Proceedings of ALGORITMY 2016
pp. 285–291

QUANTITATIVE ANALYSIS OF THE AGE STRUCTURED MATHEMATICAL MODEL OF VARICELLA SPREAD IN SLOVAKIA *

J. ZIBOLENOVÁ†, D. ŠEVČOVIČ‡, T. BAŠKA†, AND H. HUDEČKOVÁ†

Abstract. Varicella (chickenpox) is one of the most frequent infectious diseases in childhood. The aim of this work is to describe the incidence and susceptibility to varicella in Slovak population using the age-structured SIR (Susceptible - Infected - Recovered) model.

The model was calibrated to estimate disease incidence in the Slovak population. Sensitivity analysis of transmission parameters and initial conditions were performed. This model is appropriate to describe the spread of disease and susceptibility of a large population in separate age groups.

Key words. mathematical modeling, varicella, susceptibility of population, disease spread, SIR

AMS subject classifications. 92C60, 91A22

1. Introduction. Varicella (chickenpox) is one of the most frequent infectious diseases in the World [1]. In countries not covered by mass vaccination, it spreads mainly among children. The clinical course of a disease is usually uncomplicated, the risk of complications depends on age and immune status of a patient. The disease is caused by varicella zoster virus (VZV) [2, 3]. Recovering from the disease leads to a permanent immunity, so most of adults are immune against it.

Vaccination is the most effective method of prevention against varicella. In some countries (e.g. US, Germany), unlike Slovakia, it has been included into mass vaccination [4]. Vaccination against varicella is available in the Slovak Republic only recently and only a small part of the population has been vaccinated until now [5]. It is expected that significantly increased vaccination coverage could potentially cause decline in morbidity rate of the disease similarly to the countries with high vaccination coverage [6, 7, 8, 9, 10].

The objective of this work is to find a suitable mathematical model to describe the spread of varicella in Slovakia, which could be used to predict changes in the spread following increase of the vaccination rates.

2. Description of the model. To describe circulation of varicella in the Slovak Republic deterministic, age-stratified SIR (susceptible-infected-recovered) model [11, 12, 13, 14, 15] with uniform age structure was used.

Population with constant size of 5000000 inhabitants was uniformly divided into 8 age groups (0-years, 1-years, 2-4 years, 5-9 years, 10-14 years, 15-24 years, 25-64 years and 65+), including 60000 individuals in each year of age. Such population does not differ too much from the population obtained by averaging the age structure in the last 15 and the next 50 years, corresponding to the time horizon of the prediction. Each of the age groups was further divided into susceptible (S), infectious (I) and resistant (R) individuals.

* This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0096-12

† Department of Public Health, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Martin, Slovakia (jana.zibolenova@gmail.com).

‡ Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Bratislava, Slovakia (sevcovic@fmph.uniba.sk).

There are several types of transition across the groups in the model: ageing, natality, mortality, disease transmission and recovery. The influx of new susceptible individuals occurred by births and the size of population remained constant due to deaths (only in 65+ group). Mortality and birth rates have continuous character with the same value of 1.2% annually (i. e. 60000 births and deaths annually). Transmission of the disease was described with the standard SIR model [11, 12, 13, 14, 15].

Differential equations for the first age group (0-years) have the following form:

$$\begin{aligned}
\frac{dS_0}{dt} &= P - S_0 \sum_{i=1}^{8} \beta_{0i} I_i - \frac{PS_0}{S_0 + I_0 + R_0} \\
\frac{dI_0}{dt} &= S_0 \sum_{i=1}^{8} \beta_{0i} I_i - \gamma I_0 - \frac{PI_0}{S_0 + I_0 + R_0} \\
\frac{dR_0}{dt} &= \gamma I_0 - \frac{PR_0}{S_0 + I_0 + R_0}
\end{aligned}$$

The equations for other age groups have the form:

$$a \in \{1, 2-4, 5-9, 10-14, 15-24, 25-64, 65+\}$$

$$\begin{aligned}
\frac{dS_a}{dt} &= -S_a \sum_{i=1}^{8} \beta_{ai} I_i + \frac{PS_{a-1}}{S_{a-1} + I_{a-1} + R_{a-1}} - \frac{PS_a}{S_a + I_a + R_a} \\
\frac{dI_a}{dt} &= S_a \sum_{i=1}^{8} \beta_{ai} I_i - \gamma I_a + \frac{PI_{a-1}}{S_{a-1} + I_{a-1} + R_{a-1}} - \frac{PI_a}{S_a + I_a + R_a} \\
\frac{dR_a}{dt} &= \gamma I_a + \frac{PR_{a-1}}{S_{a-1} + I_{a-1} + R_{a-1}} - \frac{PR_a}{S_a + I_a + R_a}
\end{aligned}$$

