THE KERMACK - McKENDRICK EPIDEMIC MODEL AND THE FINAL SIZE RELATION

FRED BRAUER

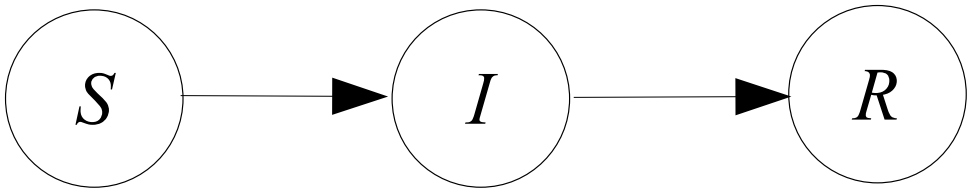
MAY 2021

THE SIMPLE KERMACK - McKENDRICK EPIDEMIC MODEL

In an epidemic, there is an initial stochastic phase until the number of members of the population who are no longer susceptible to infection becomes a significant fraction of the population. During this initial stochastic phase, a branching process model or a contact network model may be used to model the epidemic. Following this initial phase it is reasonable to switch to a compartmental model. The simplest compartmental epidemic model is the simple SIR model, a very special case of the original epidemic model of Kermack and McKendrick (1927).

The simple SIR model in a population of constant total size N is

$$\begin{aligned} S' &= -a\frac{S}{N}I & S(0) &= S_0 \\ I' &= a\frac{S}{N}I - \alpha I & I(0) &= I_0. \end{aligned}$$



The population, of size N, is divided into three compartments, susceptibles (S) who are not infected but may be infected by contact with an infected individual, infectives (I) who are infected and able to transmit infection, and removed individuals (R) who have been removed from infection risk by recovery from infection with immunity against reinfection or death caused by disease.

The model is based on the following assumptions:

(i) An average member of the population makes contact sufficient to transmit infection with βN others per unit time, where N represents total population size (mass action incidence).

(ii) Infectives leave the infective class at rate αI per unit time.

(iii) There is no entry into or departure from the population, except possibly through death from the disease.

(iv) There are no disease deaths, and the total population size is a constant N.

Contacts sufficient to transmit infection depend on the the mechanism of transmission. Many diseases are transmitted by a virus in the air from an infected individual by exhaling, coughing, or sneezing. Sexually transmitted diseases are transmitted by sexual contact. Here, we are concerned only with direct transmission, and do not consider transmission through a vector, as in malaria, or indirectly as in cholera, where infected individuals shed virus into contaminated drinking water and susceptibles are infected by drinking contaminated water. Models for such diseases may be formulated using ideas similar to those described here for directly transmitted diseases.

We think of introducing a small number of infectious individuals into a population of susceptibles and ask whether there will be an epidemic. We observe that $S' < 0$ for all t and $I' > 0$ if and only if $S > \alpha N/a$. Thus I increases so long as $S > \alpha N/a$ but since S decreases for all t, I ultimately decreases and approaches zero.

The quantity $\mathcal{R}_0 = a/\alpha$ determines whether there is an epidemic. If $\mathcal{R}_0 < 1$ the infection dies out because $I'(t) < 0$ for all t, and there is no epidemic. Ordinarily, $S_0 \approx N$.

The original formulation of the Kermack-McKendrick age of infection epidemic model was

$$\begin{aligned} v(t) &= -x'(t) \\ x'(t) &= -x(t)\left[\int_0^t A(s)v(t-s)ds + A(t)y_0\right] \\ z'(t) &= \int_0^t C(s)v(t-s)ds + C(t)y_0 \\ y(t) &= \int_0^t B(s)v(t-s)ds + B(t)y_0. \end{aligned}$$

Here, $x(t)$ is the number of susceptibles, $y(t)$ is the number of infectious individuals, and $z(t)$ is the number of removed individuals. Also $B(s)$ is the recovery rate when the age of infection is s,
$A(s) = \varphi(s)B(s), C(s) = \psi(s)B(s)$, with $\varphi(s)$ the infectivity at infection age s, and $\psi(s)$ the recovery rate at infection age s. Kermack and McKendrick did not bring the basic reproduction number into their analysis, but were able to derive a final size relation in the form

$$\log \frac{1 - \frac{y_0}{N}}{1 - p} = pN \int_0^\infty A(s)ds,$$

in which N is the total population size and p is the attack ratio

$$p = 1 - \frac{x_\infty}{N}.$$

If we define

$$S(t) = x(t), A(s) = B(s) = e^{-\gamma s}, I(t) = y(t)/\beta,$$

the model reduces to the simple Kermack - McKendrick epidemic model.

