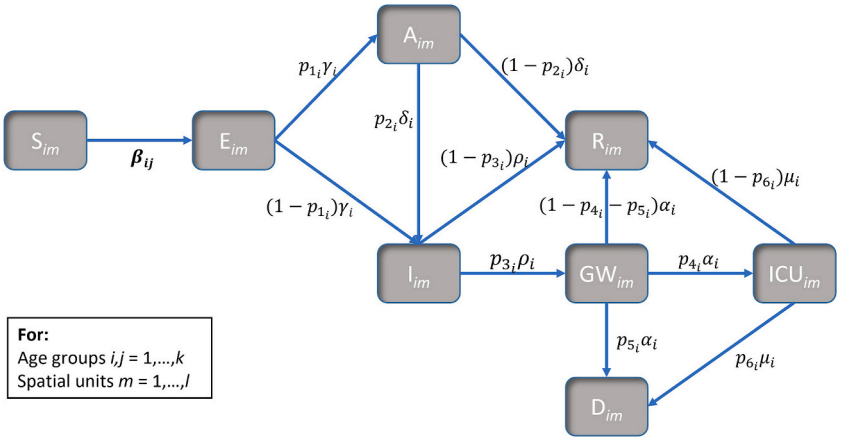

Fig. 1. Spatial age-stratified extended SEIR model structure with transition parameters.

interventions, c_{ij}^{*} is the age-specific scaling factor derived from the contact matrix, based on the number of contacts that an individual in age group i has amongst the sub-population of age group j and d_j is the mean duration of the infectious period for individuals in age group j, calculated as:

$$d_j = p_{1j}\delta_j + \left(1 - p_{1j}\right)\rho_j \tag{4}$$

with p_{1j} being the proportion of asymptomatic infections for age group j; δ_j being the number of days with asymptomatic infection, for age group j; and ρ_j being the number of days with mild symptomatic infection for age group j.

The full model is shown in Fig. 1.

2. Application to COVID-19 in South Africa

The age-stratified spatial SEIR model was developed to simulate the spread of COVID-19 in South Africa. This model incorporates variations in susceptibility, transmission dynamics, and disease severity as outlined, across different age-groups and geographic locations. The population is divided into four age-groups (0–19, 20–44, 45–59 and 60+), and the model was applied at district level (for 52 geographical areas). Only the first two waves of the pandemic are considered as this was before COVID-19 vaccines were introduced to the population. The first COVID-19 wave in South Africa occurred from June 7, 2020 to August 22, 2020, and the second COVID-19 wave occurred from November 15, 2020 to February 6, 2021. Only one wave of the pandemic is estimated at a time due to assumptions of temporary immunity.

For comparison purposes, the default (non-stratified) model was run using the average parameter values for all age-groups. The combined results from the GW and ICU components were then compared to the National Institute for Communicable Diseases (NICD) hospital surveillance data [21]. Note, however, that the NICD hospital data retrieved were not a complete reflection of all hospitalisations in South Africa since at the time of the hospitalisation peak during wave 1, only 340 of the 791 facilities had reported their COVID-19 hospital admission data.

2.1. Parameterisation

The model is parameterised using both literature and available data. The model uses data from South Africa, parameterised with estimates from global studies where necessary. The model presented in this paper is an adaption of the model presented in Refs. [9,12], which much of the parameters obtained are based off. The details of the parameters are discussed in the sections below according to their categorisation namely: wave and variant related parameters, spatial-related parameters, hospitalisation parameters and finally other important epidemiological key parameters. The details on the estimation methodology are detailed in the supplementary.

2.2. Waves, variants and associated scaling

Only the first two waves in South Africa are presented. The predominant variants during these periods are the ancestral strain and the Beta (B.1.351) variant, respectively [22]. The South African government implemented a risk-adjusted strategy with different lockdown levels, which were factored into the model by adjusting the basic reproduction number (R_0) using scaling factors derived from effective reproduction numbers during these lockdowns.[1]

2.3. Spatial weights and contact matrices

The spatial weights or mobility matrix were derived from cell phone location data (data shared with the research team for use under ethics approval NAS107/2023), indicating the movement probabilities between districts over all age groups as a whole. The contact matrix, showing average daily interactions between age-groups, was used to adjust transmission rates to be age-specific [17]. Younger individuals had higher daily contacts, influencing the spread dynamics. Age-specific susceptibility values were estimated based on risk ratios, with younger age-groups having lower susceptibility. These parameters were obtained from the literature stated and assumed to be fixed. A sensitivity analysis was performed to assess the variability resulting from the stratification.

