Rubella, age-structure, serological profile, R_0, vaccination.

David GREENHALGH[*], Nikolaos SFIKAS[**]

MATHEMATICAL MODELLING OF UK RUBELLA VACCINATION PROGRAMS

In this article we discuss mathematical modelling of vaccination programs for rubella in the UK. We briefly discuss rubella before outlining the underlying mathematical model. Age-structured serological data is used to estimate the force of infection in the absence of vaccination and hence the mixing matrix. Homogeneous, proportional and symmetric mixing are considered. The estimated mixing matrix is used to evaluate the basic reproduction number R_0 and minimum elimination vaccination programs using one stage and two stage vaccination strategies.

1. INTRODUCTION

Rubella is a mild febrile disease with a diffuse punctuate and produces a rash which has characteristics inbetween those of a macula and a papule. The rash may resemble that of measles or scarlet fever. However up to half the infections occur without evident rash. A diminution of the number of leucocytes normally present in blood is common and thrombocytopenia, a reduction in the number of platelets present in blood can occur with rare haemorrhaging [4,10]. Encephalitis can happen rarely. The most important aspect of rubella is its ability to produce abnormalities in the developing fetus. Congenital rubella syndrome (C.R.S.) occurs in at least 25% of infants born to women who acquire rubella during the first trimester of pregnancy. C.R.S. can have unpleasant side effects such as blindness or deafness in the child. In this paper we shall use mathematical models to evaluate rubella vaccination programs. We are particularly interested in one stage and two stage vaccination programs which vaccinate a given proportion of susceptibles at one or two fixed ages respectively.

[*] Department of Statistics and Modelling Science, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH, U.K.

[**] Novartis Pharma AG, Basel, Switzerland.

2. MATHEMATICAL MODEL

The basic mathematical model is that of Anderson and May [1,2], Dietz and Schenzle [5] and Greenhalgh and Dietz [8]. We are interested here in finding the impact that one and two stage vaccination programs would have and in particular the minimum proportions of susceptibles who must be vaccinated at different ages in order to eliminate rubella in the UK. We are going to investigate the effect of different mixing assumptions, in particular homogeneous, proportional and symmetric mixing. We should also bear in mind that vaccination policies can have serious implications as far as the overall incidence of C.R.S. is concerned (Anderson and May, [1,3]). So we must take great care in evaluating the effects of these vaccination strategies.

The population is divided into classes of susceptible, infected and immune individuals. Every individual starts off susceptible, at some stage catches the disease and after a short infectious period becomes permanently immune. Age-structured partial differential equations are used to model the spread of the disease. $x(a,t)$ denotes the density with respect to age of the number of susceptible individuals at time t. Hence the absolute number of susceptibles between ages A_1 and A_2 at time t is

$$\int_{A_1}^{A_2} x(t,a)da.$$

$y(t,a,c)$ is the density with respect to age a, and duration of infection c, of the number of infected individuals at time t.

$$\int_{A_1}^{A_2}\int_{c_1}^{c_2} y(t,a,c)dadc$$

Thus is the number of infecteds at time t who are aged between A_1 and A_2 and have durations of infection between c_1 and c_2.

The rate at which a susceptible of age a makes potentially infectious contacts (in other words a contact which if between a susceptible and an infected individual would cause infection) is $\beta(a,a') = kb(a,a')/N$. N is the total population size and k is a normalised contact rate. The per capita rate of acquisition of infection of a single susceptible individual of age a at time t is called the force of infection and is given by

$$\lambda(t,a) = \int_0^L \int_0^{a'} \beta(a,a')y(t,a,c)dcda'.$$

The spread of the disease is described by the following partial differential equations (Dietz and Schenzle, [5])

$$\frac{\partial x}{\partial a} + \frac{\partial x}{\partial t} = -[\lambda(t,a) + \phi(a) + \mu(a)]x(t,a) \qquad (1)$$

and $$\frac{\partial y}{\partial a}+\frac{\partial y}{\partial t}+\frac{\partial y}{\partial c}=-[\gamma(c)+\mu(a)]y(t,a,c), \qquad (2)$$

where $x(t,0) = \nu$, $y(t,0,c) = 0$ and $y(t,a,0) = \lambda(t,a)x(t,a)$. ν is the (constant) total birth rate, $\phi(a)$ is the age-dependent vaccination rate, $\gamma(c)$ is the rate at which infected individuals who have been infected for time c enter the immune class and $\mu(a)$ is the age-dependent death rate.

