Journal of Biological Dynamics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tjbd20

On the formulation of epidemic models (an appraisal of Kermack and McKendrick)

D. Breda [a], O. Diekmann [b], W. F. de Graaf [b], A. Pugliese [c] & R. Vermiglio [a]

[a] Dipartimento di Matematica e Informatica, Università degli Studi di Udine, Udine, Italy

[b] Department of Mathematics, Utrecht University, Utrecht, The Netherlands

[c] Dipartimento di Matematica, Università di Trento, Povo, TN, Italy

Version of record first published: 17 Aug 2012.

To cite this article: D. Breda , O. Diekmann , W. F. de Graaf , A. Pugliese & R. Vermiglio (2012): On the formulation of epidemic models (an appraisal of Kermack and McKendrick), Journal of Biological Dynamics, 6:sup2, 103-117

To link to this article: http://dx.doi.org/10.1080/17513758.2012.716454

Journal of Biological Dynamics
Vol. 6, Suppl. 2, September 2012, 103–117

Taylor & Francis
Taylor & Francis Group

On the formulation of epidemic models (an appraisal of Kermack and McKendrick)†

D. Breda[a], O. Diekmann[b], W.F. de Graaf[b]*, A. Pugliese[c] and R. Vermiglio[a]

[a]*Dipartimento di Matematica e Informatica, Università degli Studi di Udine, Udine, Italy;* [b]*Department of Mathematics, Utrecht University, Utrecht, The Netherlands;* [c]*Dipartimento di Matematica, Università di Trento, Povo (TN), Italy*

(*Received 15 May 2012; final version received 24 July 2012*)

The aim of this paper is to show that a large class of epidemic models, with both demography and non-permanent immunity incorporated in a rather general manner, can be mathematically formulated as a *scalar* renewal equation for the force of infection.

Keywords: infectious disease epidemiology; demographic turnover; waning immunity; endemic state; force of infection; renewal equation

1. Introduction

As early as 1927, Kermack and McKendrick published a paper 'A contribution to the mathematical theory of epidemics' in the Proceedings of the Royal Society London Ser. A [31]. The paper became a classic in infectious disease epidemiology and has been cited innumerable times. It was reprinted, with a discussion by Roy Anderson, in a special issue 'Classics of theoretical biology' (part two) in the Bulletin of Mathematical Biology [34].

But how often is it actually read? Judging from an incessant misconception of its contents, one is inclined to conclude: hardly ever! Indeed, even experienced experts often believe that the paper is just about the system

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -\beta IS$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta IS - \alpha I$$
$$\frac{\mathrm{d}R}{\mathrm{d}t} = \alpha I$$

*Corresponding author. Email: w.f.degraaf@uu.nl
Author Emails: o.diekmann@uu.nl; w.f.degraaf@uu.nl; pugliese@science.unitn.it; rossana.vermiglio@uniud.it

†Dedicated to Simon Levin on the occasion of his 70th birthday. Thank you Simon for being an organizing centre for mathematical biology and a source of inspiration for two of the authors for many years.

ISSN 1751-3758 print/ISSN 1751-3766 online

http://dx.doi.org/10.1080/17513758.2012.716454
http://www.tandfonline.com

of three ODEs. This system figures prominently in Section 3.2 of the paper, but there it is explicitly described as a very special (yet important) case of a much more general model formulated in Section 2. The key feature of the general model is that it incorporates a 'per capita rate of infectivity ϕ' which depends on the time θ elapsed since the infection took place (below we use the notation A and τ instead of ϕ and θ). The mathematical incarnation of the general model is a nonlinear renewal equation! A more detailed presentation of the underlying individual-based assumptions and of the derivation of the renewal equation can be found in [45]. In Section 5, we provide an incomplete list of papers dealing with epidemic models formulated in terms of delay equations.

The same issue of the *Bulletin of Mathematical Biology* has also reprinted two follow-up papers:

- Contributions to the mathematical theory of epidemics-II. The problem of endemicity [32,35].
- Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity [33,36].

These follow-up papers are not nearly as much cited as the 1927 paper [31,34]. A possible explanation might be that the 1927 paper arrives at a general and robust conclusion (viz., the Threshold Theorem), while, in contrast, the other two papers discuss a multitude of formulations and partial results that do not culminate in one robust 'law'. By the way, see [4] for a recent critique of one particular aspect of the 1927 paper.

In the spirit of Inaba [29], the aim of the present paper is to take up 'the problem of endemicity' and to show that, actually, one can formulate it in a way that is very reminiscent of the general epidemic model of 1927, provided one takes the force of infection as the primary unknown. In particular, we show that one can derive a *scalar* nonlinear renewal equation for the force of infection under rather general assumptions on both the demographic turnover and the waning of immunity. In order to introduce this approach, we first rederive in Section 2 the main results of the 1927 paper by employing the formulation in terms of the renewal equation for the force of infection. In this 'epidemic' setting, the population is demographically closed and infection leads to permanent immunity. In Section 3, we relax the first of these assumptions and introduce newborn individuals, which are susceptible, at a constant rate B. Rather than introducing an age-independent per capita death rate, we describe the survival probability till at least age a by a general decreasing function $\mathcal{F}(a)$ in order to be able to incorporate the rather flat age distributions that characterize modern developed countries. We derive the unsurprising result that there is a unique endemic equilibrium when $R_0 > 1$ and no endemic equilibrium when $R_0 < 1$. For the special case of a constant *per capita* death rate μ (so $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$), we show that the endemic equilibrium is locally asymptotically stable, no matter how infectiousness depends on time since infection. We are unable to extend this result to general survival functions and wonder whether or not one might have the possibility of a Hopf bifurcation in that setting.

In Section 4, we drop the assumption of permanent immunity and allow for re-infection of the same individual. We note that there are multiple options for describing partial protection but that, whichever of these we choose, the force of infection is determined by a scalar renewal equation (the difference is in the kernel). We briefly discuss the existence and uniqueness of an endemic state, but essentially leave this as an open problem (for ourselves), and we do not even start to discuss the question of local stability. Our plan is to investigate these issues in the near future. Moreover, we would like to incorporate vaccination schedules and multiple strains (perhaps in Adaptive Dynamics spirit) in order to study how (and on what time scale) vaccination may lead to strain replacement in populations with realistic demographic structure.

In the appendix, we show how the SIS[1] compartmental model fits into the framework described in this paper.

