Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

The structural identifiability of SIR type epidemic models with incomplete immunity and birth targeted vaccination ⋆

J. D. Chapman * **N. D. Evans** *

* *School of Engineering, University of Warwick, Coventry, CV4 7AL (e-mail: james.chapman@warwick.ac.uk).*

Abstract: This paper considers the implications of a structural identifiability analysis on a series of fundamental three-compartment epidemic model structures, derived around the general SIR (Susceptible-Infective-Recovered) framework. The models represent various forms of incomplete immunity acquired through natural infection, or from administration of a birth targeted vaccination programme. It is shown that the addition of a vaccination campaign has a negative effect on the structural identifiability of all considered models. In particular, the actual proportion of vaccination coverage achieved, an essential parameter, cannot be uniquely estimated from even ideal prevalence data.

Keywords: Structural Identifiability; SIR Models; Incomplete Immunity; Vaccination.

1. INTRODUCTION

The application of mathematical modelling techniques within the context of communicable disease epidemiology has been motivated by the need for more accurate predictions of the outcome of community infection, to achieve better understanding of the underlying biological mechanisms at work and to aid in the development of optimal strategies for control by public health intervention.

Postulated models of epidemic systems are often fitted to time series prevalence or incidence data of infection, usually from observations of clinical disease. However an important, and often overlooked prerequisite to interpreting results from parameter estimation and inferring characteristics of the real system is the consideration of structural identifiability. That is the theoretical problem of determining whether alternative, indistinguishable parameterisations of a given model exist, that give rise to identical input/output behaviour. Should the model structure prove to be unidentifiable, whereby a number of model parameters are not uniquely determined by the measurable behaviour of the system, parameter values estimated from fitting to real data should be treated with caution.

The most common approach to deterministic modelling of population level infection dynamics is through a compartmental representation of the various stages of the natural history of infection, written as a system of ordinary differential equations. This method was first developed by Kermack and McKendrick [1927], and was intended to approximate epidemic evolution within large constant size populations (for general examples see Jacquez [1996], Capasso [1993], and for specific examples see Weber et al. [2001], White et al. [2006] for human respiratory syncytial virus and Keeling and Grenfell [2002] for measles). Models of this type tend to be uncontrolled (free or autonomous) and nonlinear, where the problem of identifiability is typically approached using a Taylor series expansion of the output (for examples in pharmacokinetics see Godfrey and Fitch [1984]). However even with the relatively simple epidemic models discussed in this work, the technique can quickly become overly computationally intensive.

The paper by Evans et al. [2005] shows how a new general theory for nonlinear structural identifiability, proposed by Evans et al. [2002], can be successfully applied to a general SIR epidemic model. It is the objective of this work to extend the theory's application in this context to models with incomplete immunity and birth targeted vaccination. A systematic analysis of the dynamics and equilibrium properties of the suboptimal immune models considered in this paper can be found in Gomes et al. [2004].

2. SIR FRAMEWORK MODELS

The general SIR model framework can be used to characterise epidemic systems where the natural history of infection can be approximated into three distinct stages. The total population is therefore divided into three non-overlapping classes (*susceptible*, *infective* and *recovered*) representing subpopulations of individuals with a specific state of disease. The susceptible class includes all individuals who are able to contract the disease and become infectious; the infective class represents only individuals who are currently infected and infectious to susceptibles; and the recovered class contains all individuals who have recovered from infection and consequently acquired some form of immunity. Individuals are born into the susceptible class at a net birth rate μN, where N is the total population size. It is assumed that the average duration of infection, $1/v$ is small with respect to the average life expectancy, $1/\mu$, so the net mortality rate $\mu(S(t) + I(t) + R(t))$ can be assumed to equal μN, hence maintaining a constant population size.

⋆ JDC was funded by an Engineering and Physical Sciences Research Council (EPSRC) UK, research studentship.