S_a is a proportion of susceptible individuals in age group a, I_a is a proportion of infected ones, R_a is a proportion of immune individuals to the total population, t is a time. β_{ia} are transmission parameters between age groups a and i (WAIFW matrix) and γ is a recovery rate (1/length of infectious period, in this case the length of infectious period is 7 days [2]. P denotes the transition between the age groups due to ageing and also the birth rate and mortality rate. Time horizon was 80 years.

The primary output of the model represents the course of the functions of proportions of susceptible, infected and resistant individuals depending on time. However, this output cannot be directly compared against the reality. Therefore, function of proportion of infected individuals was adjusted for length of infectious period.

Moreover, to prevent distortion of results due to fluctuations in the beginning of the simulation, only the values of the last 40 years of simulation were considered. Thus, one of the most important modified outputs of the model is an average annual modelled number of infected 40-80 years since start of simulation $40I_a$ in an age group a. It was calculated for each simulation using the approximate relationship:

$$40I_a = \sum_{t=40*365}^{80*365} I_a(t) \frac{N\gamma}{40},$$

where $I_a(t)$ is the primary output of the model and it represents the proportion of infectious individuals at the time t (days) in the age group a, and N is the population size.

TABLE 4.1

Average annual numbers of varicella cases reported to the database EPIS in 1997-2012, percentage of reported cases in a given age group, I_a^{real}, the age structure of reference model population N_a.

	0	1–4	5–9	10–14	15–24	25–64	65+	overall
Average annual numbers of varicella cases reported (min-max)	498 (353-754)	6245 (4258-9756)	8627 (6722-12573)	2385 (1850-2907)	711 (520-976)	417 (259-740)	7 (2-14)	18890 (14346-27720)
Percentage of reported case in a given age group [18]	57%	31%	30%	37%	65%	90%	100%	
I_a^{real} – average value (min-max)	941 (683-1353)	21486 (15750-29236)	28486 (22426-34336)	5651 (3740-7762)	784 (529-1136)	380 (218-725)	13 (3-25)	57741 (43349-71328)
Age structure of reference model population N_a	60000	60000+ 180000	300000	300000	600000	2400000	1100000	5000000

3. Who Acquires Infection From Whom matrix (WAIFW matrix). A key precondition for the proper functioning of the model is to find the appropriate values β_{ij} in the WAIFW matrix. WAIFW matrix reflects the structure of effective contacts in population among age groups. Diagonal elements reflect probability of acquiring infection from individual within the same age group, and other elements represent effective contacts across different age groups.

The structure of WAIFW matrix has the form of a symmetric matrix:

$$WAIFW: \quad \beta = \begin{pmatrix} k_1 & k_1 & k_1 & k_1 & k_1 & k_6 & k_7 & k_8 \\ k_1 & k_2 & k_2 & k_2 & k_2 & k_6 & k_7 & k_8 \\ k_1 & k_2 & k_3 & k_3 & k_5 & k_6 & k_7 & k_8 \\ k_1 & k_2 & k_3 & k_4 & k_5 & k_6 & k_7 & k_8 \\ k_1 & k_2 & k_5 & k_5 & k_5 & k_6 & k_7 & k_8 \\ k_6 & k_6 & k_6 & k_6 & k_6 & k_6 & k_7 & k_8 \\ k_7 & k_7 & k_7 & k_7 & k_7 & k_7 & k_7 & k_8 \\ k_8 & k_8 & k_8 & k_8 & k_8 & k_8 & k_8 & k_8 \end{pmatrix}$$

where the optimal values k_i are obtained from calibration of model considering the true incidence of the disease. Sensitivity analysis to these parameters was made, as well.

4. Calibration of the model. The estimation of WAIFW parameters in the most of published models of varicella spread [16, 17, 18] are based on seroprevalence studies [12]. Since in Slovakia currently such data are not available, we decided to proceed in a different way - calibration considering the actual incidence of the disease.

Data on the incidence of the disease are available from the database of the Epidemiological Information System (EPIS) used officially in Slovakia to record reported cases of infectious diseases [19]. These data refer only to reported diseases and a large numbers of cases remain unreported. They regard the spread of the disease which is as important as reported cases. In [18] Brisson *et al.* dealt with a problem of unreported cases in each age group and estimated ratio between reported cases and all cases of the disease. Table 4.1 shows estimations of actual numbers of cases in respective age groups in Slovak population calculated through estimations prepared by above mentioned authors.