2. An epidemic in a closed population

We denote

$$F(t) := \text{force of infection at time } t$$
$$S(t) := \text{density (number per unit area) of susceptibles at time } t$$

and recall that the force of infection is, by definition, the probability per unit of time that a susceptible becomes infected. So, if numbers are large enough to warrant a deterministic description, we have

$$\text{incidence} = F(t)S(t),$$

where 'incidence' is defined as the number of new cases per unit of time and area. In a demographically closed population, the variable S only changes due to the transmission of infection, that is,

$$\dot{S}(t) = -\text{incidence},$$

if, as indeed we assume, the infection leads to permanent immunity. As the key modelling ingredient, we now introduce

$A(\tau) :=$ expected contribution to the force of infection by an individual that was itself infected τ units of time ago.

The constitutive equation

$$F(t) = \int_0^\infty F(t-\tau)S(t-\tau)A(\tau)\,\mathrm{d}\tau \tag{1}$$

now tells us how the current force of infection depends on past incidence. By integrating

$$\dot{S}(t) = -F(t)S(t),$$

we obtain

$$S(t) = S(-\infty)\,\mathrm{e}^{-\int_{-\infty}^{t} F(\sigma)\,\mathrm{d}\sigma}, \tag{2}$$

and if we substitute Equation (2) into Equation (1), we obtain a nonlinear scalar renewal equation for the unknown F.

Quite in general, a delay equation is a rule for extending a function of time towards the future on the basis of the (assumed to be) known past. Such equations lead to dynamical systems by considering the shift along the extended function, see [15,16]. The renewal equation (1), with (2) substituted, fits into this framework!

But before we rederive the main conclusion of the Kermack–McKendrick 1927 paper from Equation (1), let us briefly discuss how one would choose the function A of time-since-infection τ.

First of all, we emphasize that familiar compartmental epidemic models are included in the present framework:

$$A(\tau) = \beta \mathrm{e}^{-\alpha\tau} \Leftrightarrow \text{SIR}$$
$$A(\tau) = \beta \frac{\gamma}{\gamma-\alpha}(\mathrm{e}^{-\alpha\tau} - \mathrm{e}^{-\gamma\tau}) \Leftrightarrow \text{SEIR}$$

(we leave it to our readers to elaborate the $\Leftrightarrow$; the general idea that delay equations with kernels defined in terms of matrix exponentials correspond to systems of ODEs is called 'Linear Chain

Figure 1. A typical infectivity function $A(\tau)$.

Trickery', see e.g. [41]). So, certain parameterized families of A correspond to well-known ODE models and the parameters can thus be interpreted as the mean duration of the infectious period, etc. Quite generally, A has two components:

- contact intensity
- infectiousness (i.e. probability of transmission, given a contact with a susceptible).

Detailed medical information may yield an idea of relative infectiousness as a function of τ. The specification of one scalar factor is then needed to turn this into a graphical specification of A (Figure 1). As we will see below, some indicators of 'severity' (in particular, R_0 and the final size) depend only on the integral of A and not on the form of the graph of A.

In order to be mathematically precise, we assume

$$A(\tau) \geq 0$$
$$A : [0, \infty) \longrightarrow [0, \infty) \text{ is integrable .}$$

Our aim is now to derive a general conclusion from Equations (1) and (2). To do so, we introduce the cumulative force of infection

$$y(t) := \int_{-\infty}^{t} F(\sigma)\, \mathrm{d}\sigma. \tag{3}$$

By integrating Equation (1) from $-\infty$ to t, interchanging the order of the integrals and using

$$F(t-\tau)S(t-\tau) = -\dot{S}(t-\tau)$$

and then Equation (2), we obtain the scalar nonlinear renewal equation

$$y(t) = \int_{0}^{\infty} (1 - \mathrm{e}^{-y(t-\tau)})S(-\infty)A(\tau)\, \mathrm{d}\tau \tag{4}$$

(note, incidentally, that Equation (4) is of convolution type, whereas Equation (1), with Equation (2) inserted, is not). Since $F \geq 0$, we know that y should be an increasing function of t and with a little bit of effort one can check that the rule for extension (4) does indeed preserve this property; that is, if y is increasing on $(-\infty, 0)$, then y is increasing on $(0, +\infty)$ [11]. Moreover, since $1 - \mathrm{e}^{-y} \leq 1$ for $y \geq 0$ and A is integrable, y is bounded. So, the limit $y(\infty) = \lim_{t\to\infty} y(t)$

exists. Using again the integrability of A, one shows that $y(\infty)$ has to satisfy the equation

$$y(\infty) = R_0(1 - \mathrm{e}^{-y(\infty)}), \tag{5}$$

where

$$R_0 := S(-\infty)\int_0^\infty A(\tau)\,\mathrm{d}\tau \tag{6}$$

can be interpreted as the expected number of secondary cases caused by a primary case introduced in a population with susceptible density $S(-\infty)$ or, in the standard jargon, R_0 is the *basic reproduction ratio*.

Using elementary properties of the function $y \mapsto 1 - \mathrm{e}^{-y}$ and simple graphical arguments, one proves that Equation (5) has a unique strictly positive solution if $R_0 > 1$ and no strictly positive solution if $R_0 \leq 1$. In order to translate this information into the Kermack–McKendrick threshold theorem, we first observe that Equation (2) implies

$$1 - \frac{S(\infty)}{S(-\infty)} = 1 - \mathrm{e}^{-y(\infty)} = \frac{1}{R_0}y(\infty), \tag{7}$$

where the left-hand side represents the *final size* of the epidemic, that is, the fraction of the population that is infected, sooner or later, during an outbreak. (See Chapters 18–20 in [51] for an encompassing discussion of the final size equation. And see Appendix B in [47] as well as [2,7] for multi-type generalizations.) Thus, we are led to conclude that

(i) when $R_0 > 1$, introduction of the infective agent leads to an outbreak with the final size given by Equation (7), where $y(\infty)$ is the strictly positive solution of Equation (5)
(ii) when $R_0 \leq 1$, introduction of the infective agent leads to an outbreak with the final size close to zero.

But several comments are essential for both the derivation of this conclusion and its interpretation:

Concerning (i): this is the deterministic formulation that takes as a starting point that a small but positive *fraction* of the large population is infected. When we start with a small *number* of infected individuals, we should describe the initial stages by a branching process. The condition $R_0 > 1$ then amounts to the branching process being supercritical. As a branching process may go extinct, even when supercritical, we may get a *minor outbreak*. Equation (7) describes the expected size of a major outbreak. See, for instance, the textbook of Diekmann and Heesterbeek [13].