The mean rate (per host) of contacts that result in disease transmission, β, is a product of the average contact rate between individuals within the population, c, and the probability of transmission upon contact, p. The average rate at which infectives make contacts that transmit disease is $\beta I(t)$, however only the fraction $S(t)/N$ are with susceptible individuals. Therefore the rate at which susceptibles become infected can be modelled as $\beta S(t)I(t)/N$, (where $\lambda = \beta I(t)/N$ represents the force of infection), see Jacquez [1996]. A system of ordinary differential equations can then be defined:

$$\dot{S}(t) = \mu N - \frac{\beta}{N} S(t)I(t) - \mu S(t) + \alpha R(t), \quad (1)$$

$$\dot{I}(t) = \frac{\beta}{N}(S(t) + \sigma R(t))I(t) - (\mu + v)I(t), \quad (2)$$

$$\dot{R}(t) = vI(t) - (\mu + \alpha)R(t) - \sigma\frac{\beta}{N} R(t)I(t), \quad (3)$$

and if necessary reduced to a two state system given that $R(t) = N - S(t) - I(t)$. In the case of the general SIR model (Fig. 1), recovered individuals are no longer able to transmit or re-contract the disease (i.e parameters α and σ associated with incomplete immunity are 0), making the structure most appropriate for modelling MMR (Measles, Mumps and Rubella) type infections where lifelong immunity to the entire pathogen population is induced following recovery from infection.

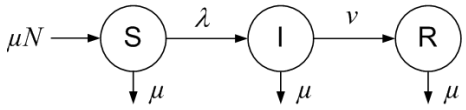

Fig. 1. General SIR Compartmental Model

The *basic reproduction number*, denoted R_0, is often considered to be the most important quantity in mathematical epidemiology of infectious diseases. It is primarily defined as the average number of secondary cases of infection produced by an average primary case in a completely susceptible (naive or virgin) population, and can be expressed as a product of the transmission parameter β, and the average duration of time an individual remains infectious:

$$R_0 = \frac{\beta}{\mu + v}. \quad (4)$$

This definition gives rise to an invasion threshold at $R_0 = 1$ that can be used to determine whether or not an infection will be able to successfully invade and sustain within a given susceptible host population. Any infective case that leads to more than one secondary case, i.e. $R_0 > 1$ has the ability to invade the host population and cause an epidemic, whereas a situation where less than 1 secondary case is produced will lead to the infection fading out and failing to survive. Given that transmission is heavily dependent on social (behavioural) and environmental variables, which differ between demographically and geographically distinct populations, R_0 is unique both for different diseases and different populations within which it is being considered.

2.1 Models With Incomplete Immunity

In reality most pathogens are able to evade natural immunity and re-infect their hosts through either static antigenic variability or antigenic evolution over a given period. The consequence of significant antigenic variability within a pathogen population is that natural immunity acquired through experience of infection serves only to protect an individual against a proportion of the circulating infectious agents, leading to only *partial* immunity within the host population. Similarly, if the pathogen population experiences a rapid rate of antigenic evolution then any acquired immunity following infection will appear only temporarily effective. It should be noted that the extreme case of incomplete immunity is the SIS (Susceptible, Infected, Susceptible) model structure, in which all individual hosts recover directly back into the fully susceptible state and are immediately able to re-contract the disease. SIS type structures are most appropriate for infections such as gonorrhoea, which does not produce immunity against reinfection.

In order to model an epidemic system where immune hosts experience waning of acquired immunity with time since previous infection, the general SIR model is extended to include an additional transfer from compartment R to S, with rate coefficient α, that describes the rate at which recovered (immune) hosts return to being fully susceptible (note that the average duration of infection is then $1/\alpha$) where they are able to re-contract the infection upon contact with an infective individual. It is assumed that before waning off, acquired immunity is solid and provides protection against all current variants of the infectious agent. The model structure is shown in Fig. 2, and the system equations can be derived from (1)-(3) by setting only $\sigma = 0$. A partial immunity model describes the

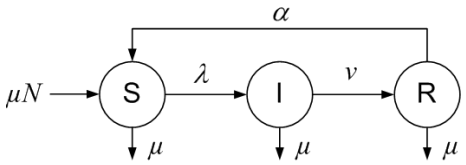

Fig. 2. SIRS Temporary Immunity Model

situation when immunity serves only to protect the individual against a proportion of the pathogen population, and an element of susceptibility to some antigenic variants is always retained. The model can be considered as a

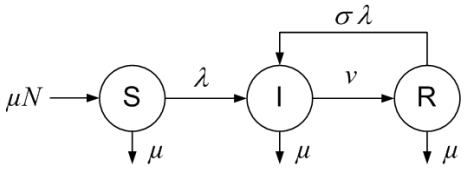

Fig. 3. SIRp Partial Immunity Model

combination of an SIR submodel (primary infection) and an SIS submodel (re-infections), hence the recovered class now acts as a second susceptible class where hosts have a reduced susceptibility to the force of infection. The system equations can be derived from (1)-(3) by setting only $\alpha = 0$, and the structure is shown diagrammatically in Fig. 3. The partial immunity or reduced susceptibility parameter $\sigma \in [0, 1]$, and it is assumed that any acquired immune protection does not wane with time and is not altered by subsequent re-infections. The dynamic behaviour and

endemicity of the system can be considered as an equilibria between the contributing characteristics of the SIR and SIS components of the system. This equilibria is governed by the re-infection threshold, which describes a limit for R_0 with respect to the partial immunity parameter σ, beyond which a significant increase in transmission can occur, see Gomes et al. [2004].