Over the last 16 years (years 1997-2012), in average about 19000 cases per year were reported to the EPIS database (see Table 4.1), [7, 10]). Taking into account the unreported cases, the total number increased up 58000 cases per year, which roughly corresponds with the influx of new (susceptible) individuals.

TABLE 4.2
Optimal values of k_i in the WAIFW matrix.

k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	$Criterion(k)$
0.200	1.299	1.560	13.882	4.484	0.359	0.051	0.004	1.47×10^{-6}

It was necessary to take the different demographic structure of the real and the model population into account. The parameter, whereby the model was calibrated, was called I_a^{real} and it represents the annual average expected number of all cases (both reported and unreported) in a uniform population after direct standardization [20].

Optimal values of k_i (see Table 4.2) were obtained by minimization of criterion:

$$Criterion(k) = \left(\frac{1}{8} \sum_{a=1}^{8} \frac{(40I_a(k) - I_a^{real})^2}{(I_a^{real})^2} \right)^{\frac{1}{2}}$$

where $40I_a(k)$ is average model number of infected in last 40 years of simulation in age group a. Minimization was performed in Matlab using the function `fminsearch`, yielding the minimal value of criterion 1.47×10^{-6}. It could be considered as very satisfactory result. Model calculations differed from the actual number of cases on the level of decimal places, which far exceeds the accuracy of the estimate of the actual number of cases.

The next output parameter was a proportion of susceptible individuals in age groups $40S_a$, defined as the mean proportion of susceptible individuals in a given age group a and referred to a number of individuals in this age group during last 40 years of simulation. This output can be compared with data from seroprevalence studies [21].

5. Sensitivity of the model with respect to elements in the WAIFW matrix. After finding the most appropriate combination of parameters k_i, it was necessary to analyse the sensitivity of model to k_i parameters. After deviating of each k_i parameter by 20% down and up, 3^8 combinations have arisen. For each combination of k_i parameter, the simulation was generated and outputs $40S_a$ and $40I_a$ were calculated. These results were compared subsequently with I_a^{real}. Initial conditions were the same in all simulations.

TABLE 5.1
Average, minimum and maximum values of $40I_a$ in appropriate age groups of across all simulations.

	0	1	2-4	5-9	10-14	15-24	25-64	65+	overall
average	940	5353	16109	28390	5682	799	386	13	57671
max	1139	6439	18823	32939	7163	1275	609	20	58507
min	739	4262	13521	24425	4424	491	244	9	56394

Following deeper analysis of the dependence of output parameters ($40I_a$, Table 5.1) on k_i parameters, we found that all simulations are plausible, although some of them are beyond of the interval specified by I_a^{real}. Respective k_i parameters affect output elements accordingly with their position in WAIFW matrix (see Figure 5.1). Total number of infected is mostly affected with k_i parameters which are related to 3^{rd}, 4^{th} and 5^{th} age group. This corresponds with the fact that the disease is the most frequent in the mentioned age groups.

Output $40S_a$ can be compared with the seroprevalence studies. Complete data

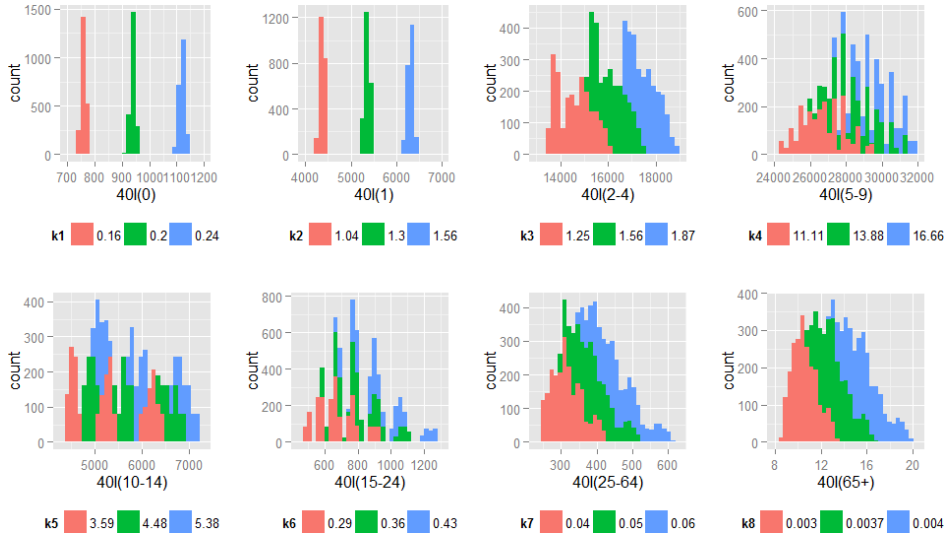

FIG. 5.1. *The sensitivity analysis of selected combination of* $40I_a$ *with respect to* k_i*. Histograms are shown only for the most important combinations of* k_i *with appropriate age group. All simulations were used.*