Concerning (ii): for $R_0 \leq 1$, the relevant solution of Equation (7) is zero, so where does the 'close to zero' come from? In deriving Equation (5), we assumed that y satisfied Equation (4) for all $t \in (-\infty, \infty)$. When providing an initial condition that captures the introduction of a small number of infectives, we essentially replace Equation (4) by

$$y(t) = \int_0^t (1 - \mathrm{e}^{-y(t-\tau)})S(-\infty)A(\tau)\,\mathrm{d}\tau + f(t), \tag{8}$$

where $f(t)$ is a given function. Thus, we obtain

$$y(\infty) = R_0(1 - \mathrm{e}^{-y(\infty)}) + f(\infty), \tag{9}$$

and conclusion (ii) should be read as follows:

(ii′) when $R_0 \leq 1$, the positive solution $y(\infty)$ of Equation (9) converges to zero when $f(\infty) \downarrow 0$.

See Exercise 1.12.iv in [13] or [51] for details.

If transmission occurs on a time scale which is fast relative to the time scale of demographic changes (birth, death and migration), we may conceive the host population as static during the outbreak. If, moreover, infection confers permanent immunity, the fraction of the population that is susceptible necessarily decreases monotonically and the concept of 'final size' makes sense. In this section, we have shown that the expected final size can be characterized in terms of the scalar equation (5) involving only one parameter, the basic reproduction number R_0. In turn, R_0 is defined by Equation (6), that is, as the product of the original density of susceptibles and the expected total contribution to the force of infection of one newly infected individual. Equation (5) was derived from a nonlinear scalar renewal equation. The entire content of this section is presented in the 1927 paper of Kermack–McKendrick! The only thing we have done is to rewrite the derivation in such a way that the transition to the next sections, dealing with demographic turnover and with waning immunity, is as smooth as possible.

3. The endemic balance

Even if individuals remain immune for the rest of their life after an infection, new susceptibles arise as a result of reproduction. In this section, we consider the situation where, at the population level, there is a constant birth rate B. So from the point of view of the infectious agent, there is a constant inflow at rate B of 'resource'.

Individuals come and go: we now also need to specify how long would newborn individuals live. We do so in terms of the probability $\mathcal{F}(a)$ that an individual stays alive for at least a units of time (see [20] for a very early use of this description in terms of $\mathcal{F}$ in the context of an epidemic model). In particular, we assume that this survival probability does not depend on whether or not the individual becomes infected (so we investigate the influence of demographic turnover on disease transmission and not the influence of a deadly infectious disease such as HIV on population dynamics). We would like to keep $\mathcal{F}$ general and not restrict to the special family $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$ with $\mu > 0$, simply since in modern developed countries the observed $\mathcal{F}$ is hugely different from exponential. Note that we have introduced two more model ingredients (on top of $A(\tau)$), viz., B and $\mathcal{F}(a)$. We assume that $\mathcal{F}(0) = 1$, that $\mathcal{F}$ is monotonically non-increasing and that $\mathcal{F} \to 0$ exponentially for $a \to \infty$. We consider $B > 0$.

If no infectious disease is circulating, we observe the stable age distribution

$$S(t, a) = B\mathcal{F}(a),$$

where S is now the number of susceptibles per unit of area and age. If we assume that 'age' has no impact on susceptibility or on the force of infection that an individual experiences, we obtain

$$S(t, a) = B\mathcal{F}(a)\,\mathrm{e}^{-\int_0^a F(t-a+\sigma)\,\mathrm{d}\sigma} \tag{10}$$

when $F(t)$ is the prevailing force of infection at time t (indeed, an individual that has age a at time t experienced a force of infection $F(t - a + \sigma)$ at age σ). If we substitute Equation (10) into

$$\text{incidence} = F(t) \int_0^\infty S(t, a)\,\mathrm{d}a,$$

we, exactly as before, express the incidence in terms of past values of the force of infection. Instead of the constitutive equation (1), we now postulate

$$F(t) = \int_0^\infty F(t - \tau) \int_0^\infty S(t - \tau, a) \frac{\mathcal{F}(a + \tau)}{\mathcal{F}(a)}\,\mathrm{d}a A(\tau)\,\mathrm{d}\tau \tag{11}$$

since we interpret A as an expected contribution given survival, and so the true expected contribution has the conditional survival probability $\mathcal{F}(a + \tau)/\mathcal{F}(a)$ as an extra factor.

If we substitute Equation (10) into Equation (11), we arrive at a nonlinear scalar renewal equation for the unknown F. Apart from the disease-free state $F(t) \equiv 0$, there might be an endemic steady state. To find it, we need to solve

$$1 = B \int_0^\infty \int_0^\infty \mathrm{e}^{-Fa} \mathcal{F}(a+\tau)\, \mathrm{d}aA(\tau)\, \mathrm{d}\tau. \tag{12}$$

The right-hand side is a monotonically decreasing function of F, converging to zero for $F \to \infty$. So, if we assume that the value at $F = 0$, which we denote by R_0 (since it is indeed the expected number of secondary cases caused by a primary case introduced in a susceptible population with age distribution $B\mathcal{F}(a)$), exceeds one, there exists exactly one steady endemic force of infection. And if $R_0 < 1$, no such steady endemic force of infection exists.

The Principle of Exchange of Stability [37] guarantees that the endemic steady state is stable for R_0 slightly greater than one. To ascertain the local stability for larger values of R_0, we linearize (11)–(10) and look for solutions of the form $\mathrm{e}^{\lambda t}$ for the linearized equation

$$\begin{aligned} u(t) = {} & B \int_0^\infty u(t-\tau) \int_0^\infty \mathrm{e}^{-Fa} \mathcal{F}(a+\tau)\, \mathrm{d}aA(\tau)\, \mathrm{d}\tau \\ & - BF \int_0^\infty \int_0^\infty \mathrm{e}^{-Fa} \mathcal{F}(a+\tau) \int_0^a u(t-\tau-a+\sigma)\, \mathrm{d}\sigma\, \mathrm{d}aA(\tau)\, \mathrm{d}\tau. \end{aligned}$$

This leads to the *characteristic equation*

$$1 = B \int_0^\infty \int_0^\infty \mathrm{e}^{-Fa} \mathcal{F}(a+\tau) \left(1 - \frac{F}{\lambda}(1 - \mathrm{e}^{-a\lambda})\right) \mathrm{e}^{-\lambda\tau}\, \mathrm{d}aA(\tau)\, \mathrm{d}\tau, \tag{13}$$

with F being the unique positive solution of Equation (12) for $R_0 > 1$. It is an open problem to determine whether or not Equation (13) has solutions with $\mathrm{Re}(\lambda) > 0$ for general $\mathcal{F}$. Remarkably, we can show that when $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$, all roots satisfy $\mathrm{Re}(\lambda) < 0$ (for *any* non-negative and integrable function A!).

