Let $R_A$ denote the resistant density under aggressive treatment and let $R_C$ denote the resistant density under containment. Because we assume that aggressive treatment immediately removes the entire drug-sensitive population, the expansion of the resistant density under aggressive treatment is described by

$$
\dot{R}_A = (1 - c_I)r R_A(1 - (1 + c_C)\delta R_A) - \mu R_A. \quad (S.1)
$$

Assuming that the immune response $\mu$ is constant the solution to this equation is

$$
R_A(t) = \frac{\left(1 - \frac{\mu}{(1-c_I)r}\right) R(0) \exp \left[\left((1 - c_I)r - \mu\right)t\right]}{\left(1 - \frac{\mu}{(1-c_I)r}\right) + R(0)(1 + c_C)\delta \left(\exp \left[\left((1 - c_I)r - \mu\right)t\right] - 1\right)}, \quad (S.2)
$$

where $R(0)$ is the resistant density at the start of the management period. If aggressive treatment fails at time $t = t_A$ then $R_A(t_A) = P_{\text{max}}$. Substituting this equality into Equation (S.2) gives,

$$
t_A = \frac{1}{(1 - c_I)r - \mu} \ln \left[\frac{P_{\text{max}}(1 - c_I)r(1 - R(0)(1 + c_C)\delta) - \mu}{R(0)(1 - c_I)r(1 - P_{\text{max}}(1 + c_C)\delta) - \mu}\right]. \quad (S.3)
$$

Under containment the expansion of the resistant density is described by

$$
\dot{R}_C = (1 - c_I)r R_C(1 - (1 + c_C)\delta P_{\text{max}}) - \mu R_C + \epsilon r (1 - \delta P_{\text{max}})(P_{\text{max}} - R_C).
$$

Assuming that the immune response $\mu$ is constant the solution to this equation is

$$
R_C(t) = (B + R(0)) \exp \left[\left((1 - c_I)r(1 - (1 + c_C)\delta P_{\text{max}}) - \epsilon r(1 - \delta P_{\text{max}}) - \mu\right)t\right] - B, \quad (S.4)
$$

where $R(0)$ is the resistant density at the beginning of the management period and

$$
B = \frac{\epsilon r (1 - \delta P_{\text{max}}) P_{\text{max}}}{(1 - c_I)r(1 - (1 + c_C)\delta P_{\text{max}}) - \epsilon r(1 - \delta P_{\text{max}}) - \mu}.
$$

If containment fails at time $t_C$ then $R_C(t_C) = P_{\text{max}}$. Substituting this equality into Equation (S.4) gives,

$$
t_C = \frac{1}{(1 - c_I)r D} \ln \left[\frac{P_{\text{max}}}{R(0)} \left(\frac{(1 - c_I)D + \epsilon (1 - \delta P_{\text{max}}) P_{\text{max}}}{(1 - c_I)D + \epsilon (1 - \delta P_{\text{max}}) R(0)}\right)\right], \quad (S.5)
$$

where

$$
D = 1 - \frac{\mu}{(1-c_I)r} - (1 + c_C)\delta P_{\text{max}} - \frac{\epsilon}{(1-c_I)(1-\delta P_{\text{max}})}.
$$
Therefore, from Equation (S.3) and (S.5) we have,

\[ \frac{t_C}{t_A} = \frac{1}{(1-c_I)rD} \ln \left[ \frac{P_{\text{max}}}{R(0)} \left( \frac{(1-c_I)D+\epsilon(1-\delta P_{\text{max}})}{(1-c_I)D+\epsilon(1-\delta P_{\text{max}})} \right) \right]. \]  

(S.6)

Now, from Equation (S.1), \( \dot{R}_A = 0 \) when \( R_A = \frac{1}{1+c_C} \delta \left( 1 - \frac{\mu}{(1-c_I)r} \right) \). This is the self-limiting density discussed in the main text, \( R_{\text{lim}} = \frac{1}{1+c_C} \delta \left( 1 - \frac{\mu}{(1-c_I)r} \right) \).

Now define the variables: \( \tilde{R}_{\text{balance}} = \frac{R_{\text{balance}}}{R_{\text{lim}}} \), \( \tilde{P}_{\text{max}} = \frac{P_{\text{max}}}{R_{\text{lim}}} \) and \( \tilde{R}_0 = \frac{R(0)}{R_{\text{lim}}} \). Substituting these variables into Equation (S.6) allows us to express the ratio \( \frac{t_C}{t_A} \) in terms of how the three pathogen densities \( R(0), P_{\text{max}} \) and \( R_{\text{balance}} \) compare to \( R_{\text{lim}} \). That is,

\[ \frac{t_C}{t_A} = \frac{1}{1 - \tilde{P}_{\text{max}} - \tilde{R}_{\text{balance}}} \ln \left[ \frac{\tilde{P}_{\text{max}}}{\tilde{R}_0} \left( \frac{(1-\tilde{P}_{\text{max}})}{(1-\tilde{P}_{\text{max}})+\tilde{R}_{\text{balance}}(\frac{\tilde{P}_{\text{max}}}{\tilde{R}_0}-1)} \right) \right]. \]  