3. VACCINATION MODELS

Vaccination strategies are employed to protect susceptible hosts, both at individual and population level, against parasite infection, and subsequently reduce the prevalence or burden of disease. Vaccines work by presenting a foreign antigen to the immune system in order to evoke a specific immune response with less clinical disease than natural infection, the intention is that subsequent immune stimulation to natural infection is achieved more rapidly, preventing the onset of severe disease. At population level immunisation can be considered as a process of fast-tracking a proportion of susceptible hosts to the recovered state, without experiencing natural infection, severe disease and hence a period of time during which they are infectious to others. The work documented in this paper focuses on the study of models where vaccines produce identical properties of immunity to that of natural infection, i.e mechanisms such as variability and evolution which may allow the infectious agent to evade natural immune protection are also viable for vaccine induced immunity, hence vaccinated hosts are included in the recovered population class $R(t)$.

The most simple form of immunisation program is an untargeted blanket vaccination of a proportion of the total population. In this instance the vaccination is applied to a random selection of individuals in the population regardless of which state (experience of infection) they reside in. This is a particularly inefficient strategy given that some individuals who have already attained a degree of immune protection through natural infection, or who are currently infected, will receive the vaccine to no effect. This approach can often be the only applicable strategy given the frequent difficulties that arise with identifying an individual vaccinee's prior experience of infection (for example in cases of wild animal vaccination). The most

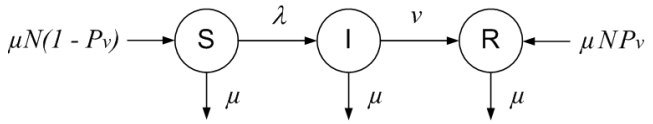

Fig. 4. SIR Model With Birth Targeted Vaccination

common example of a targeted vaccination model is a strategy applied to a proportion of all newborn children. Although in reality vaccination is very rarely administered to babies due to complications with the presence of maternal antibodies, it provides a useful approximation to situations where immunisation is targeted below the average age of infection. This type of program, perhaps applied to a particular age group of young children, is also much more straight forward to implement and monitor within a typical health care infrastructure. A basic vaccination model of this strategy can be implemented on the SIR framework structure (1)-(3) by replacing the inflow of births μN into the susceptible compartment (1), by an inflow of susceptibles $\mu N(1-P_v)$ and an inflow of vaccinees to the recovered compartment (3) $\mu N P_v$, see Fig. 4, where the parameter P_v is the *actual* proportion of newborns that successfully take the vaccine and develop immunity to infection.

Considering the case of disease free equilibrium, where there is no infection prevalent (i.e $S = N$), the completely susceptible *proportion*, $X = S/N$, can be reduced through vaccination to $1 - P_v$ (provided immunity is solid and lifelong i.e in the case of the general SIR). Given that the *effective* reproduction number, $R = XR_0$, is required to be greater than unity for an epidemic to occur, a potential pathogen invasion can be prevented by reducing X to below a critical value $X^* = 1/R_0 = 1-P_v$. The proportion of the population that are required to be immunised in order to prevent an epidemic is then:

$$P_v \ \geq \ 1 - \frac{1}{R_0}, \tag{5}$$

which is also a threshold for the eradication of an infection in an endemic situation. This concept is often referred to as herd immunity, and it shows how population level protection against an infection can be achieved without necessarily vaccinating all individual hosts within it. Equation (5) also shows that diseases with a relatively low R_0 such as smallpox, can be eradicated much more easily than those with a high R_0 such as measles.