Notation: $\bar{A}$ is the Laplace transform of A, that is, $\bar{A}(z) := \int_0^\infty A(\tau)\, \mathrm{e}^{-z\tau}\, \mathrm{d}\tau$.

THEOREM 3.1 *Assume that* $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$ *with* $\mu > 0$ *and that* $R_0 > 1$ *where*

$$R_0 = \frac{B}{\mu} \bar{A}(\mu).$$

The solution

$$F = B\bar{A}(\mu) - \mu \tag{14}$$

of (12) *is locally asymptotically stable as a steady state of* (11)–(10).

Proof By the theory developed in [12,16], it suffices to show that all roots of Equation (13) have a negative real part. A straightforward computation, using also Equation (14), shows that

Equation (13) takes the form

$$1 = \frac{\lambda + \mu}{\lambda + B\bar{A}(\mu)} \frac{\bar{A}(\lambda + \mu)}{\bar{A}(\mu)} \tag{15}$$

when $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$. Since (recall that $A(\tau) \geq 0$!)

$$|\bar{A}(\lambda + \mu)| \leq \bar{A}(\mathrm{Re}(\lambda) + \mu)$$

and for $\mathrm{Re}(\lambda) \geq 0$

$$\bar{A}(\mathrm{Re}(\lambda) + \mu) \leq \bar{A}(\mu),$$

the absolute value of the second factor on the right-hand side is bounded above by one. Since $R_0 > 1$, we have

$$B\bar{A}(\mu) + \mathrm{Re}(\lambda) > \mu + \mathrm{Re}(\lambda) > 0$$

for $\mathrm{Re}(\lambda) \geq 0$ and hence

$$\left| \frac{\lambda + \mu}{\lambda + B\bar{A}(\mu)} \right| = \frac{\sqrt{(\mu + \mathrm{Re}(\lambda))^2 + (\mathrm{Im}(\lambda))^2}}{\sqrt{(B\bar{A}(\mu) + \mathrm{Re}(\lambda))^2 + (\mathrm{Im}(\lambda))^2}} < 1.$$

So for $\mathrm{Re}(\lambda) \geq 0$, the right-hand side of Equation (15) is in absolute value less than one and consequently (15) cannot hold for such λ. ■

Remark The proof given above is a much simplified version of the elaboration of [13, Exercise 3.10]. See [52, Theorem 5] for a different proof and various generalizations. Global stability has been shown in [42] and a text-book version of that proof can be found in [49, Section 9.9].

In this section, we have shown that when demographic turnover is described by a constant population birth rate and a general survival function, we can still capture the infectious disease dynamics by a scalar renewal equation for the force of infection. This equation has a positive steady state if and only if $R_0 > 1$ and there is at most one such steady state. For the special case of a constant *per capita* death rate, we can find that this steady state is locally asymptotically stable. Whether or not it is for general survival is an open problem. But note that Andreasen [1], Magal and Ruan [43] and Thieme [50] found that instability is certainly possible if we allow the susceptibility or the infectiousness of an individual to depend on its age.

4. Waning immunity

When immunity reduces susceptibility, but has no impact on infectiousness once the infection occurs, we can stick to the already introduced $A(\tau)$ to describe the expected contribution to the force of infection at time τ after infection, given survival.

There are at least two ways to describe temporary partial immunity. One may assume that all individuals, at time τ after the last infection took place, have their susceptibility reduced by a factor $Q(\tau)$, meaning that the probability of getting infected upon contact with an infectious individual is $Q(\tau)$ times what it is for a fully susceptible individual (so $Q(\tau) = 0$ corresponds to full protection and $Q(\tau) = 1$ to full susceptibility). Another possibility is to allow only for full protection or full susceptibility and to describe the probability that at time τ after the last infection an individual is fully susceptible again by $P(\tau)$. Note that SIS and SIRS compartment models (see Appendix) are of this type. Note that Inaba [28–30] followed Kermack and McKendrick [32,33,35,36] in assuming that a new clock starts when an event transforms an infected individual

into a recovered individual and that the reduction in susceptibility is a function of the time shown by the second clock (see [54] for a similar approach; also see [18,43,46,53]). Also note that one can, of course, introduce a compartment of intermediate protection (see e.g. [39] and [48] and the references given there) and thus obtain a description that combines features of both gradual return of susceptibility and variability in the length of the time window of protection. One might also assume that susceptibility does not simply depend on the time since last infection, but also depends on the history of previous infections [39].

Note that we have to make sure that an infected individual that contributes to the positivity of $P(\tau)$ should *not* contribute to the positivity of $A(\sigma)$ for $\sigma \geq \tau$, as otherwise the possibility of re-infection does have an influence on expected infectiousness (for compartmental models of SIS and SIRS type, etc., this is indeed guaranteed). When we work with Q, the safest assumption is to require that $Q(\tau) = 0$ for $\tau \in \mathrm{supp}(A)$. In any case, we can assume that Equation (11), which we repeat here as

$$F(t) = \int_0^\infty F(t-\tau) \int_0^\infty S(t-\tau, a) \frac{\mathcal{F}(a+\tau)}{\mathcal{F}(a)} \, \mathrm{d}aA(\tau) \, \mathrm{d}\tau, \tag{16}$$

remains valid, interpreting (in case of reduced susceptibility) $S(t, a)$ in terms of susceptible equivalents, in the sense that one individual with reduced susceptibility $Q(\tau)$ is counted as $Q(\tau)$ susceptibles. But we have to change expression (10) for the number of susceptibles per unit of area and age, in order to account for regained susceptibility (long) after infection.

Another important observation is that, implicitly, P and Q refer to an individual while conditioning that this individual is not yet re-infected. The possibility of re-infection leads, in the case of partial susceptibility, to the conclusion that of the individuals that become infected at time $t - \tau$, a fraction

$$\mathrm{e}^{-\int_0^\tau Q(\sigma) F(t-\tau+\sigma) \, \mathrm{d}\sigma}$$

will not have been re-infected before time t. If we work in the P-framework and consider again an individual infected at time $t - \tau$, the probability that it is susceptible at time t is given by the Stieltjes integral:

$$\int_0^\tau \mathrm{e}^{-\int_0^{\tau-\sigma} F(t-\tau+\sigma+\eta) \, \mathrm{d}\eta} P(\mathrm{d}\sigma)$$

(indeed, if the individual regains full susceptibility at time σ after infection, it has to escape from re-infection on a time interval of length $\tau - \sigma$ beginning at $t - \tau + \sigma$). These observations motivate us to define

$$\mathcal{G}(\tau, F_t) := \begin{array}{l} \text{either } Q(\tau) \, \mathrm{e}^{-\int_0^\tau Q(\sigma) F(t-\tau+\sigma) \, \mathrm{d}\sigma} \\ \text{or } \int_0^\tau \mathrm{e}^{-\int_0^{\tau-\sigma} F(t-\tau+\sigma+\eta) \, \mathrm{d}\eta} P(\mathrm{d}\sigma) \end{array} \tag{17}$$

in order to capture Q-models and P-models in one uniform notation. Here, we use the notational convention

$$F_t(\sigma) := F(t+\sigma), \sigma \leq 0 \tag{18}$$

of the theory of delay equations [15,21].