The population is divided into n disjoint age classes $I_1, I_2, \ldots I_n$ and for $a \in I_i$, $a' \in I_j$, $\beta(a,a') = \beta_{ij}$. The matrix β_{ij} is called the who-acquires-infection-from-whom, WAIFW, matrix. $\hat{\lambda}_0(a)$, the force of infection in the absence of vaccination, is estimated from the age-serological profile using the non-parametric maximum likelihood method given in Keiding [9]. As, at least prior to the start of vaccination, rubella is a disease of childhood, there are a much greater number of observed cases at relatively small ages (i.e. 0-5 years) than larger ones (i.e. adult cases). This fact means that if we use a constant kernel smoothing bandwidth then the quantities such as the estimated force of infection are much more reliable at smaller ages than larger ones. We ensure that our estimates are more equally reliable across the whole age range by using a variable smoothing bandwidth which is small at small ages and large at large ages. Following Keiding [9] we use the Epanechnikov kernel, but use a truncated Epanechnikov kernel at the ends of the age range.

Once $\hat{\beta}_{ij}$ has been estimated it is then used with the equilibrium versions of equations (1) and (2) to estimate β_{ij} by $\hat{\beta}_{ij}$. However we have n linear equations in n^2 unknowns and need to make some assumptions on (β_{ij}) to reduce the number of unknowns to n. For a mixing assumption to be feasible we must have $\hat{\beta}_{ij} \geq 0$ for all i,j. In this paper we shall consider homogeneous mixing ($\beta_{ij} = \beta$ for $i,j = 1,2, \ldots n$), proportional mixing ($\beta_{ij} = p_i p_j$ for $i,j = 1,2, \ldots n$) and symmetric mixing ($\beta_{ij} = \beta_{ji}$ for $i,j = 1,2, \ldots n$).

AGE (YEARS)	SEROPO-SITIVE	TESTED	AGE (YEARS)	SEROPO-SITIVE	TESTED
1	25	206	47	14	15
2	25	145	48	13	15
3	31	168	49	23	23
4	54	188	50	14	16
5	92	218	51	13	13
6	98	194	52	11	11
7	86	164	53	14	15
8	91	145	54	15	15
9	134	180	55	15	16
10	108	160	56	8	8
11	108	148	57	12	12
12	145	178	58	16	18
13	137	176	59	9	9
14	139	165	60	3	5
15	50	67	61	6	6
16	45	58	62	12	14
17	72	81	63	11	11
18	67	79	64	6	6
19	95	111	65	13	15
20	63	76	66	11	11
21	72	82	67	2	3
22	84	101	68	4	4
23	80	88	69	3	5
24	77	85	70	5	5
25	89	94	71	8	9
26	84	91	72	4	4
27	81	89	73	4	4
28	72	76	74	5	6
29	71	79	75	6	6
30	50	56	76	9	9
31	44	52	77	4	4
32	45	48	78	5	5
33	35	37	79	4	4
34	39	41	80	3	4
35	34	40	81	7	7
36	37	38	82	4	4
37	36	39	83	3	4
38	36	41	84	1	1
39	27	30	85	2	2
40	26	27	87	1	2
41	25	25	91	2	2
42	21	22	94	1	1
43	18	19	98	1	1
44	18	18	99	1	1
45	16	17			
46	17	17			

Tab.1. Serological data for rubella, showing the age of the individuals, the number who were found to have experienced the disease and the number of people who were tested respectively. (Data taken from Farrington, [7].)

R_ϕ, the reproduction number under steady-state vaccination effort ϕ, is defined as the expected number of secondary infections with constant vaccination effort ϕ due to a single infected

individual entering the population at the disease-free equilibrium. We expect the disease to take off if $R_\phi > 1$ and die out if $R_\phi \leq 1$. If we define

$$A_i^* = \int_{I_i} e^{-\Phi(\xi)-M(\xi)} d\xi,$$

then R_ϕ is the spectral radius of the *n*x*n* matrix ($\hat{\beta}_{\phi,ij}$) where

$$\hat{\beta}_{\phi,ij} = \frac{\beta_{ij} k \nu D}{N} A_j^*(\phi)$$

[8]. The basic reproduction number R_0 is R_ϕ when ϕ=0.

Once the WAIFW matrix has been obtained it is used to estimate R_0 and R_ϕ. R_ϕ will allow us to evaluate a given vaccination campaign. We will look at the practically relevant situations of a one stage vaccination campaign, where a given proportion of susceptible individuals are vaccinated at a fixed age, and a two stage vaccination campaign where given proportions of individuals are vaccinated at two fixed ages. A two stage vaccination campaign allows coverage of those individuals missed by the first vaccination. Further details are given in Greenhalgh and Dietz [8].