2.4. Age-dependent parameters and hospitalisation rates

Hospitalisation data from the Gauteng DATCOV system (data shared with the research team for use under ethics approval NAS107/2023) provided age-specific rates for the progression of cases from general wards to ICUs and recovery or death. The length of stay in general wards and ICUs varied by age-group, with median values used in the model. These parameters allowed for a detailed simulation of the progression of severe cases within the hospital system. DATCOV was established as the national surveillance system for COVID-19 hospital admissions, with public and private hospitals reporting to the system by the end of 2020 [35]. It is the most

[1] https://www.google.com/url?q=https://www.cogta.gov.za/index.php/2020/04/29/risk-adjusted-strategy-gazetted-regulations-in-the-schedule/&sa=D&source=docs&ust=1720519204854447&usg=AOvVaw30bNaAwWEbi7IkZRHrUdfn.

complete source of COVID-19 hospitalisations and several routine data quality assurance steps were made to minimise errors.

2.5. Other key parameters from literature

For the age-stratified model, the key parameters given in Table 1 are assumed to be the same across all ages and across both first and second waves. This assumption could be relaxed if there are data available to obtain these estimates.

2.6. Under-counting and seeding

The model accounted for under-counting of cases by scaling reported infections and deaths based on estimates of the actual versus reported numbers. It assumed that asymptomatic cases were three times the number of symptomatic infections [9,23]. Initial values for susceptible, infected, and other compartments were calculated based on mid-year population estimates for 2020 [24], adjusted for under-counting of infections.

2.7. Uncertainty analysis

To investigate the sources of uncertainty due to age-dependent parameters, this paper focuses primarily on sensitivity analysis from a parametric perspective particularly on the age-dependent parameters to assess the model's sensitivity towards the stratification. A sensitivity analysis was conducted to understand the impact of parameter changes on the performance of the model [S5.9]. Sobol indices quantified the influence of individual age-varying parameters on the output and a one-by-one analysis was done on other key parameters. The analysis revealed that total hospitalisations were not highly sensitive to transmission probability but were more sensitive to age-specific hospitalisation rates, particularly for mild infections progressing to hospitalisation in younger age-groups. For the non age-varying parameters, total hospitalisations were most sensitive to the proportion of asymptomatic infections. A one-at a time sensitivity analysis was also conducted on the key parameters as stated in Table 1 to further investigate the uncertainty that arises from the other parameters. For the one at a time sensitivity analysis of the key parameters, each parameter was sampled uniformly from a range stated in Table S8 and the total hospitalisations were calculated based on the outcome since the hospitalisations are a key feature in the analysis.

The model was then simulated over 1000 simulations to produce a range of outcomes in an attempt to capture the inherent uncertainty. The sources of uncertainty in the simulations arise from the variations in the key parameters that are given in Table 1, capturing some of the parameter uncertainty that would exist. The parameters which are varied and the ranges are listed in Table 1. It is important to note that variability in the parameters obtained from literature were not included in the simulations. Variation in the spatial and age-related parameters would impact these results, and though they are not included in the stochastic simulations, their impact is understood by means of the sensitivity analysis conducted on those.

3. Results

3.1. First wave

Fig. 2 shows a comparison of the hospitalisation results per age-group for both the age-stratified and default (non-stratified) models for wave 1. It can be seen from this figure that the hospitalisation of children (age-group 1) and younger adults (age-group 2)are greatly over-estimated in the default model. This large over-estimation is due to these age-groups having the largest population yet having the same levels of susceptibility and hospitalisation as the older age-groups in the default model. The age-stratified model, however, uses the age-specific parameters and therefore clearly indicates that children did not require the same need for hospitalisation as the older age-groups during the COVID-19 pandemic.

The recorded hospitalisation data indicate that in-hospital COVID-19 cases peaked on August 1, 2020 during wave 1. Fig. 3 shows the number of hospitalisations per district and per age-group that were predicted by the model at that time. Although Fig. 3 only displays one snapshot in time, the full set of results indicated that the spatial age-stratified model was able to capture the spatial variation in terms of magnitude and timing of hospital peaks across the districts. The highest levels of hospitalisations are predicted for the working age groups (ages 20–59 years), in the three metropolitan municipalities, Johannesburg, Ekurhuleni and Tshwane, forming the Gauteng City Region. Higher hospitalisation rates are also predicted for the Ethekwini and Ehlanzeni municipalities, which, like the Gauteng City Region, are the economic hubs of the Kwazulu Natal and Mpumalanga provinces respectively. There is a higher degree of mobility in these areas in the age groups 20–59, as these are the economically active age groups whose mobility is associated with work in these large city regions.