We are now ready to replace expression (10) by the equation

$$S(t, a) = B\mathcal{F}(a) \, \mathrm{e}^{-\int_0^a F(t-a+\sigma) \, \mathrm{d}\sigma} + \int_0^a F(t-\tau) S(t-\tau, a-\tau) \frac{\mathcal{F}(a)}{\mathcal{F}(a-\tau)} \mathcal{G}(\tau, F_t) \, \mathrm{d}\tau. \tag{19}$$

Thus, the model is described by the two equations (16) and (19) for the two unknowns $F(t)$ and $S(t, a)$ and it has as its ingredients B, $\mathcal{F}(a)$, $A(\tau)$ and either $Q(\tau)$ or $P(\tau)$. For a given F, Equation (19) is linear in S, so solving Equation (19) is in principle straightforward, and upon

substitution of the result into Equation (16), we end up with a scalar nonlinear renewal equation for F. In fact, one can solve Equation (19) by successive approximation, which amounts to generation expansion: the term $B\mathcal{F}(a)\,\mathrm{e}^{-\int_0^a F(t-a+\sigma)\,\mathrm{d}\sigma}$ corresponds to individuals that have never been infected and the next term

$$B\mathcal{F}(a)\int_0^a F(t-\tau)\,\mathrm{e}^{-\int_0^{a-\tau} F(t-a+\sigma)\,\mathrm{d}\sigma}\mathcal{G}(\tau,F_t)\,\mathrm{d}\tau$$

to individuals that have been infected once, etc. Whenever the support of Q or P is bounded away from zero (meaning that the time window between two successive infections has a strictly positive minimal length), there are, at any given finite age, only finitely many non-zero terms in the generation expansion. So, if in addition $\mathcal{F}(a) = 0$ for any $a \geq a_{\max}$, we find in this way an exact representation of the solution of Equation (19) as a finite sum of explicit expressions involving the history of F. We conclude that the dynamics of an epidemic model that incorporates a general survival function $\mathcal{F}(a)$ and various quite general forms of waning immunity is still fully described by a *scalar* nonlinear renewal equation!

If we define $q(t,a)$ by the relation

$$S(t,a) = B\mathcal{F}(a)q(t,a), \tag{20}$$

then q describes the relative susceptibility (Q-version) or the probability to be susceptible (P-version) of an individual of age a at time t, given survival. Upon substitution of this relation into Equation (19), we find for q the equation

$$q(t,a) = \mathrm{e}^{-\int_0^a F(t-a+\sigma)\,\mathrm{d}\sigma} + \int_0^a F(t-\tau)q(t-\tau,a-\tau)\mathcal{G}(\tau,F_t)\,\mathrm{d}\tau, \tag{21}$$

while substitution into Equation (16) leads to

$$F(t) = B\int_0^\infty F(t-\tau)\int_0^\infty q(t-\tau,a)\mathcal{F}(a+\tau)\,\mathrm{d}aA(\tau)\,\mathrm{d}\tau. \tag{22}$$

Clearly, $F(t) \equiv 0$, $q(t,a) \equiv 1$ corresponds to the disease-free steady state. An endemic steady state is characterized by

$$q(a) = \mathrm{e}^{-Fa} + F\int_0^a q(a-\tau)\mathcal{G}(\tau,F)\,\mathrm{d}\tau, \tag{23}$$

$$1 = B\int_0^\infty\int_0^\infty q(a)\mathcal{F}(a+\tau)\,\mathrm{d}aA(\tau)\,\mathrm{d}\tau. \tag{24}$$

The right-hand side of Equation (24) for $q(a) \equiv 1$ is again equal to R_0, just as it was in the preceding section (the possibility of re-infection has no impact on R_0 for the simple reason that when infection is very rare, it can be excluded that the same individual is infected again). If, as before, we solve the linear equation (23) by generation expansion and substitute the result into Equation (24), we obtain the scalar equation from which we have to determine the unknown $F > 0$.

In the special case that $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$, Equation (24) reduces to

$$1 = B\bar{A}(\mu)\bar{q}(\mu). \tag{25}$$

Equation (23) is a linear renewal equation, so it can be solved by Laplace transformation, leading to

$$\bar{q}(\mu) = \frac{1}{1 - F\bar{\mathcal{G}}(\mu, F)} \frac{1}{\mu + F}. \tag{26}$$

In the P-version, we have

$$\bar{\mathcal{G}}(\mu, F) = \frac{1}{\mu + F}\bar{P}(\mu), \tag{27}$$

where $\bar{P}(\mu)$ is the Laplace–Stieltjes transform of P:

$$\bar{P}(\mu) := \int_0^\infty \mathrm{e}^{-\mu\tau} P(\mathrm{d}\tau). \tag{28}$$

So for the P-version, we find

$$\bar{q}(\mu) = \frac{1}{\mu + F(1 - \bar{P}(\mu))}, \tag{29}$$

and since $\bar{P}(\mu) < 1$, the right-hand side is a decreasing function of F. Accordingly, Equation (25) has a unique positive solution whenever $R_0 > 1$ and no such solution when $R_0 \leq 1$. We conjecture that the same is true for the Q-version, but so far we did not manage to prove this. We also conjecture that this conclusion holds for general $\mathcal{F}(a)$. In order to show this, we might rewrite the factor

$$\int_0^\infty q(a)\mathcal{F}(a + \tau)\,\mathrm{d}a$$

in Equation (24) by integration by parts as

$$-\int_0^\infty \int_0^a q(\alpha)\,\mathrm{d}\alpha \mathcal{F}'(a + \tau)\,\mathrm{d}a,$$

(where, for convenience, we have assumed that $\mathcal{F}$ is differentiable; since $\mathcal{F}$ is anyhow monotone decreasing, the general case can be dealt with by using a Stieltjes integral). Now, note that $\int_0^a q(\alpha)\,\mathrm{d}\alpha$ is the expected time that an individual that survived till age a spent being susceptible (weighed in the appropriate manner when we work with the Q-version). So, it seems very reasonable to conjecture that

$$\frac{\mathrm{d}}{\mathrm{d}F} \int_0^a q(\alpha)\,\mathrm{d}\alpha < 0$$

(note that q itself may depend in a more complicated way on F, as an increase in F leads first to a decrease of the expected age of the first infection but next to an increase of *return* to the susceptible class at an early age). It remains to study in detail how the solution of the renewal equation (23) depends on the parameter F.