3. NUMERICAL RESULTS

We used age-structured serological data provided to us by Farrington [7]. This is shown in Table 1 and consists of a large sample of males tested for rubella and gives the number seropositive at each age. We do not have any serological data for England and Wales prior to the start of the vaccination of women, so we have to use the men only to calculate the age-serological profile in the absence of vaccination. This is not perfectly correct as the immunisation of women will influence the force of infection and thus indirectly affect the age-serological profile. But the women who were immunised before our age-serological profile was collected were women around fifteen years of age. This was done to vaccinate the few remaining susceptible women before they entered the child-bearing age-classes. Hence the vast majority of women experienced the disease in childhood and so achieved natural immunity. Thus although a significant percentage of women were immunised, the majority of these were naturally immune prior to vaccination. So the influence of these immunisations on both the force of infection and the male age-serological profile is very small. Hence it is reasonable to treat the male age-serological profile as though it were the age serological profile in the absence of vaccination and use it to evaluate immunisation programs. Data on age-related mortality rates in England and Wales were taken from Preston, Keyfitz and Schoen [11].

3.1. HOMOGENEOUS MIXING

R_0 is an important epidemiological measure. It must exceed one as we know that the disease persists in the absence of vaccination. A disease such as measles where R_0 is large will spread

quickly and require a very high vaccination coverage to eliminate it. Calculation of R_0 tells us how quickly the disease spreads in the absence of vaccination. The following values for R_0 were obtained.

Bandwidth	Age at division	Estimated R_0
5 and 15	25 years (Case A)	4.051
5 and 15	15 years (Case B)	4.121
5 and 25	25 years (Case C)	3.914

Tab.2. Estimated value of R_0 for variable bandwidths used.

Figure 1 gives the minimum elimination vaccination proportions for the variable bandwidth case B.

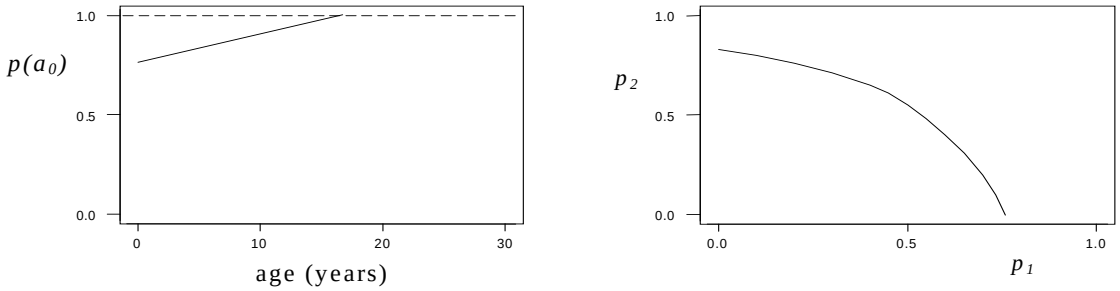

Fig.1. Homogeneous mixing. Variable bandwidth with b_1 = 5 years and b_2 = 15 years (Case B). Estimated minimum elimination coverage proportions (*a*) $p(a_0)$, assuming vaccination at a fixed age a_0 and (*b*) p_2 at age A_2 = 5 years given a coverage p_1 at age A_1 = 2 years.

3.2. PROPORTIONAL MIXING

Table 3 gives the values for R_0 which were obtained. Figure 2 gives the minimum elimination vaccination proportions for the variable bandwidth case B.

3.3. SYMMETRIC MIXING

The last mixing assumption that we are going to examine is symmetric mixing. As we have already mentioned one of the difficulties with symmetric mixing is that we must make assumptions to reduce the number of elements in the who-acquires-infection-from-whom-matrix (WAIFW) matrix from n^2 to n. It is sometimes difficult to decide what assumptions to make for the WAIFW

Bandwidth	Age at division	Estimated R_0
5 and 15	25 years (Case A)	3.421
5 and 15	15 years (Case B)	3.659
5 and 25	25 years (Case C)	3.101

Tab.3. Estimated value of R_0 for variable bandwidths used.