ANALYSIS OF THE SIMPLE KERMACK - McKENDRICK MODEL

Since the model is a two-dimensional autonomous system of differential equations, the natural approach would be to find equilibria and linearize about each equilibrium to determine its stability. However, since every point with $I = 0$ is an equilibrium, the system the model has a line of equilibria and this approach is not applicable (the linearization matrix at each equilibrium has a zero eigenvalue). The standard linearization theory for systems of ordinary differential equations is not applicable, and it is necessary to develop a new mathematical approach.

The sum of the two equations of the model is

$$(S + I)' = -\alpha I.$$

Thus $S + I$ is a non-negative smooth decreasing function and therefore tends to a limit as $t \to \infty$. Also, it is not difficult to prove that the derivative of a smooth decreasing function must tend to zero, and this shows that

$$I_\infty = \lim_{t \to \infty} I(t) = 0.$$

Thus $S + I$ has limit S_∞.

Integration of the sum of the two equations of the model from 0 to ∞ gives

$$-\int_0^\infty (S(t)+I(t))'dt = S_0 + I_0 - S_\infty = N - S_\infty$$
$$= \alpha \int_0^\infty I(t)dt.$$

Division of the first equation by S and integration from 0 to ∞ gives

$$\log \frac{S_0}{S_\infty} = \frac{a}{N}\int_0^\infty I(t)dt$$
$$= \frac{a}{\alpha N}[N - S_\infty]$$
$$= \mathcal{R}_0 \left[1 - \frac{S_\infty}{N}\right].$$

This equation is called the *final size relation*. It gives a relation between the basic reproduction number and the size of the epidemic. Note that the final size of the epidemic, the number of members of the population who are infected over the course of the epidemic, is $N - S_\infty$. This is often described in terms of the attack rate $(1 - S_\infty/N)$. [Technically, the attack rate should be called an attack ratio, since it is dimensionless and is not a rate].

Then

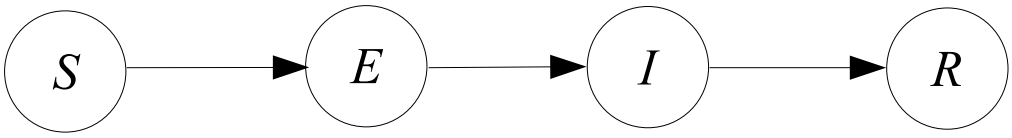
- The basic reproduction number is
$$\mathcal{R}_0 = \frac{a}{\alpha}.$$
- The final size relation determining $S_\infty = \lim_{t\to\infty} S(t)$ is
$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N}\right].$$

More complicated compartmental structures are possible

Exposed periods

An $SEIR$ model is

$$\begin{aligned} S' &= -a\frac{S}{N}I & S(0) &= S_0 \\ E' &= a\frac{S}{N}I - \kappa E & E(0) &= E_0 \\ I' &= \kappa E - \alpha I & I(0) &= I_0. \end{aligned}$$

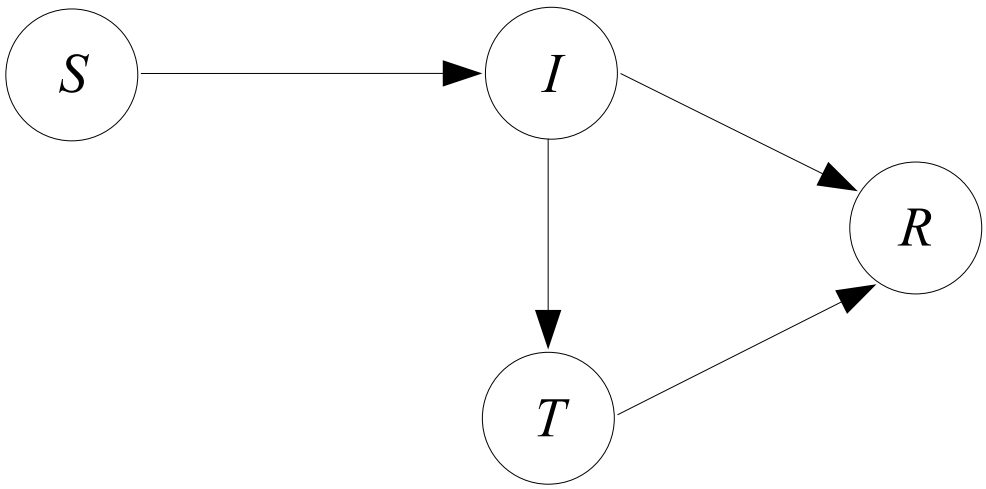


- The basic reproduction number and final size relation are the same as for the SIR model.