5. Concluding remarks

Often when epidemic models are formulated as delay equations, the aim is to derive conditions for the (in)stability of the endemic steady state and to find out, by way of the Hopf bifurcation theorem, when one should expect to find persistent oscillations [5,6,10,14,22–26,29,43,55,56].

The trigger for the approach sketched in the present paper is somewhat different: our ultimate goal is to find conditions for strain replacement as a result of mass vaccination [17,19,27,38,44]. In this context, the need arises to incorporate

(i) a realistic form of demographic turnover,
(ii) the waning of immunity and
(iii) cross-immunity.

The above text describes our approach for dealing with issues (i) and (ii). We realise that dealing in an effective yet meaningful way with issue (iii) presents a major challenge. Since Simon Levin has considerable expertise in this matter [3,8,9,40], his extremely valuable advice on this matter will be highly appreciated.

Note

1. S = susceptible, I = infectious, R = removed and E = exposed.

Acknowledgements

We are grateful to Horst Thieme for providing us with a substantial list of very relevant references! Comments by Horst and by Hisashi Inaba have been very helpful.

References

[1] V. Andreasen, *Instability in an SIR-model with age-dependent susceptibility*, in *Mathematical Population Dynamics: Analysis of Heterogeneity Volume 1: Theory of Epidemics*, O. Arino, D. Axelrod, M. Kimmel, and M. Langlais, eds., Wuerz Publishing, Winnipeg, 1995, pp. 3–14.
[2] V. Andreasen, *The final size of an epidemic and its relation to the basic reproduction number*, Bull. Math. Biol. 73 (2011), pp. 2305–2321.
[3] V. Andreasen, J. Lin, and S.A. Levin, *The dynamics of cocirculating influenza strains conferring partial cross-immunity*, J. Math. Biol. 35 (1997), pp. 825–842.
[4] N. Bacaër, *The model of Kermack and McKendrick for the plague epidemic in Bombay and the type reproduction number with seasonality*, J. Math. Biol. 64 (2012), pp. 403–422.
[5] E. Beretta and D. Breda, *An SEIR epidemic model with constant latency time and infectious period*, Math. Biosci. Eng. 8 (2011), pp. 931–952.
[6] K.B. Blyuss and Y.N. Kyrychko, *Stability and bifurcations in an epidemic model with varying immunity period*, Bull. Math. Biol. 72 (2009), pp. 490–505.
[7] F. Brauer, *Heterogeneous mixing in epidemic models*, preprint (2012), to appear in Can. Appl. Math. Q.
[8] C. Castillo-Chavez, H.W. Hethcote, V. Andreasen, S.A. Levin, and W.M. Liu, *Cross-immunity in the dynamics of homogeneous and heterogeneous populations*, in *Mathematical Ecology*, T. Hallam, L. Gross and S.A. Levin, eds., Proceedings of the Autumn Course Research Seminars, Trieste, 1986, World Scientific, Singapore, 1988, pp. 303–316.
[9] C. Castillo-Chavez, H.W. Hethcote, V. Andreasen, S.A. Levin, and W.M. Liu, *Epidemiological models with age structure, proportionate mixing, and cross-immunity*, J. Math. Biol. 27 (1989), pp. 233–258.
[10] K.L. Cooke and P. Van Den Driessche, *Analysis of an SEIRS epidemic model with two delays*, J. Math. Biol. 35 (1996), pp. 240–260.
[11] O. Diekmann, *Limiting behaviour in an epidemic model*, NLA-TMA 1 (1977), pp. 459–470.
[12] O. Diekmann and M. Gyllenberg, *Equations with infinite delay: Blending the abstract and the concrete*, J. Differ. Equ. 252 (2012), pp. 819–851.
[13] O. Diekmann and J.A.P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation (Mathematical and Computational Biology)*, John Wiley, Chichester, 2000.
[14] O. Diekmann and R. Montijn, *Prelude to Hopf bifurcation in an epidemic model: Analysis of a characteristic equation associated with a nonlinear Volterra integral equation*, J. Math. Biol. 14 (1982), pp. 117–127.
[15] O. Diekmann, S.A. van Gils, S.M. Verduyn Lunel, and H.-O. Walther, *Delay Equations: Functional, Complex, and Nonlinear Analysis*, Applied Mathematical Sciences, Vol. 110, Springer-Verlag, Berlin, 1995.
[16] O. Diekmann, Ph. Getto, and M. Gyllenberg, *Stability and bifurcation analysis of Volterra functional equations in the light of suns and stars*, SIAM J. Math. Anal. 39 (2007), pp. 1023–1069.
[17] I. Dorigatti and A. Pugliese, *Analysis of a vaccine model with cross-immunity: When can two competing infectious strains coexist?* Math. Biosci. 234 (2011), pp. 33–46.