3.1 Vaccination and Eradication Thresholds

The invasion threshold, which determines the basic reproduction number R_0, and the minimum vaccination threshold can both be derived from an eigenvalue analysis of the stability of the disease free equilibrium. Given that $P_v \in [0, 1]$ is a proportion, it can be seen from the expression for herd immunity (5) that eradication of an SIR type infection is always possible provided 100% vaccination coverage can be achieved. For the general SIR with birth targeted vaccination (Fig. 4), the disease free steady state of the (reduced two state) system can be derived:

$$\begin{bmatrix} S \\ I \end{bmatrix} = \begin{bmatrix} N(1-P_v) \\ 0 \end{bmatrix}.$$

The influence of vaccination on the stability of the disease free equilibrium can be determined from the corresponding eigenvalues of the system Jacobian matrix:

$$Jacobian = \begin{bmatrix} -(\mu N + \beta I)/N & -\beta S/N \\ \beta I/N & \beta S/N - (\mu + v) \end{bmatrix}$$

where the corresponding eigenvalues for the disease free equilibrium are: $\lambda_1 = -\mu$ and $\lambda_2 = \beta(1 - P_v) - (\mu + v)$. It can then be seen that λ_1 and λ_2 are always real, and from rearranging λ_2, that the disease free steady state can be forced stable (i.e negative real parts) if $P_v \geq 1 - 1/R_0$ hence reiterating the condition for herd immunity (5).

SIRS Temporary/Waning Immunity For the SIRS model (Fig. 2), where acquired immunity serves only to protect individuals for a limited period of time, the system Jacobian matrix, when evaluated at the corresponding disease free equilibrium gives rise to the following eigenvalues:

$$\lambda_1 = -\alpha - \mu \qquad \text{and} \qquad \lambda_2 = \frac{\beta(\mu(1-P_v)+\alpha)}{\alpha+\mu} - (\mu + v).$$

It can be seen that λ_1 is always negative and real, however the stability of λ_2 is also dependent on the waning immunity parameter α. Given that $P_v = 1$ indicates ideal 100% vaccination coverage, a new threshold term can be derived that determines the maximum rate of loss of immunity α before eradication of the infection through vaccination becomes impossible. For eradication to be possible:

$$\alpha \;<\; \frac{\mu}{R_0 - 1}. \tag{6}$$

Provided the temporary immunity parameter does not exceed this threshold, the critical vaccination coverage required to force the disease free steady state to be stable and hence successfully eradicate the infection is found to be:

$$P_v \;=\; \frac{(R_0 - 1)(\alpha + \mu)}{R_0 \mu}. \tag{7}$$

SIRp Partial Immunity For the partial immunity model (Fig. 3), when $P_v = 1$ the model structure can be reduced to that of an SIS given that only secondary infections can occur. This implies that for eradication to be possible, the corresponding reproduction number for the SIS submodel, σR_0, must be less than unity, hence:

$$\sigma \;<\; \frac{1}{R_0}. \tag{8}$$

Provided σ does not exceed this threshold, the critical vaccination coverage required to eradicate the infection can be shown to be:

$$P_v \;=\; \frac{(1 - R_0)}{R_0(\sigma - 1)}. \tag{9}$$

4. STRUCTURAL IDENTIFIABILITY

The rate at which susceptible hosts become infected $\beta S(t)I(t)/N$ is an inherently non-linear term given that it has an explicit dependency on both the susceptible and infective state variables. Therefore in the following sections on structural identifiability all SIR framework models are considered in the following standard form for uncontrolled nonlinear systems:

$$\dot{x}(t,p) = f(x(t,p),p), \tag{10}$$

$$y(t,p) = h(x(t,p),p), \tag{11}$$

$$x(0,p) = x_0(p), \tag{12}$$

where $x(t,p) \in \mathbb{R}^n$ and $y(t,p) \in \mathbb{R}^m$ denote the state variables and the output respectively; $p \in \Omega$ (an open subset of $\mathbb{R}^q$) corresponds to a constant parameter vector, and for all $p \in \Omega$, $f(\cdot,p)$ and $h(\cdot,p)$ are analytic.

Definition 1. Two particular parameter vectors $p, \tilde{p} \in \Omega$ are said to be *indistinguishable* ($\tilde{p} \sim p$) if they give rise to identical output data, $y(t,p) = y(t,\tilde{p})$ for all $t \geq 0$. Hence it is impossible to distinguish between p and $\tilde{p}$ from an ideal noise free observation via the output.

Definition 2. A model conforming to (10)-(12) can be said to be *globally identifiable* at $p \in \Omega$ if $\tilde{p} \sim p$ and $\tilde{p} \in \Omega$ imply that $\tilde{p} = p$, and *locally identifiable* at $p \in \Omega$ if there exists some open neighbourhood $\mathcal{N}$ of p in Ω such that $\tilde{p} \sim p$ for $\tilde{p} \in \mathcal{N}$ implies that $\tilde{p} = p$.

Definition 3. The model is said to be *structurally globally (locally) identifiable* if it can be shown to be globally (locally) identifiable at p, for almost all $p \in \Omega$, (except for a subset of Ω of measure zero). If the model is shown not to be structurally locally identifiable then it is said to be *unidentifiable*.