A comparison of total predicted hospital cases to the hospital cases recorded nationally for wave 1 are given in Fig. 5. In this figure, it can be seen that the hospitalisations are much closer to the actual reported hospital admissions than those given by the default model. However, one should note that the actual cases are under-reported due to not all the facilities reporting their data.

3.2. Second wave

The hospitalisation results for the second wave give a similar pattern by age-group to that of the first wave, as shown in Fig. 4 but with higher peaks. Age-group 3 (ages 45–59), however, does have the highest number of hospitalisations in the age-stratified model,

Table 1
Key parameters from literature.

Parameter	Description	Estimate	Interval	Reference
r_{a_i}	Relative infectivity of asymptomatic cases	0.75	[0.7,0.8]	[23]
$\frac{1}{\gamma_i}$	Number of days in incubation	2	[2,6]	[9]
$\frac{1}{\delta_i}$	Number of days with asymptomatic infection	7	[3,7]	[9,23]
$\frac{1}{\rho_i}$	Number of days with symptomatic infection	7	[3,7]	[23]
p_{1i}	Proportion of asymptomatic infections	0.75	[0.65,0.85]	[9,23]

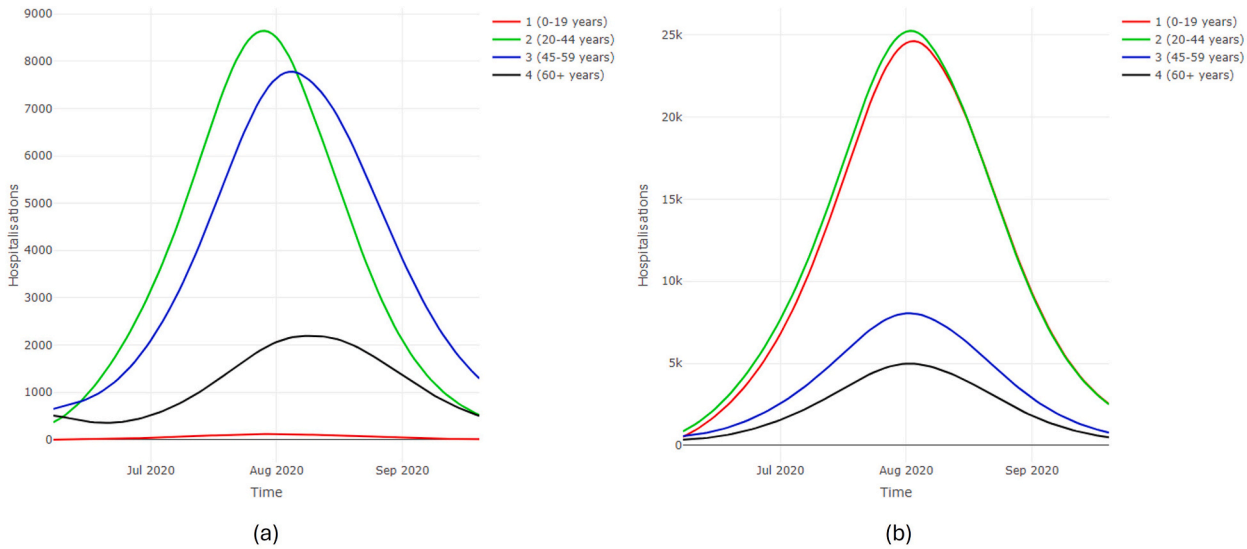

Fig. 2. Comparing (a) age-stratified to (b) default model for wave 1.

while the default model indicates age-group 1 (ages 0–19) and age-group 2 (ages 20–44) had the highest number of hospitalisations since these are the age-groups with the largest population.

Similarly to wave 1, the age-stratified model gives better predictions of hospitalisations than that of the default model, as shown in Fig. 5. Fig. 5 shows the predictions and actual values of wave 1 and wave 2 side by side. From this plot one can see the difference in magnitude of the hospitalisation peaks between wave 1 and wave 2, which was a clear effect of the increased transmissibility of the Beta variant that was prevalent during wave 2. However, discrepancies are observed between the model's predictions and actual hospitalisations during the intervals between and after waves. These differences may be attributed to limitations in the model, such as assumptions around contact rates and changes in public health interventions, as well as variability in the accuracy of reported hospitalisation data during these periods.