[18] Z. Feng and H.R. Thieme, *Endemic models with arbitrarily distributed periods of infection I: Fundamental properties of the model*, SIAM J. Appl. Math. 61 (2000), pp. 803–833.
[19] Z. Feng, M. Iannelli, and F.A. Milner, *A two-strain tuberculosis model with age of infection*, SIAM J. Appl. Math. 62 (2002), pp. 1634–1656.
[20] G. Gripenberg, *On a nonlinear integral equation modelling an epidemic in an age-structured population*, J. Reine Angew. Math. 341 (1983), pp. 54–67.
[21] J.K. Hale and S.M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Applied Mathematical Sciences, Vol. 99, Springer-Verlag, Berlin, 1993.
[22] H.W. Hethcote and S.A. Levin, *Periodicity in epidemiological models. Applied mathematical ecology (Trieste, 1986)*, Biomathematics 18 (1989), pp. 193–211.
[23] H.W. Hethcote and D.W. Tudor, *Integral equation models for endemic infectious diseases*, J. Math. Biol. 9 (1980), pp. 37–47.
[24] H.W. Hethcote and P. Van Den Driessche, *An SIS epidemic model with variable population size and a delay*, J. Math. Biol. 34 (1995), pp. 177–194.
[25] H.W. Hethcote, H.W. Stech, and P. Van Den Driessche, *Nonlinear oscillations in epidemic models*, SIAM J. Appl. Math. 40 (1981), pp. 1–9.
[26] G. Huang, E. Beretta, and Y. Takeuchi, *Global stability for epidemic model with constant latency and infectious periods*, Math. Biosci. Eng. 9 (2012), pp. 297–312.
[27] M. Iannelli, M. Martcheva, and X.-Z. Li, *Strain replacement in an epidemic model with super-infection and perfect vaccination*, Math. Biosci. 195 (2005), pp. 23–46.
[28] H. Inaba, *Mathematical analysis for an evolutionary epidemic model*, in *Mathematical Models in Medical and Health Sciences*, M.A. Horn, G. Simonett, and G.F. Webb, eds., Vanderbilt University, Nashville, TN, 1998, p. 213.
[29] H. Inaba, *Kermack and McKendrick revisited: The variable susceptibility model for infectious diseases*, Japan J. Indust. Appl. Math. 18 (2001), pp. 273–292.
[30] H. Inaba, *Endemic threshold and stability in an evolutionary epidemic model*, in *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, C. Castillo-Chavez, S. Blower, P. van den Driessche, D. Kirschner, and A.-A. Yakubu, eds., Springer, New York, 2002, p. 337.
[31] W.O. Kermack and A.G. McKendrick, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. Lond. Ser. A 115 (1927), pp. 700–721.
[32] W.O. Kermack and A.G. McKendrick, *Contributions to the mathematical theory of epidemics-II. The problem of endemicity*, Proc. R. Soc. Lond. Ser. A 138A (1932), pp. 55–83.
[33] W.O. Kermack and A.G. McKendrick, *Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity*, Proc. R. Soc. Lond. Ser. A 141A (1933), pp. 94–122.
[34] W.O. Kermack and A.G. McKendrick, *A contribution to the mathematical theory of epidemics in classics of theoretical biology (part two)*, Bull. Math. Biol. 53 (1991), pp. 33–55.
[35] W.O. Kermack and A.G. McKendrick, *Contributions to the mathematical theory of epidemics-II. The problem of endemicity in classics of theoretical biology (part two)*, Bull. Math. Biol. 53 (1991), pp. 57–87.
[36] W.O. Kermack and A.G. McKendrick, *Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity in classics of theoretical biology (part two)*, Bull. Math. Biol. 53 (1991), pp. 89–118.
[37] H. Kielhöfer, *Bifurcation Theory: An Introduction with Applications to PDEs*, Applied Mathematical Sciences, Vol. 156, Springer, New York, 2004.
[38] S. Kryazhimskiy, U. Dieckmann, S.A. Levin, and J. Dushoff, *On state-space reduction in multi-strain pathogen models, with an application to antigenic drift in influenza A*, PLoS Comput. Biol. 3 (2007), pp. 1513–1525.
[39] J.S. Lavine, A.A. King, and O.N. Bjørnstad, *Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure*, PNAS 108 (2011), pp. 7259–7264.
[40] J. Lin, V. Andreasen, and S.A. Levin, *Dynamics of influenza A drift: The linear three-strain model*, Math. Biosci. 162 (1999), pp. 33–51.
[41] N. MacDonald, *Biological Delay Systems: Linear Stability Theory*, Cambridge University Press, Cambridge, 1989.
[42] P. Magal, C.C. McCluskey, and G.F. Webb, *Lyapunov functional and global asymptotic stability for an infection-age model*, Appl. Anal. 89 (2010), pp. 1109–1140.
[43] P. Magal and S. Ruan, *Sustained oscillations in an evolutionary epidemiological model of influenza A drift*, Proc. R. Soc. A 466 (2010), pp. 965–992.
[44] M. Martcheva, M. Iannelli, and X.-Z. Li, *Subthreshold coexistence of strains: The impact of vaccination and mutation*, Math. Biosci. Eng. 4 (2007), pp. 287–317.
[45] J.A.J. Metz, *The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections*, Acta Biotheoretica 27 (1978), pp. 75–123.
[46] C.M. Pease, *An evolutionary epidemiological mechanism, with applications to type A influenza*, Theoret. Popul. Biol. 31 (1987), pp. 422–452.
[47] L. Rass and J. Radcliffe, *Spatial Deterministic Epidemics*, AMS, Providence, RI, 2003.
[48] A.N. Swart, M. Tomasi, M. Kretzschmar, A.H. Havelaar, and O. Diekmann, *The protective effects of temporary immunity under imposed infection pressure*, Epidemics 4 (2012), pp. 43–47.
[49] H.L. Smith and H.R. Thieme, *Dynamical Systems and Population Persistence*, AMS, Providence, RI, 2011.
[50] H.R. Thieme, *Stability change of the endemic equilibrium in age-structured models for the spread of SIR type infectious diseases*, in *Differential Equations Models in Biology, Epidemiology and Ecology*, S. Busenberg and M. Martelli, eds., Springer, Berlin, 1991, pp. 139–158.

[51] H.R. Thieme, *Mathematics in Population Biology, Princeton Series in Theoretical and Computational Biology*, Princeton University Press, Princeton, NJ, 2003.
[52] H.R. Thieme and C. Castillo-Chavez, *How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS?* SIAM J. Appl. Math. 53 (1993), pp. 1447–1479.
[53] H.R. Thieme and P. van den Driessche, *Global stability in cyclic epidemic models with disease fatalities*, Fields Inst. Commun. 21 (1999), pp. 459–472.
[54] H.R. Thieme and J. Yang, *An endemic model with variable re-infection rate and application to influenza*, Math. Biosci. 180 (2002), pp. 207–235.
[55] H.R. Thieme, A. Tridane, and Y. Kuang, *An epidemic model with post-contact prophylaxis of distributed length. I. Thresholds for disease persistence and extinction*, J. Biol. Dyn. 2 (2008), pp. 221–239.
[56] H.R. Thieme, A. Tridane, and Y. Kuang, *An epidemic model with post-contact prophylaxis of distributed length. II. Stability and oscillations if treatment is fully effective*, Math. Model. Nat. Phenom. 3 (2008), pp. 267–293.

Appendix

Here, we show how the familiar SIS compartmental model described by the ODE system

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= B - \mu S - \beta IS + \alpha I\\ \frac{\mathrm{d}I}{\mathrm{d}t} &= -\mu I + \beta IS - \alpha I\end{aligned} \tag{A1}$$

fits into the framework described in this paper. Hopefully, our presentation is such that the reader is convinced that also models with additional exposed and/or removed compartments, such as the SEIRS model described by

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t} &= B - \mu S - \beta IS + \delta R\\ \frac{\mathrm{d}E}{\mathrm{d}t} &= -\mu E + \beta IS - \gamma E\\ \frac{\mathrm{d}I}{\mathrm{d}t} &= -\mu I + \gamma E - \alpha I\\ \frac{\mathrm{d}R}{\mathrm{d}t} &= -\mu R + \alpha I - \delta R,\end{aligned} \tag{A2}$$

are covered by our approach.