A treatment model

Select a fraction γ per unit time of infectives for treatment. Treatment reduces infectivity by a fraction δ and changes recovery rate to η. Then the $SITR$ model, where T is the treatment class, is

$$\begin{aligned} S' &= -a\frac{S}{N}[I + \delta T] \\ I' &= a\frac{S}{N}[I + \delta T] - (\alpha + \gamma)I \\ T' &= \gamma I - \eta T. \end{aligned}$$



- The reproduction number is

$$\mathcal{R}_0 = \frac{a}{\alpha + \gamma} + \frac{\gamma}{\alpha + \gamma}\frac{\delta a}{\eta}$$

- An infective in a totally susceptible population causes a new infections in unit time, and the mean time spent in the infective compartment is $1/(\alpha + \gamma)$. In addition, a fraction $\gamma/(\alpha + \gamma)$ of infectives are treated. While in the treatment stage the number of new infections caused in unit time is δa, and the mean time in the treatment class is $1/\eta$.
- The final size relation is

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N}\right].$$

There are many other epidemic models that can be formulated, all of which have the same final size relation. This suggests that a more general formulation is possible. An age of infection approach used by Diekmann, Heesterbeek, and Metz, establishes this.

THE AGE OF INFECTION MODEL

The general Kermack-McKendrick age of infection model for a population of constant total population size N is

$$\begin{aligned} S' &= -a\frac{S}{N}\varphi \\ \varphi(t) &= \varphi_0(t) + \int_0^t a\frac{S(t-\tau)}{N}\varphi(t-\tau)A(\tau)d\tau \\ &= \varphi_0(t) + \int_0^t [-S'(t-\tau)]A(\tau)d\tau. \end{aligned}$$

Here, a is the number of contacts sufficient to transmit infection per individual in unit time, $\varphi(t)$ represents the total infectivity of all infectious individuals at time $t, \varphi_0(t)$ represents the total infectivity at time t of all individuals who were already infected at time $t = 0$, and $A(\tau)$ represents the mean infectivity of all individuals who had been infected τ time units previously, including those who are no longer infectious. The function $A(\tau)$ is the product of the function representing the fraction of infected members still infected at infection age τ and the mean relative infectivity at infection age τ.

For this model,

$$\mathcal{R}_0 = a \int_0^\infty A(\tau) d\tau,$$

and the final size relation between the reproduction number and the number S_∞ of individuals who go through the epidemic without being infected is

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left[1 - \frac{S_\infty}{N}\right] - \frac{a}{N} \int_0^\infty [(N - S_0)A(t) - \varphi_0(t)] dt.$$

The initial term satisfies

$$\int_0^\infty [(N - S_0)A(t) - \varphi_0(t)] dt \geq 0.$$

If all initial infectives have infection-age zero at $t = 0$, $\varphi_0(t) = [N - S_0]A(t)$, and

$$\int_0^\infty [\varphi_0(t) - (N - S_0)A(t)] dt = 0.$$

Then the final size relation is

$$\log \frac{S_0}{S_\infty} = \mathcal{R}_0 \left(1 - \frac{S_\infty}{N}\right).$$

THE GENERAL $SEIR$ MODEL

The simple $SEIR$ epidemic model may be interpreted as an age of infection model with

$$A(\tau) = \frac{\kappa}{\kappa - \alpha}\left[e^{-\alpha t} - e^{-\kappa t}\right].$$

In order to derive this, we let $u(\tau)$ be the fraction of infected members with infection age τ who are not yet infective and $v(\tau)$ the fraction of infected members who are infective. Then the rate at which members become infective at infection age τ is $\kappa u(\tau)$, and we have

$$\begin{aligned} u'(\tau) &= -\kappa u(\tau), \qquad u(0) = 1 \\ v'(\tau) &= \kappa u(\tau) - \alpha v(\tau), \quad v(0) = 0 \end{aligned}$$