The following technique used in this paper (see Evans et al. [2002]), utilises the existence of a smooth mapping, $\lambda(x)$, that connects the state trajectories of indistinguishable parameter vectors such that:

$$\lambda(x(t,\tilde{p})) = x(t,p).$$

The Lie derivative of $h \in C^\infty(M(p))$ along the vector field f is the smooth function given by

$$L_f h(x) = \frac{\partial h}{\partial x}(x) f(x).$$

Let $f^p = f(\cdot,p)$ and, for some $1 \leq l \leq m$, $h_l^p(\cdot) = h_l(\cdot,p)$. For any n smooth functions $u_1(x,p), \ldots, u_n(x,p)$ of the form $h_j(x,p)$, for some j, or $L_{f^p}^r h_l^p(x)$, for some l and r, a function H can be defined by

$$H(x,p) = (u_1(x,p), \ldots, u_n(x,p))^T$$

where for a particular $p \in \Omega$, H_p denotes the vector field $H(\cdot,p)$. The system (10)-(12) can be shown to satisfy the Observability Rank Criterion (ORC) at the initial condition $x_0(p)$ if a function H exists such that the Jacobian matrix of H_p, evaluated at $x_0(p)$, is nonsingular (Hermann and Krener [1977]).

Theorem 4. Given the system satisfies the ORC at the initial condition for a particular parameter vector $p \in \Omega$, then $\tilde{p} \in \Omega$ is indistinguishable, $\tilde{p} \sim p$, provided a $\tau(p) > 0$, an open neighbourhood $V_{\tilde{p}}$ of $x_0(\tilde{p})$ and a smooth mapping $\lambda : V_{\tilde{p}} \to \lambda(V_{\tilde{p}})$ exist, such that

$$H_p(\lambda(x)) = H_{\tilde{p}}(x) \tag{13}$$

for all $x \in V_{\tilde{p}}$, and

$$\lambda(x_0(\tilde{p})) = x_0(p), \tag{14}$$

$$f^p(\lambda(x(t,\tilde{p})),p) = \frac{\partial \lambda}{\partial x}(x(t,\tilde{p})) f^{\tilde{p}}(x(t,\tilde{p}),\tilde{p}), \tag{15}$$

$$h^p(\lambda(x(t,\tilde{p})),p) = h^{\tilde{p}}(x(t,\tilde{p})\tilde{p}) \tag{16}$$

for all $t \in [0, \tau(p))$ with $x(t,\tilde{p}) \in V_{\tilde{p}}$, where $x(t,\tilde{p})$ is the solution of the system for parameter vector $\tilde{p}$.

For full proof of *Theorem 4* see Evans et al. [2002].

5. IDENTIFIABILITY ANALYSIS

It has been shown by Evans et al. [2005] that a general SIR model is unidentifiable given that the parameters k, N, S_0 and I_0 (hence β/N) are not uniquely determined by an output structure corresponding to a prevalence observation of infection. However it was also shown that the parameters μ, v and β are uniquely determined by the output and hence globally identifiable.

Although this analysis shows that it is inappropriate to use this model and output structure to estimate the

proportion of contacts between infective and susceptible individuals that result in infection β/N, it can be used to uniquely determine the basic reproduction number R_0 (4), and hence from (5) also the proportion of vaccination coverage required to eradicate the infection and provide herd immunity against future epidemics.

5.1 SIR With Birth Targeted Vaccination

The model shown in Fig. 4 is reduced to a two state problem using $R(t) = N - S(t) - I(t)$, and expressed in the form given by (10)-(12):

$$f(x(t,p),p) = \begin{pmatrix} \mu N(1-P_v) - \beta S(t,p)I(t,p)/N - \mu S(t,p) \\ \beta S(t,p)I(t,p)/N - (\mu+v)I(t,p) \end{pmatrix}$$
$$h(x(t,p),p) = kI(t,p)$$
$$x_0(p) = (S_0 \;\; I_0)^T,$$

where $x(t,p) = (S(t,p), I(t,p))^T$, the output structure, $kI(t,p)$, corresponds to an observation (with unknown gain k) of the prevalence of infection; S_0 and I_0 are the respective initial conditions for the susceptible and infective states; and $p = (\mu, N, P_v, \beta, v, k)^T$ is a vector of the model parameters (assumed to be indistinguishable from $\tilde{p} = (\tilde{\mu}, \tilde{N}, \tilde{P_v}, \tilde{\beta}, \tilde{v}, \tilde{k})^T$).