4. Discussion

The COVID-19 spatial age-stratified SEIR-type model presented can predict hospitalisations better than a non-stratified model would.

In order to obtain realistic infectious disease models, age needs to be factored in, since age-groups mix heterogeneously, and most risk of infection is dependent on age, and vaccination programmes are age specific [25]. In addition, the number of contacts that people in the same age-groups have will be more similar than those in different age-groups, and that results in different transmissibility rates across age-groups. Infectious diseases such as measles are more prevalent in children due to the inherent immunity that develops once infected and/or because of national immunisation programmes that countries offer for babies and children. The clinical severity of measles, however, is worse for children under five years old and for adults older than 20 years of age for unvaccinated people[2] [26]. Rubella, another infectious disease, usually results in mild illness, but affects infected pregnant women through their unborn children as it can cause miscarriage or major birth defects, and so any model needs to differentiate between the different transition rates for women of reproductive age as well their contacts with unvaccinated children who may acquire the infection through school and

[2] https://www.cdc.gov/measles/symptoms/complications.html.

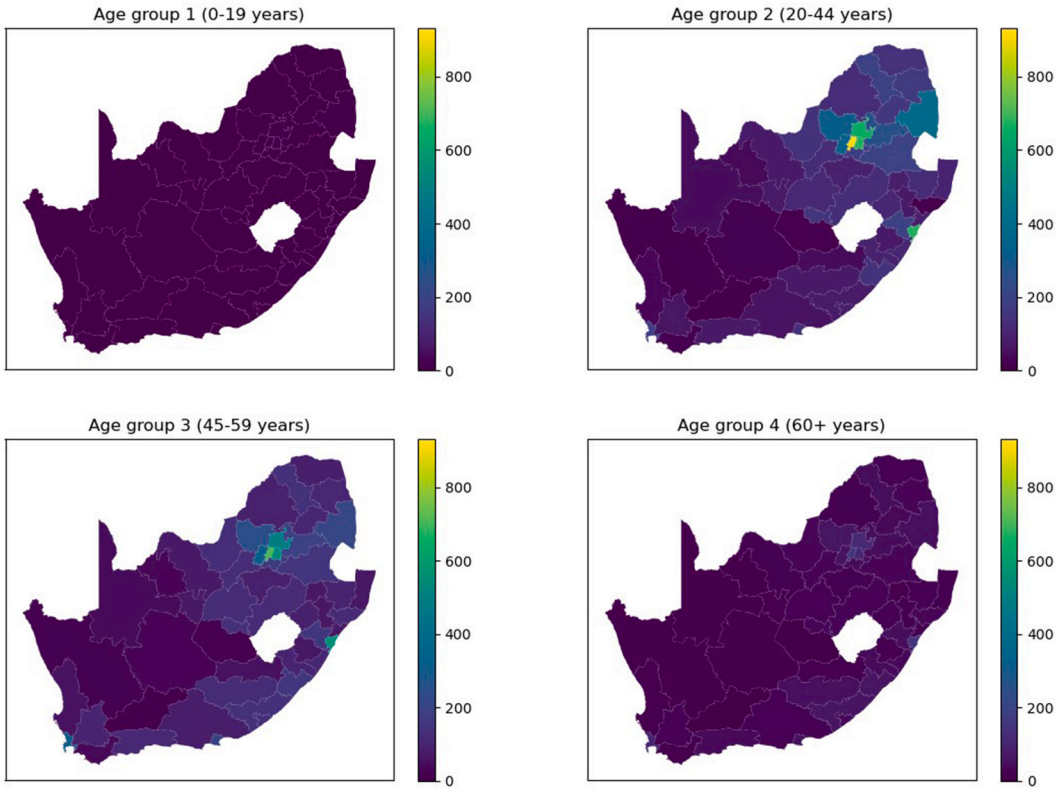

Fig. 3. Predicted hospitalisations per district and per age-group for wave 1 at the recorded hospital peak (August 1, 2020).