TABLE 5.2
Average, minimum and maximum values of $40S_a$ *in appropriate age groups of across all simulations.*

	0	1	2-4	5-9	10-14	15-24	25-64	65+	overall
average	98%	89%	62%	15%	6%	5%	4%	4%	9%
max	99%	91%	68%	19%	8%	7%	5%	5%	11%
min	98%	87%	57%	12%	4%	3%	3%	3%	8%

for Slovakia are not available, but the decreasing prevalence of susceptibility with age is consistent with assumptions. (see Table 5.2), [21]).

6. Sensitivity of the model with respect to initial conditions. For values $S_a(0), I_a(0), R_a(0)$ the following condition is valid:

$$S_a(0) + I_a(0) + R_a(0) = \frac{N_a}{N},$$

where N_a is the size of appropriate age groups, and N is the total population size. The above condition results in:

$$\sum_{a=0}^{a=65+} S_a(0) + I_a(0) + R_a(0) = 1.$$

The parameters connected to the transition between age groups, mortality and natality are set in such a way to be valid during the whole simulation time (in order to preserve size and age structure of the population).

For the best combination of k_i parameters a brief analysis of sensitivity to initial conditions was performed. Since there are many options how initial conditions can be changed, we have analysed only those, that could have a significant impact on the spread of the disease. Values $S_3(0), S_4(0), S_5(0)$ were deviated by 20% up and down, whole vector S was reduced by 10, 20, 30 and 40%, and whole vector I was shifted

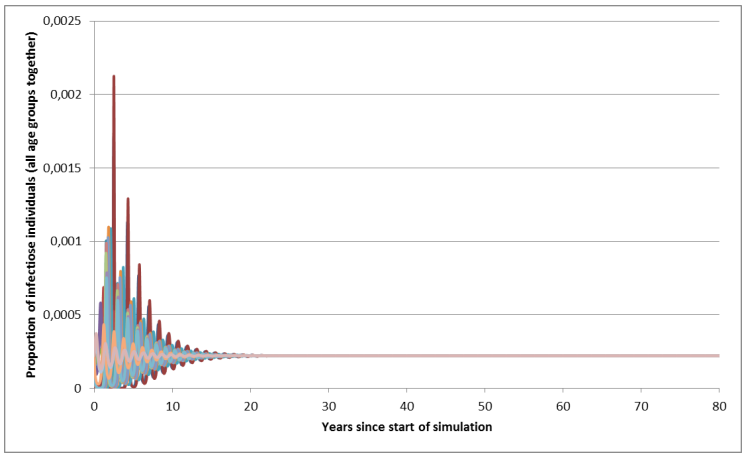

FIG. 6.1. *20 random selected simulations with equal parameters (k_i, P, γ) but different initial conditions.*

by 20% up and down from minimized value. Hence, $3^4 \times 5$ simulations with equal parameters, but different initial conditions were generated.

If initial conditions $(S_a(0), I_a(0), R_a(0))$ are too much different from endemic equilibrium state, damped oscillations in proportion of susceptible, infected and resistant individuals in the beginning of simulation are created (see Figure 6.1). These oscillations disappear with time, and 40 years from the beginning of the simulation are no longer relevant (from practical aspect). Values for $40S_a, 40I_a$ differ only on the level of decimal places. Hence, initial conditions do not affect substantially the behaviour of the model in its final stages.

Conclusion. Varicella is a disease the spread of which in a large population can be described on an acceptable level using deterministic models. Currently, the spread of the disease in Slovakia has not been influenced by vaccination, however, in the future the situation can be changed. The given model appropriately describes transition of the disease and the model population is divided to allow introduction of vaccination of one years old children (along with transition between 0-years and 1-year old age group). This makes it suitable to estimate effects of vaccination of small children on the population. Then it will be possible to observe the output parameters ($40S_a$ and $40I_a$) depending on vaccination rates.