For a given function $F = F(t)$, let $M_i(t;s)$, $i = 1, 2$, be defined by

$$\begin{aligned}\frac{\mathrm{d}M_1}{\mathrm{d}t}(t;s) &= -F(t)M_1(t;s) + \alpha M_2(t;s), \quad M_1(s;s) = 1,\\ \frac{\mathrm{d}M_2}{\mathrm{d}t}(t;s) &= F(t)M_1(t;s) - \alpha M_2(t;s), \quad M_2(s;s) = 0.\end{aligned} \tag{A3}$$

So, $(M_1, M_2)^{\mathrm{T}}$ is the first column of the fundamental matrix solution of the linear ODE system generated by the matrix

$$\begin{pmatrix} -F(t) & \alpha \\ F(t) & -\alpha \end{pmatrix},$$

and if we define

$$\begin{pmatrix} S \\ I \end{pmatrix}(t) = \int_{-\infty}^{t} B\mathrm{e}^{-\mu(t-s)} M(t;s)\,\mathrm{d}s, \tag{A4}$$

it follows that

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t}(t) &= B - \mu S(t) - F(t)S(t) + \alpha I(t),\\ \frac{\mathrm{d}I}{\mathrm{d}t}(t) &= -\mu I(t) + F(t)S(t) - \alpha I(t).\end{aligned} \tag{A5}$$

Our task is to relate Equation (22) to the equation

$$F(t) = \beta I(t) \tag{A6}$$

that allows us to identify (A5) and (A1). But in order to do so, we need to uncover first the relation between Equation (21) and the definition of M.

By using for each of the two equations of (A3) the variation-of-constants formula, we obtain

$$M_1(t;s) = \mathrm{e}^{-\int_s^t F(\sigma)\,\mathrm{d}\sigma} + \alpha \int_s^t \mathrm{e}^{-\int_\sigma^t F(\eta)\,\mathrm{d}\eta} M_2(\sigma;s)\,\mathrm{d}\sigma$$

$$M_2(t;s) = \int_s^t \mathrm{e}^{-\alpha(t-\sigma)} F(\sigma) M_1(\sigma;s)\,\mathrm{d}\sigma,$$

and upon substitution of the second of these into the first, we obtain

$$\begin{aligned} M_1(t;s) &= \mathrm{e}^{-\int_s^t F(\sigma)\,\mathrm{d}\sigma} + \alpha \int_s^t \mathrm{e}^{-\int_\sigma^t F(\eta)\,\mathrm{d}\eta} \int_s^\sigma \mathrm{e}^{-\alpha(\sigma-\theta)} F(\theta) M_1(\theta;s)\,\mathrm{d}\theta\,\mathrm{d}\sigma \\ &= \mathrm{e}^{-\int_s^t F(\sigma)\,\mathrm{d}\sigma} + \int_s^t \int_\theta^t \mathrm{e}^{-\int_\sigma^t F(\eta)\,\mathrm{d}\eta} \alpha \mathrm{e}^{-\alpha(\sigma-\theta)} F(\theta) M_1(\theta;s)\,\mathrm{d}\sigma\,\mathrm{d}\theta \\ &= \mathrm{e}^{-\int_s^t F(\sigma)\,\mathrm{d}\sigma} + \int_0^{t-s} F(t-\tau) M_1(t-\tau;s) \int_0^\tau \mathrm{e}^{-\int_{t-\tau-\sigma}^t F(\eta)\,\mathrm{d}\eta} \alpha \mathrm{e}^{-\alpha\sigma}\,\mathrm{d}\sigma\,\mathrm{d}\tau. \end{aligned}$$

If we now put $M_1(t;s) = q(t, t-s)$ and $s = t-a$, we recover exactly Equation (21) for the special case that $P(\tau) = 1 - \mathrm{e}^{-\alpha\tau}$ (and hence $P(\mathrm{d}\sigma) = \alpha \mathrm{e}^{-\alpha\sigma}\,\mathrm{d}\sigma$).

From Equation (A4), we have

$$I(t) = B \int_0^\infty \mathrm{e}^{-\mu a} M_2(t; t-a)\,\mathrm{d}a,$$

and from the variation-of-constants formula, we have

$$M_2(t; t-a) = \int_{t-a}^t \mathrm{e}^{-\alpha(t-\sigma)} F(\sigma) M_1(\sigma; t-a)\,\mathrm{d}\sigma = \int_0^a \mathrm{e}^{-\alpha(a-\eta)} F(t-a+\eta) M_1(t-a+\eta; t-a)\,\mathrm{d}\eta.$$

If we insert the second of these into the first and write again $M_1(t;s) = q(t, t-s)$, we find upon changing the order of integration

$$\begin{aligned} I(t) &= B \int_0^\infty \int_\eta^\infty \mathrm{e}^{-\mu a - \alpha(a-\eta)} F(t-a+\eta) q(t-a+\eta, \eta)\,\mathrm{d}a\,\mathrm{d}\eta \\ &= B \int_0^\infty \int_0^\infty \mathrm{e}^{-\mu(\theta+\eta) - \alpha\theta} F(t-\theta) q(t-\theta, \eta)\,\mathrm{d}\theta\,\mathrm{d}\eta \\ &= B \int_0^\infty F(t-\theta) \int_0^\infty q(t-\theta, \eta)\,\mathrm{e}^{-\mu\eta}\,\mathrm{d}\eta \mathrm{e}^{-(\mu+\alpha)\theta}\,\mathrm{d}\theta. \end{aligned}$$

If we now multiply both sides by β and insert $A(\tau) = \beta \mathrm{e}^{-\alpha\tau}$ and $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$ into Equation (22), we find that the right-hand side of Equation (22) coincides with the right-hand side of the identity for $\beta I(t)$. In other words, starting from Equation (A1), we have derived Equation (22) for $F(t)$ defined by $F(t) = \beta I(t)$, in the special case that $A(\tau) = \beta \mathrm{e}^{-\alpha\tau}$ and $\mathcal{F}(a) = \mathrm{e}^{-\mu a}$ and with q defined by Equation (21) in the special case that $\mathcal{G}(\tau, F_t) = \int_0^\tau \mathrm{e}^{-\int_0^{\tau-\sigma} F(t-\tau+\sigma+\eta)\,\mathrm{d}\eta} \alpha \mathrm{e}^{-\alpha\sigma}\,\mathrm{d}\sigma$.