It was found that the first stages of the analysis regarding the ORC and the generation of the smooth mapping λ were the same as the unvaccinated SIR, as in Evans et al. [2005]. The smooth mapping $\lambda = (\lambda_1 \; \lambda_2)^T$ given by:

$$\lambda(x) = \left(\frac{N(\tilde{N}(\mu + v - (\tilde{\mu} + \tilde{v})) + \tilde{\beta}S)}{\tilde{N}\beta}, \frac{\tilde{k}I}{k} \right)^T \qquad (17)$$

where $x = (S, I)^T$, can be seen to automatically satisfy Theorem 4 (13) and (16), thus in order to satisfy (15) it is necessary that

$$\begin{pmatrix} \mu N(1-P_v) - \beta\lambda_1\lambda_2/N - \mu\lambda_1 \\ \lambda_2(\beta\lambda_1 - (\mu+v)) \end{pmatrix} = \begin{pmatrix} (N\tilde{\beta}/\tilde{N}\beta)(\tilde{\mu}\tilde{N}(1-\tilde{P_v}) - \tilde{\beta}SI/\tilde{N} - \tilde{\mu}S) \\ (\tilde{k}/k)I(\tilde{\beta}S - \tilde{N}(\tilde{\mu}+\tilde{v})) \end{pmatrix}$$

It can subsequently be seen that the second row of (15) is automatically equal and the resulting expression from the first component can be rearranged into the following multivariate polynomial form:

$$q_1 + q_2 S(t,\tilde{p}) + q_3 I(t,\tilde{p}) + q_4 S(t,\tilde{p})I(t,\tilde{p}) = 0. \qquad (18)$$

It can be shown, by solving a series of simultaneous equations derived by setting all $t = 0$ in (18) and successively differentiating with respect to t, that the only solution to (18) is $q_1 = q_2 = q_3 = q_4 = 0$. Each of the four coefficients $q_i = 0$ for $i = \{1, 2, 3, 4\}$ can then be solved along with (14) to give all possible conditions for $\tilde{p} \sim p$:

$$\{\tilde{\mu} = \mu, \;\; \tilde{v} = v, \;\; \tilde{\beta}(\tilde{P_v} - 1) = \beta(P_v - 1), \;\; \tilde{k}\tilde{S_0} = kS_0, \;\; \tilde{k}\tilde{I_0} = kI_0, \;\; \frac{\tilde{k}\tilde{N}}{\tilde{\beta}} = \frac{kN}{\beta}\}$$

It can be seen that the parameters μ and v are still globally identifiable, however the effective coverage of the applied vaccination P_v can not be uniquely determined from the output. Further to this, the basic reproduction number R_0 is no longer uniquely identifiable following vaccination, given that with $P_v > 0$ the transmission parameter β is not uniquely determined by the output.

5.2 SIRS Temporary/Waning Immunity

The SIRS model (Fig. 2) is reduced and expressed in the form given by (10)-(12), in a similar manner to the previous example, where the output structure $y(t,p) = kI(t,p)$, initial conditions S_0 and I_0, and parameter vector $p = (\mu, N, \alpha, \beta, v, k)^T$ (assumed indistinguishable from $\tilde{p} = (\tilde{\mu}, \tilde{N}, \tilde{\alpha}, \tilde{\beta}, \tilde{v}, \tilde{k})^T$). It was again found that the first stages of the analysis regarding the ORC and the generation of the smooth mapping λ were consistent with the SIR models, where the smooth mapping λ, as in (17), is known to automatically satisfy Theorem 4. (13) and (16). In order to satisfy (15) it is necessary that:

$$\begin{pmatrix} \mu N - \beta\lambda_1\lambda_2/N - \mu\lambda_1 + \alpha(N - \lambda_1 - \lambda_2) \\ \lambda_2(\beta\lambda_1 - (\mu+v)) \end{pmatrix} = \begin{pmatrix} (N\tilde{\beta}/\tilde{N}\beta)(\tilde{\mu}\tilde{N} - \tilde{\beta}SI/\tilde{N} - \tilde{\mu}S + \tilde{\alpha}(\tilde{N} - S - I)) \\ (\tilde{k}/k)I(\tilde{\beta}S - \tilde{N}(\tilde{\mu} + \tilde{v})) \end{pmatrix}$$

which along with (14) gives all possible conditions for $\tilde{p} \sim p$

$$\{\tilde{v} = v, \;\; \tilde{\mu} - \tilde{\beta} = \mu - \beta, \;\; \tilde{\alpha} + \tilde{\beta} = \alpha + \beta, \;\; \tilde{k}\tilde{S_0} + \tilde{\alpha} = kS_0 + \alpha, \;\; \tilde{k}\tilde{I_0} = kI_0, \;\; \frac{\tilde{k}\tilde{N}}{\tilde{\beta}} = \frac{kN}{\beta}\}$$