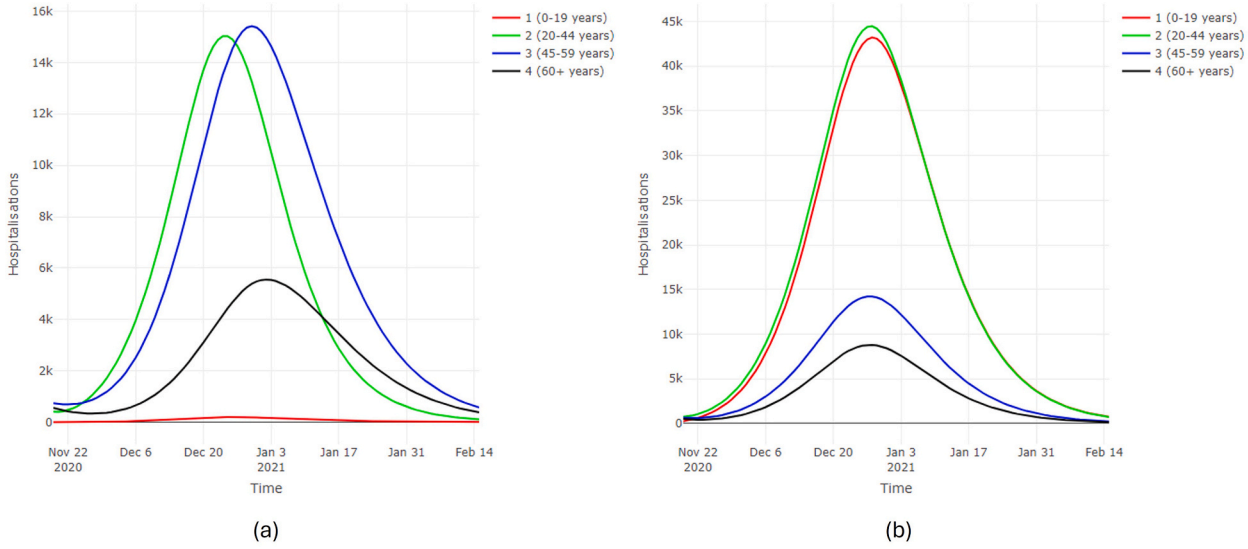

Fig. 4. Comparing (a) age-stratified model to (b) default model for wave 2.

transmit the disease to their families.[3] Age-stratification has been incorporated in models of COVID-19 [27–29]. They have determined that by accounting for age-specific rates and the age structure, an important element of the transmission risk of COVID-19 has been incorporated which results in better predictions.

[3] https://www.cdc.gov/rubella/about/transmission.html.