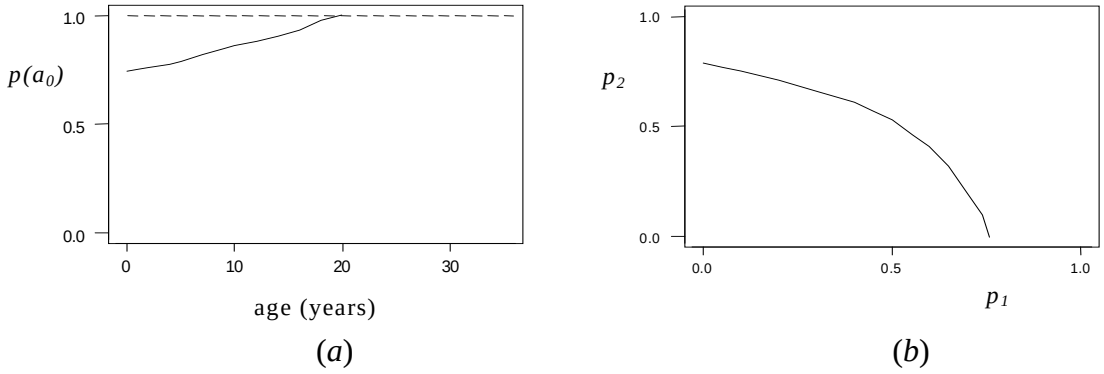

Fig.2. Proportional mixing. Variable bandwidth with b_1 = 5 years and b_2 = 15 years (Case B). Estimated minimum elimination coverage proportions (*a*) $p(a_0)$, assuming vaccination at a fixed age a_0 and (*b*) p_2 at age A_2 = 5 years given a coverage p_1 at age A_1 = 2 years.

matrix to remain feasible. The first priority is to determine matrices which give feasible results and can be motivated by biological considerations. We examined the following matrices based on previous work by Anderson and May [3] and Greenhalgh and Dietz [8].

Matrix A

$$\begin{pmatrix} \beta_1 & \beta_1 & \beta_3 & \beta_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_3 & \beta_3 & \beta_3 & \beta_4 \\ \beta_4 & \beta_4 & \beta_4 & \beta_4 \end{pmatrix}$$

Matrix B

$$\begin{pmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_2 \\ \beta_2 & \beta_2 & \beta_2 & \beta_2 \\ \beta_3 & \beta_2 & \beta_4 & \beta_4 \\ \beta_2 & \beta_2 & \beta_4 & \beta_4 \end{pmatrix}$$

Matrix C

$$\begin{pmatrix} \beta_1 & \beta_1 & \beta_1 & \beta_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_1 & \beta_3 & \beta_3 & \beta_4 \\ \beta_4 & \beta_4 & \beta_4 & \beta_4 \end{pmatrix}$$

Matrix D

$$\begin{pmatrix} \beta_1 & \beta_1 & \beta_1 & \beta_1 \\ \beta_1 & \beta_2 & \beta_2 & \beta_2 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_1 & \beta_2 & \beta_4 & \beta_4 \end{pmatrix}$$

Matrix E

$$\begin{pmatrix} \beta_1 & \beta_4 & \beta_4 & \beta_4 \\ \beta_4 & \beta_2 & \beta_4 & \beta_4 \\ \beta_4 & \beta_4 & \beta_3 & \beta_4 \\ \beta_4 & \beta_4 & \beta_4 & \beta_4 \end{pmatrix}$$

Matrix A has high transmission within the second age category. In Matrix B there is a high level of transmission both from contacts within the second age category and from contacts between this age category and other age categories. β_2 is the corresponding disease transmission coefficient. Matrix C is a variation on Matrix A and Matrix D is the reverse pattern of transmission than Matrix A. Matrix E is a special configuration where the transmission is high within each of the first three age classes but not between age classes. This is intended to model the spread of common childhood diseases among school children which spread predominantly among children of the same age groups.

(*a*)

b_1/b_2	Matrix A	Matrix B	Matrix C	Matrix D	Matrix E
5/5	2.600	**	3.800	**	2.701
15/15	3.375	**	3.560	**	3.457
5/15	2.906	**	3.925	**	3.214

(*b*)

b_1/b_2	Matrix A	Matrix B	Matrix C	Matrix D	Matrix E
5/5	2.678	**	3.450	**	2.665
15/15	3.331	**	3.381	**	2.852
5/15	2.987	**	3.501	**	2.714

Tab.4. Value of the basic reproduction number R_0 for the cases of a bandwidth of five, fifteen years and a variable bandwidth of 5 years up to the age of 15 years and of 15 years thereafter. (*a*) Age class division 1-5, 6-10, 11-15 and 16-99 and (*b*) Age class division 1-7, 8-12, 13-20 and 21-99. The notation '**' means that there was at least one negative element in this estimated matrix which made the configuration infeasible.

We obtained the following results which are shown in Table 4. In Figure 3 we give the minimum elimination vaccination proportions for the one and two stage vaccination campaigns when considering the configuration of Matrix A for a constant bandwidth of 5 years and the age division of Table 4(*a*). We decided to present this case only, because this is the matrix, age division and bandwidth that gave the lowest value for R_0 and the highest value for the minimum elimination vaccination proportions. So we are particularly interested in this worst possible case as if we vaccinate these proportions of susceptible individuals we can be reasonably certain to eliminate rubella in the UK.