The result shows that only the recovery rate coefficient v is globally identifiable (except also the combinations $\mu - \beta$, $\alpha + \beta$ and $\alpha + \mu$). The consequence of this outcome is that neither parameters associated with R_0 or the duration of immunity can be uniquely determined from the model output. This implies that it is impossible to determine what level of vaccine coverage would be required to eradicate an infection of this type (7), or even if such an outcome were possible (6).

Following the addition of vaccination:

$$\{\tilde{v} = v, \;\; \tilde{\mu} + \tilde{\alpha} = \mu + \alpha, \;\; \frac{\tilde{k}\tilde{N}}{\tilde{\beta}} = \frac{kN}{\beta}, \;\; \frac{\tilde{\beta}\tilde{S_0}}{\tilde{N}} + \tilde{\alpha} = \frac{\beta S_0}{N} + \alpha, \;\; \tilde{k}\tilde{I_0} = kI_0, \;\; \tilde{\mu}\tilde{\beta}\tilde{P_v} + (\tilde{\mu} - \tilde{\beta})(\tilde{\mu} + \tilde{\alpha}) = \mu\beta P_v + (\mu - \beta)(\mu + \alpha)\}$$

It can be seen that even with prior knowledge of the parameter combination $(\mu - \beta)$, from pre-vaccination model fitting, the vaccine efficacy parameter P_v cannot be globally identified from prevalence data.

5.3 SIRp Partial Immunity

The first step of the analysis is to show that the ORC is satisfied at $x_0(p)$. Let $H_p(x) = (u_1(x,p), \; u_2(x,p))^T$ where

$$u_1(x,p) = h^p(x) = kI$$
$$u_2(x,p) = L_{f^p} u_1^p(x) = kI((S + \sigma(N - S - I))\beta/N - (\mu + v))$$

It can be seen that the ORC is satisfied given that the Jacobian matrix of $H_p(x)$:

$$\begin{pmatrix} 0 & k \\ k\beta I(1-\sigma)/N & k((\beta(S(1-\sigma) - 2\sigma I) - N(\mu + v - \beta\sigma)))/N \end{pmatrix}$$

has full rank for any $p \in \Omega$ and for all $x \in W = \{x \in \mathbb{R}^n : x \neq 0\}$. Given that $x(t,p) \in W$ for all $t \geq 0$ it can be seen from Theorem 4. that if the parameter vectors $p, \tilde{p} \in \Omega$ are indistinguishable $\tilde{p} \sim p$, the open neighbourhood $V_{\tilde{p}}$ of $x_0(\tilde{p})$ exists and the smooth mapping λ is a diffeomorphism on $V_{\tilde{p}}$ onto its range [Evans et al., 2005].

Solving for Theorem 4. (13) yields the following smooth map, which also satisfies (16):

$$\left(\frac{kN(\tilde{N}(\tilde{v}-v+\tilde{\mu}-\mu+\beta\sigma-\tilde{\beta}\tilde{\sigma})+\tilde{\beta}(\tilde{\sigma}-1)S)+(kN\tilde{\beta}\tilde{\sigma}-\tilde{k}\tilde{N}\beta\sigma)I}{k\tilde{N}\beta(\sigma-1)}, \frac{\tilde{k}I}{k}\right)^T$$

Further satisfying equations (15) and (14) yields the outcome of a structural identifiability analysis of an SIRp:

$$\{\tilde{\beta}=\beta,\ \tilde{\mu}=\mu,\ \tilde{v}=v,\ \tilde{\sigma}=\sigma,\ \tilde{k}\tilde{S}_0=kS_0,\ \tilde{k}\tilde{I}_0=kI_0,\ \tilde{k}\tilde{N}=kN\}$$