REFERENCES

[1] *EUVAC.NET. Varicella surveillance report 2000-2007. [Online] 2009. [13. 8 2014.]* http://ecdc.europa.eu/en/publications/Publications/varicella_report_2000_2007_euvacnet.pdf
[2] D. DRAŽAN, *Varicella*, Pediatr Praxi. 2007, 8(6), 374-3788. (In Czech.)
[3] H. HUDEČKOVÁ, V. ŠVIHROVÁ, Vaccination. Martin: Osveta; 2013. (In Slovak.)
[4] *European Centre for Disease Prevention and control. Varicella vaccine in the European Union. [Online] 2014. [1. 8 2014.]* www.ecdc.europa.eu/en/publications/Publications/Varicella-guidance-2014-consultation.pdf
[5] *Annual report of the activities of the regional public health authorities in the Slovak Republic*, Bratislava: Public Health Authority of the Slovak Republic; 2012 [cited 2015 Apr 25] (In Slovak). www. uvzsr.sk/index.php?option=com_content&view=category&layout=blog&id=25&Itemid=34

[6] D. Guris, A.O. Jumaan, L. Mascola, B.M. Watson, J.X. Zhang, S.S. Chaves, *et al..*, *Changing varicella epidemiology in active surveillance sites - United States, 1995-2005*, J Infect Dis. 2008 Mar 1;197 Suppl 2:S71-75.
[7] E. Malobická, H. Hudečková, *Varicela*, Pediatria. 2014;9(6):303-5. (In Slovak.)
[8] H. Hudečková, Š. Straka, Š. Rusňáková, *Epidemiological features and economic evaluation of a potential chickenpox vaccination strategy in Slovak Republic*, Cent Eur J Public Health. 2000, Zv. 8, 4, s. 227-228.
[9] E.M. Halloran, S.L. Cochi, T.A. Lieu, M. Wharton, L. Fehrs, *Theoretical Epidemiologic and Morbidity Effects of Routine Varicella Immunization of Preschool Children in the United States*, Am. J. Epidemiol. 1994; 140(2): 81-104.
[10] J. Zibolenová, V. Szabóová, T. Baška, D. Ševčovič, H. Hudečková, *Mathematical modeling of varicella spread in Slovakia*, Cent Eur J Public Health. 2015; 23 (3): 227 - 232.
[11] M.J. Keeling, P. Rohani, *Modelling infectious diseases in humans and animals*, Princeton: Princeton University Press; 2008.
[12] E. Vynnycky, R. White, *An introduction to infectious disease modelling*, New York: Oxford University Press; 2010.
[13] S. Mishra, D.N. Fisman, M.C. Boily, *The ABC of terms used in mathematical models of infectious diseases*, J. Epidemiol Community Health. 2011;65:87-94.
[14] Z. Chladná, E. Moltchanova, *Incentive to vaccinate: a synthesis of two approaches*, Acta Math. Univ. Comenianae. 2015; 84(2): 283-296
[15] J. Zibolenová, D. Ševčovič, T. Baška, D. Rošková, E. Malobická, V. Szabóová, V. Švihrová, H. Hudečková, *Mathematical modeling of infectious childhood diseases*, Česko-Slovenská pediatrie, 2015; 70(4), 210-214.
[16] M. Brisson, W.J. Edmunds, N.J. Gay, B. Law, G. De Serres, *Modelling the impact of immunization on the epidemiology of varicella zoster virus*, Epidemiol Infect. 2000 Dec;125(3):651-69.
[17] M. Brisson, W.J. Edmunds, *Varicella vaccination in England and Wales: cost-utility analysis*, Arch Dis Child. 2003 Oct;88(10):862-9.
[18] M. Brisson, W.J. Edmunds, B. Law, *et al..*, *Epidemiology of varicella zoster virus infection in Canada and the United Kingdom*, Epidemiology and Infection. 2001;127(2):305-314.
[19] *Epidemiological Information System (EPIS) [Internet]. Bratislava: Public Health Authority of the Slovak Republic; 2015 [cited 2015 Apr 25]* (In Slovak.)
http://www.epis.sk/?lang=sk-SK
[20] P. Bakoss, *Epidemiology*, Bratislava: Univerzita Komenskho Bratislava; 2005 (In Slovak).
[21] A. Nardone, F. de Ory, M. Carton, D. Cohen, P. van Damme, I. Davidkin, *et al..*, *The comparative sero-epidemiology of varicela zoster virus in 11 countries in the European region*, Vaccine. 2007 Nov 7;25(45):7866-72