The solution of this system is

$$u(\tau) = e^{-\kappa\tau}, \quad v(\tau) = \frac{\kappa}{\kappa - \alpha}[e^{-\alpha\tau} - e^{-\kappa\tau}].$$

Thus we have

$$A(\tau) = v(\tau) = \frac{\kappa}{\kappa - \alpha}[e^{-\alpha\tau} - e^{-\kappa\tau}].$$

As another example of formulation of a model in age of infection form, we consider an $SEIR$ model with general distributions of stay in both the exposed and infectious period. Here, we consider the exposed period to be the time period from the acquisition of infection to the time when an individual can transmit infection, and the infective period to be the time period during which an individual can transmit infection.
Suppose the fraction of exposed individuals who are still in the exposed class s time units after being exposed is $P_E(s)$ and the fraction of individuals who are still in the infectious class s time units after entering the infectious class is $P_I(s)$, with $P_E(s), P_I(s)$ non-negative, non-increasing functions such that

$$P_E(0) = 1, \quad \int_0^\infty P_E(s)ds < \infty,$$

$$P_I(0) = 1, \quad \int_0^\infty P_I(s)ds < \infty.$$

Then P_E and P_I represent survival probabilities in the classes E and I respectively.

We assume that E_0 newly exposed members enter the exposed class at time $t = 0$. Then

$$S' = -a\frac{S}{N}I$$

$$E(t) = E_0 P_E(t) + \int_0^t [-S'(s)] P_E(t-s) ds.$$

Differentiation of the equation for $E(t)$ gives

$$E'(t) = E_0 P'_E(t) - S'(t) + \int_0^t [-S'(s)] P'_E(t-s) ds,$$

and this shows that the input to the infectious stage at time u is

$$-E_0 P'_E(u) - \int_0^u [-S'(s)] P'_E(u-s) ds.$$

and

$$I(t) = -E_0 \int_0^t P'_E(u) P_I(t-u) du$$

$$- \int_0^t \int_0^u [-S'(s)] P'_E(u-s) ds P_I(t-u) du.$$

The first term in this expression may be written as $I_0(t)$. Using interchange of the order of integration in an iterated integral, the second term may be simplified to give the equation

$$I(t) = I_0(t) + \int_0^t [-S'(s)] A_I(t-s) ds,$$

with

$$A_I(z) = -\int_0^z P_E'(z-v) P_I(v) dv,$$

Then the model is

$$\begin{aligned} S' &= -a\frac{S}{N} I \\ E(t) &= E_0 P_E(t) + \int_0^t [-S'(s)] P_E(t-s) ds \\ I(t) &= I_0(t) + \int_0^t [-S'(s)] A_I(t-s) ds, \end{aligned}$$

which is in age of infection form with $\Phi = I$ and $A(z) = A_I(z)$.

Then

$$\begin{aligned}\mathcal{R}_0 &= a\int_0^\infty A(z)dz \\ &= -a\int_0^\infty \int_0^z P_E'(z-u)P_I(u)dudz \\ &= a\int_0^\infty P_I(u)du,\end{aligned}$$

using $-\int_0^\infty P_E'(v)dv = P_E(0) - P_E(\infty) = 1$.

In all these calculations we have assumed that there are no disease deaths. If there are disease deaths, the final size relation becomes an inequality and it is possible to estimate the final size of the epidemic by numerical simulations.

We have assumed homogeneous mixing of the population in all of these calculations. It is possible to extend the results to include heterogeneity of mixing, and the extension may be found in the book by Brauer, Castillo-Chavez, and Feng.

The deeper information Faust sought
Could not from the Devil be bought
But now we are told
By theorists bold
That all you need is R nought

R.M.May (Baron May of Oxford)

REFERENCES

W.O. Kermack & A.G. McKendrick, Contributions to the mathematical theory of epidemics, Part 1, Proc. Roy. Soc. Edinburgh A, 15(1927), 700–721. [Reprinted as Bull. Math. Biol. 53(1991), 33–55.]

O. Diekmann, J.A.P. Heesterbeek, & J.A.J. Metz, The legacy of Kermack and McKendrick, in Epidemic Models: Their Structure and Relation to Data, D. Mollison, ed. Cambridge University Press (1995), pp. 95–115.

F, Brauer, C. Castillo-Chavez, & Z. Feng, Mathematical Models in Epidemiology, Springer (2019), Chapter 5.