Similar to the general SIR result, the partial immunity model parameters associated with the basic reproduction number (i.e β, μ, and v) are globally identifiable, as well as the reduced susceptibility parameter σ. This result means that from (8) the possibility of eradication with 100% vaccination coverage can be uniquely determined from the model output, and from (9) the effective proportion of the population required to be immunised can also be determined uniquely. However, following the application of the vaccination programme:

$$\{\tilde{\mu}=\mu,\ \tilde{\sigma}=\sigma,\ \tilde{v}-\tilde{\beta}\tilde{\sigma}=v-\beta\sigma,\ \tilde{k}\tilde{S}_0=kS_0,\ \tilde{k}\tilde{I}_0=kI_0,\ \tilde{\beta}(\tilde{P}_v-1)=\beta(P_v-1),\ \frac{\tilde{k}\tilde{N}}{\tilde{\beta}}=\frac{kN}{\beta}\}$$

It can be seen that P_v is not uniquely determined by the output and R_0 becomes unidentifiable once vaccination has been applied. However it should be noted that if either v, β or kN are known, i.e from fitting to pre-vaccination data, then R_0 and P_v can be uniquely identified.

6. CONCLUSIONS

The key results of this paper are that the actual vaccination coverage, P_v, achieved after employing a birth targeted immunisation campaign on any of the discussed SIR framework models, cannot be uniquely determined from ideal prevalence data, and that the addition of vaccination serves to force important parameters associated with the natural basic reproduction number and the re-infection threshold to be unidentifiable. This outcome may prove important given that the proportion of vaccinees that successfully take an administered vaccine and acquire sufficient protection is often very difficult to measure directly, and this work suggests that the discussed vaccination models are not appropriate for estimating this effective coverage. It is also shown, in the case of the SIRS model, that it is not possible to uniquely determine the potential success of even an ideal birth targeted vaccination programme with respect to eradication of the infectious agent.

It should be noted that in both the SIR and SIRp cases, if R_0 (specifically β) or the combination kN is known, perhaps from fitting the unvaccinated model to pre-vaccination prevalence data, then P_v can be uniquely identified. However, estimates from pre-intervention data are only appropriate if the parameters β, k and N can be confidently considered to have remained constant over the period of time corresponding to pre and post vaccination. Confidence in the consistency of these parameters is limited given the variable nature of population sizes and the observation gain, and the dependency of infection transmission on unpredictable social and environmental variables. The consequence of this is that assuming the selected model structure remains appropriate for the system, it cannot be uniquely determined whether an applied vaccination programme has failed due to an increase in R_0 or from an inadequate P_v.

Although the models considered in this work are very basic, and would not be the primary basis for a national public health intervention, extended models with greater depth of realism and additional complexity are unlikely to reduce the identifiability problem given increasing degrees of freedom and continued limitations on the observation of the system.

REFERENCES

V. Capasso. *Mathematical Structures of Epidemic Systems*. Springer-Verlag, 1993.

N. D. Evans, M. J. Chapman, K. R Godfrey, and M. J. Chappell. Identifiability of uncontrolled nonlinear rational systems. *Automatica*, 38:1799–1805, 2002.

N. D. Evans, L. J. White, M. J. Chapman, K. R Godfrey, and M. J. Chappell. The structural identifiability of the susceptible infected recovered model with seasonal forcing. *Math Biosci*, 194(2):175–197, Apr 2005.

K. R. Godfrey and W. R. Fitch. The deterministic identifiability of nonlinear pharmacokinetic models. *Pharmacokinetics and Biopharmaceutics*, 12(2):177–191, 1984.

M. Gabriela M Gomes, Lisa J White, and Graham F Medley. Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives. *J Theor Biol*, 228(4):539–549, Jun 2004.

R. Hermann and A. J. Krener. Nonlinear controllability and observability. *IEEE Trans. Automat. Control*, AC-22(3):728–740, 1977.

J. A. Jacquez. *Compartmental Analysis in Biology and Medicine*. BioMedware, third edition, 1996.

M. J. Keeling and B. T Grenfell. Understanding the persistence of measles: reconciling theory, simulation and observation. *Proc Biol Sci*, 269(1489):335–343, 2002.

W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics, part i. *Proc. Roy. Soc. A*, 115:700–721, 1927.

A. Weber, M. Weber, and P. Milligan. Modeling epidemics caused by respiratory syncytial virus (rsv). *Math Biosci*, 172(2):95–113, Aug 2001.

L. J. White, J. N. Mandl, M. G M Gomes, A. T. Bodley-Tickell, P. A. Cane, P. Perez-Brena, J. C. Aguilar, M. M. Siqueira, S. A. Portes, S. M. Straliotto, M. Waris, D. J. Nokes, and G. F. Medley. Understanding the transmission dynamics of respiratory syncytial virus using multiple time series and nested models. *Math Biosci*, 209(1):222–239, Sep 2006.