4. SUMMARY AND CONCLUSIONS

In this article we have used mathematical models to evaluate rubella vaccination programs in the UK. The basic reproduction number R_0 is an important epidemiological quantity and gives an estimate of how fast the disease will spread in the absence of vaccination. Starting with an age-serological profile we estimated both R_0 and minimum elimination vaccination proportions for one stage and two stage immunisation strategies. Future work will use age-structured serological data to similarly evaluate vaccination programs for mumps in the UK and hepatitis A in Bulgaria and use the bootstrap method to estimate confidence and percentile intervals for the estimated epidemiological parameters.

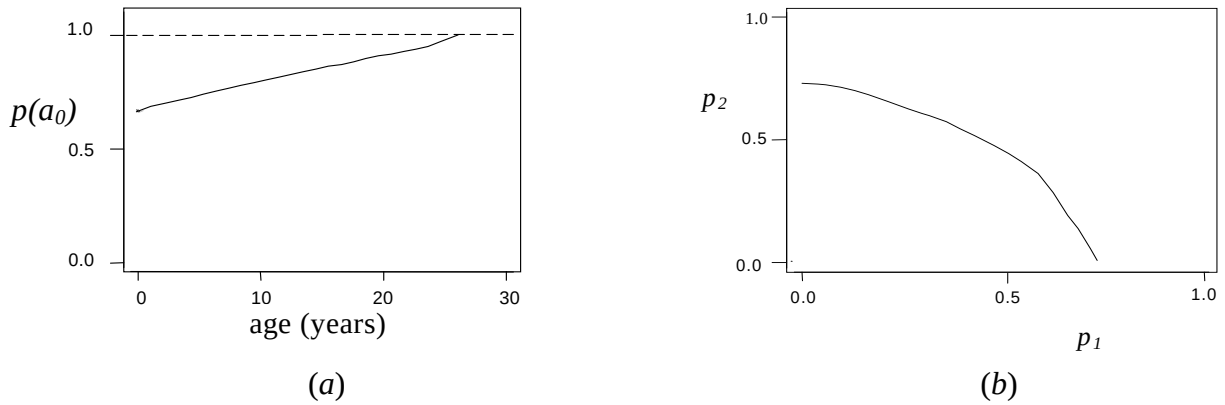

Fig.3. Symmetric mixing. Matrix A and age division 0 – 5, 6 – 10, 11 – 15, 16 – 99 years. Estimated minimum elimination coverage proportions (*a*) $p(a_0)$, assuming vaccination at a fixed age a_0 and (*b*) p_2 at age $A_2 = 5$ years given a coverage p_1 at age $A_1 = 2$ years.

BIBLIOGRAPHY

[1] ANDERSON, R.M. and MAY, R.M., Vaccination against rubella and measles: quantitative investigations of different policies, Cambridge Journal of Hygiene, Vol. 90, pp.259-325, 1983.
[2] ANDERSON, R.M. and MAY, R.M., Age-related changes in the rate of disease transmission: implication for the design of vaccination programmes, Cambridge Journal of Hygiene, Vol. 94, pp.365-436, 1985.
[3] ANDERSON, R.M. and MAY, R.M., Infectious diseases of humans: dynamics and control, Oxford University Press, Oxford, 1991.
[4] BENENSON, A.S., Control of communicable diseases in man, Sixteenth Edition, American Public Health Association, Washington D.C., 1990.
[5] DIETZ, K. and SCHENZLE, D., Proportionate mixing for age-dependent disease transmission, Journal of Mathematical Biology, Vol. 22, pp.117-120, 1985.
[6] FARRINGTON, C.P., Modelling forces of infection for measles, mumps and rubella, Statistics in Medicine, Vol. 9, pp.953-967, 1990.
[7] FARRINGTON, C.P., Private Communication, 1995.
[8] GREENHALGH, D. and DIETZ, K., Some bounds on estimates for reproductive ratios derived from the age-specific force of infection. Mathematical Biosciences, Vol. 124, pp.9-57, 1994.
[9] KEIDING, N., Age-specific incidence and prevalence: a statistical perspective, Journal of the Royal Statistical Society, Series A, Vol. 154, pp.371-412, 1991.
[10] MACNALTY, A.S., Butterworths Medical Dictionary, Butterworth, 1965.
[11] PRESTON, S.H., KEYFITZ, N. and SCHOEN, R., Causes of Death, Lifetables for National Populations. Studies in Population Series, Seminar Press, London/New York, 1972.