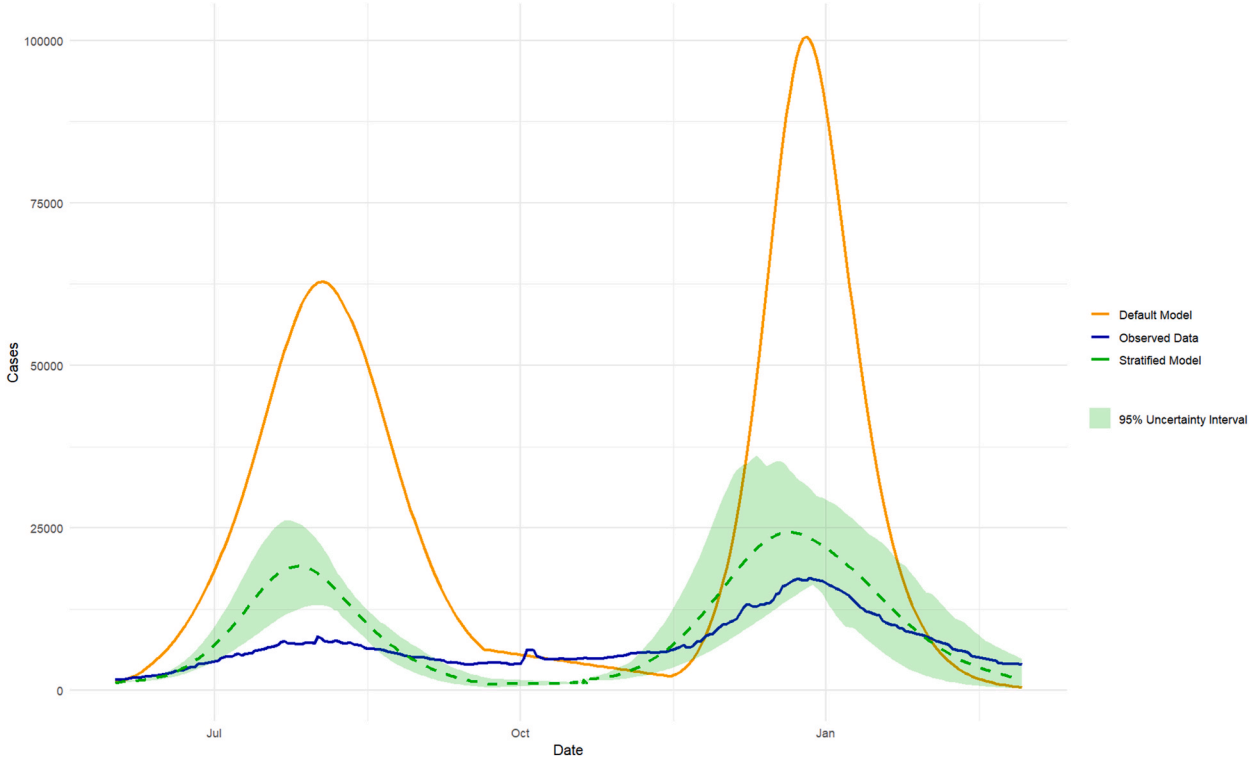

Fig. 5. Comparison of hospitalisations for waves 1 and 2 - model predictions with 95 % uncertainty intervals and actual recorded data.

Infectious diseases that are widespread over an entire country or across geographic boundaries such as Ebola virus disease, require an element of spatial modelling and the inclusion of mobility data, to understand the spread and for countries to determine quick and effective response strategies [30]. In addition, the spatial disparities that exist within a country in terms of resource capacities, societal capacities and different movement patterns requires that spatial stratification is included within models to provide more nuanced predictions [9]. In a world that is increasingly feeling the impacts of climate change, the risk from pathogens is heightened resulting in more potential transmission events [31]. Emerging and re-emerging infectious diseases such as SARS, MERS, H1N1, avian influenza, yellow fever, and the most widespread, COVID-19, have demonstrated the enormous need for the public health community to understand the spatial spread in the transmission dynamics of these diseases [32].

Grave, Viguerie, Barros et al. [33] have used partial differential equations to incorporate spatial heterogeneity in the spread of COVID-19 and have found that the predictions are in good agreement with real-world epidemiological data in different location settings. Fabris-Rotelli, Holloway, Kimmie et al. [9] have shown how a spatially-explicit model can improve the planning and preparation for COVID-19, especially at lower-level spatial units.

There has been a wave of severe infectious disease outbreaks in the 21st century, especially in the face of major global changes such as population growth, technological developments and climate change [34]. These emerging and re-emerging diseases are becoming more transmissible and widely spread and require modellers to develop more real-time, accurate models to determine the spread and impacts in terms of deaths and hospitalisations. This requires models to incorporate different demographics, such as age, to account for the fact that different demographic groups are heterogeneous and often have differing dynamics; as well as spatial-level heterogeneity. Diseases often begin with a localised outbreak that has the potential to spread at a local-scale, national-scale and across country boundaries. The more realistic a model is, and if able to provide more accurate predictions, the more useful it is to governmental institutions making intervention decisions, especially in the beginning of a disease outbreak or pandemic. It allows them to be better prepared for mitigating the impacts of these diseases in a timely and efficient manner, especially in resource-constrained settings.

5. Conclusion

Models that account for age differentials in disease transmission dynamics and describe variations in spatial distributions are more biologically plausible, whilst providing more nuanced and accurate predictions. This allows for better planning tools for governments to use, especially in the face of an increasing and devastating number of emerging and re-emerging infectious diseases.

CRediT authorship contribution statement

Raeesa Manjoo-Docrat: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Nada Abdelatif:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Jenny Holloway:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology. **Nontembeko Dudeni-Tlhone:** Writing – review & editing, Writing – original draft, Visualization, Validation. **Claudia Dresselhaus:** Writing – review & editing, Visualization, Resources. **Elona Mbayise:** Writing – original draft, Methodology. **Charl Janse van Rensburg:** Methodology, Conceptualization. **Inger Fabris-Rotelli:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Conceptualization. **Pravesh Debba:** Writing – review & editing, Resources, Project administration, Methodology, Conceptualization. **Sibusisiwe Makhanya:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged under grants 137785, COE-MaSS #2022-086-LIF-COVID-19 and #2024-025-LIF-Water Surveillance. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF. This research is also funded by Canada's International Development Research Centre (IDRC) (Grant No. 109559-001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e43171.

Data Availability and Ethics Statement

The data used in this study are not publicly available due to confidentiality agreements with the data providers. Access to the data was granted for research purposes only and remains restricted. Ethical approval for this study was obtained from the Ethics Committee of the Faculty of Natural and Agricultural Sciences, University of Pretoria (Reference number: NAS107/2023, approved October 2023).

References

[1] T.T. Yoshikawa, Epidemiology of aging and infectious diseases, Infectious Disease in the Aging (2009) 3–9, https://doi.org/10.1007/978-1-60327-534-7.
[2] J. Piret, G. Boivin, Pandemics throughout history, Front. Microbiol. 11 (2020) 631736, https://doi.org/10.3389/fmicb.2020.631736.
[3] Anthony S. Fauci, Infectious diseases: considerations for the 21st century, Clin. Infect. Dis. 32 (5) (2001) 675–685, https://doi.org/10.1086/319235.
[4] G. Sorci, B. Faivre, Age-dependent virulence of human pathogens, PLoS Pathog. 18 (9) (2022) e1010866.
[5] Zeinab Abdelrahman, Mengyuan Li, Xiaosheng Wang, Comparative review of SARS-CoV-2, SARS-CoV, MERS-CoV, and influenza a respiratory viruses, Front. Immunol. 11 (2020) 552909.
[6] J.R. Glynn, P.A.H. Moss, Systematic analysis of infectious disease outcomes by age shows lowest severity in school-age children, Sci. Data 7 (1) (2020 Oct 15) 329, https://doi.org/10.1038/s41597-020-00668-y. PMID: 33057040; PMCID: PMC7566589.
[7] M. Monod, A. Blenkinsop, X. Xi, D. Hebert, S. Bershan, S. Tietze, M. Baguelin, V.C. Bradley, Y. Chen, H. Coupland, S. Filippi, Age groups that sustain resurging COVID-19 epidemics in the United States, Science 371 (6536) (2021) eabe8372.
[8] W. He, I. Aboderin, D. Adjaye-Gbewonyo, Africa aging 2020, international population reports. https://www.census.gov/content/dam/Census/library/publications/2020/demo/p95201.pdf, 2020, 30.
[9] I. Fabris-Rotelli, J. Holloway, Z. Kimmie, S. Archibald, P. Debba, R. Manjoo-Docrat, A. le Roux, N. Dudeni-Tlhone, C. Janse van Rensburg, R. Thiede, N. Abdelatif, S. Makhanya, A. Potgieter, A spatial SEIR model for COVID-19 in South Africa, journal of data science, Statistics, and Visualisation 2 (7) (2022) 14–45.
[10] D. Daley, J. Gani, Epidemic Modelling: an Introduction, Cambridge University Press, 2001.
[11] Pratha Sah, et al., Revealing mechanisms of infectious disease spread through empirical contact networks, PLoS Comput. Biol. 17 (12) (2021) e1009604.
[12] A. Potgieter, I.N. Fabris-Rotelli, Z. Kimmie, N. Dudeni-Tlhone, J.P. Holloway, C. Janse van Rensburg, R.N. Thiede, P. Debba, R. Manjoo-Docrat, N. Abdelatif, S. Khuluse-Makhanya, Modelling representative population mobility for COVID-19 spatial transmission in South Africa, Frontiers in Big Data 4 (2021) 82.
[13] R.M. Anderson, R.M. May, Infectious Diseases of Humans: Dynamics and Control, Oxford university press, 1991.
[14] N.G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, R.M. Eggo, Age dependent effects in the transmission and control of COVID-19 epidemics, Nat. Med. 26 (8) (2020) 1205–1211, 31.
[15] A.T. Levin, W.P. Hanage, N. Owusu-Boaitey, K.B. Cochran, S.P. Walsh, G. Meyerowitz-Katz, Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and pulic policy implications, Eur. J. Epidemiol. 35 (12) (2020) 1123–1138.
[16] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G.S. Tomba, J. Wallinga, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases, PLoS Med. 5 (3) (2008) e74.
[17] K. Prem, K.v. Zandvoort, P. Klepac, R.M. Eggo, N.G. Davies, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, A. R. Cook, M. Jit, Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era, PLoS Comput. Biol. 17 (7) (2021) e1009098.
[18] G. Gavazzi, K.-H. Krause, Ageing and infection, Lancet Infect. Dis. 2 (11) (2002) 659–666.
[19] G. Pawelec, A. Akbar, C. Caruso, R. Effros, B. Grubeck-Loebenstein, A. Wikby, Is immunosenescence infectious? Trends Immunol. 25 (8) (2004) 406–410.

[20] T.R. Kollmann, B. Kampmann, S.K. Mazmanian, A. Marchant, O. Levy, Protecting the newborn and young infant from infectious diseases: lessons from immune ontogeny, Immunity 46 (3) (2017) 350–363.
[21] National Institute for Communicable Diseases, National COVID-19 daily report. https://www.nicd.ac.za/diseases-a-z-index/disease-index-covid-19/surveillance-reports/daily-hospitalsurveillance-datcov-report/. (Accessed 7 October 2021).
[22] S. Zhao, Z. Peng, M. Wang, A tentative assessment of the changes in transmissibility and fatality risk associated with Beta SARS-CoV-2 variants in South Africa: an ecological study, Pathogens and, Glob. Health 116 (2021).
[23] S. Silal, J. Pulliam, G. Meyer-Rath, B. Nichols, L. Jamieson, Z. Kimmie, H. Moultrie, On behalf of SACMC, estimating cases for COVID-19 in South Africa update: 19 may 2020. https://www.nicd.ac.za/wp-content/uploads/2020/05/SACMC19052020slides-for-MoH-media-briefing.pdf, 2020. (Accessed 10 June 2020).
[24] StatsSA, Mid-year population estimates 2020. https://www.statssa.gov.za/publications/P0302/P03022020.pdf, 2020. (Accessed 7 October 2021).
[25] H.W. Hethcote, The mathematics of infectious diseases, SIAM Rev. 42 (4) (2000) 599–653.
[26] T.H.T. Nguyen, C. Faes, N. Hens, Measles epidemic in Southern Vietnam: an age-stratified spatio-temporal model for infectious disease counts, Epidemiol. Infect. 150 (2022) e169.
[27] F. Balabdaoui, D. Mohr, Age-stratified discrete compartment model of the COVID-19 epidemic with application to Switzerland, Sci. Rep. 10 (1) (2020) 21306.
[28] A. Ludwig, P. Berthiaume, H. Orpana, C. Nadeau, M. Diasparra, J. Barnes, D. Hennessy, A. Otten, N. Ogden, Contact Tracing of COVID-19 Transmission during Restrictive Closures in Canada, CCDR 46 (11/12), 2020.
[29] B. Robinson, J.D. Edwards, T. Kendzerska, C.L. Pettit, D. Poirel, J.M. Daly, M. Ammi, M. Khalil, P.J. Taillon, R. Sandhu, et al., Comprehensive compartmental model and calibration algorithm for the study of clinical implications of the population-level spread of COVID-19: a study protocol, BMJ Open 12 (3) (2022) e052681.
[30] J.P. D'Silva, M.C. Eisenberg, Modeling spatial invasion of Ebola in west Africa, J. Theor. Biol. 428 (2017) 65–75.
[31] I. Kurane, The effect of global warming on infectious diseases, O song public health and research perspectives 1 (1) (2010) 4–9.
[32] H.-J. Li, Q. Cheng, L. Wang, Understanding spatial spread of emerging infectious diseases in contemporary populations Comment on "Pattern transitions in spatial epidemics: Mechanisms and emergent properties" by gui-quan sun et al, Phys. Life Rev. 19 (2016) 95–97.
[33] M. Grave, A. Viguerie, G.F. Barros, A. Reali, A.L. Coutinho, Assessing the spatio-temporal spread of COVID-19 via compartmental models with diffusion in Italy, USA, and Brazil, Arch. Comput. Methods Eng. 28 (2021) 4205–4223.
[34] R.E. Baker, A.S. Mahmud, I.F. Miller, M. Rajeev, F. Rasambainarivo, B.L. Rice, S. Takahashi, A.J. Tatem, C.E. Wagner, L.-F. Wang, et al., Infectious disease in an era of global change, Nat. Rev. Microbiol. 20 (4) (2022) 193–205.
[35] W. Jassat, C. Cohen, S. Tempia, M. Masha, S. Goldstein, T. Kufa, et al., Risk factors for Covid-19-related in-hospital mortality in a high HIV and tuberculosis prevalence setting in South Africa: a cohort study, Lancet HIV (8) (2021) e554–e567, https://doi.org/10.1016/S2352-3018(21)00151-X.