(S.7)

Equation (S.7) is the equation that was used to generate Fig. 3 of the main text. In Fig. 3 the acceptable burden was allowed to vary from 10% to 80% of \( R_{\text{lim}} \) (i.e, \( \tilde{P}_{\text{max}} \in [0, 0.8] \)) and the resistant density at the start of the management period was allowed to vary from the balance threshold to 80% of \( R_{\text{lim}} \) (i.e, \( \tilde{R}_0 \in [\tilde{R}_{\text{balance}}, 0.8] \)). Here we reproduce Fig. 3 from the main text (S4 Fig; Panel A) but also include the possibility that the starting resistant density is below the balance threshold (S4 Fig; Panel B). Together, Panel A and Panel B of S4 Fig allow the starting resistant density to vary from \( 10^{-8}\% \) to 80% of \( R_{\text{lim}} \) (i.e, \( \tilde{R}_0 \in [10^{-10}, 0.8] \)). These choices cover a wide range of possibilities.

The chosen parameter range, \( \tilde{R}_{\text{balance}} \in [0, 0.01] \) requires a bit more explanation. Note that,

\[ \tilde{R}_{\text{balance}} = \frac{\epsilon (1-\delta P_{\text{max}})}{(1-c_I)(1+c_C)\delta} \frac{(1+c_C)\delta}{1 - \frac{\mu}{(1-c_I)r}}, \]

\[ = \frac{\epsilon (1-\delta P_{\text{max}}) r}{(1-c_I)r - \mu}, \]

\[ \leq \frac{\epsilon}{1 - c_I} \frac{1}{1 - \frac{\mu}{(1-c_I)r}}. \]
The quantity \( \frac{(1-c_I)r}{\mu} \) can be thought of as the expected number of progeny produced by an average resistant pathogen (assuming there is no competition). If \( \frac{(1-c_I)r}{\mu} \geq 1.1 \) and the reduction in intrinsic replication \( (1 - c_I) \geq 0.1 \) then we have,

\[
\tilde{R}_{\text{balance}} \leq \epsilon 110.
\]

Under these assumptions \( \tilde{R}_{\text{balance}} \) will be less than 0.01 provided the probability of mutation is not too large (i.e., \( \epsilon < 9.1 \times 10^{-5} \)). Alternatively, if \( \frac{(1-c_I)r}{\mu} \geq 2 \) and the reduction in intrinsic replication \( (1 - c_I) \geq 0.1 \) then \( \tilde{R}_{\text{balance}} \) will be less than 0.01 provided the mutation rate \( \epsilon \) is less than \( 5 \times 10^{-4} \).

We can also use this example to gain some insight into the situation when \( R(0) < \tilde{R}_{\text{balance}} < P_{\text{max}} \) (i.e., the cases depicted in Figure 2 C-D in the main text). In particular we will show that there is a resistant density \( \tilde{R}^*(0) \) such that aggressive treatment is best whenever \( R(0) < \tilde{R}^*(0) \) and containment is best whenever \( R(0) > \tilde{R}^*(0) \).

Containment will be at least as good as aggressive treatment whenever \( t_C \geq t_A \). By Equation (S.7) this will occur whenever

\[
\tilde{R}_0 \leq f(\tilde{R}_0)
\]

where

\[
f(\tilde{R}_0) = \hat{P}_{\text{max}} \left[ \tilde{R}_0 \left( 1 - \hat{P}_{\text{max}} \right)/\hat{P}_{\text{max}} \left( 1 - \tilde{R}_0 \right) \right]^A - \tilde{R}_{\text{balance}} \left( \hat{P}_{\text{max}} - \tilde{R}_0 \right)/1 - \hat{P}_{\text{max}}
\]

and \( A = 1 - \hat{P}_{\text{max}} - \tilde{R}_{\text{balance}} \). We will now show that the equality in Equation (S.8) can hold for at most one value of \( \tilde{R}_0 \in [0, \hat{P}_{\text{max}}] \). First note that if \( \tilde{R}_0 = 0 \) then \( f(0) = -\tilde{R}_{\text{balance}} \hat{P}_{\text{max}}/(1 - \hat{P}_{\text{max}}) < 0 \) and hence (assuming that \( \epsilon \neq 0 \)) Equation (S.8) is not satisfied when \( \tilde{R}_0 = 0 \). Additionally, when \( \tilde{R}_0 = \hat{P}_{\text{max}} \) we have that \( f(\hat{P}_{\text{max}}) = \hat{P}_{\text{max}} \) and \( \partial f/\partial \tilde{R}_0 \bigg|_{\hat{P}_{\text{max}}} = 1 \). Note also that if \( \tilde{R}_0 = \hat{P}_{\text{max}} - \epsilon \) then

\[
f(\hat{P}_{\text{max}} - \epsilon) = \hat{P}_{\text{max}} \left[ (\hat{P}_{\text{max}} - \epsilon)(1 - \hat{P}_{\text{max}})/\hat{P}_{\text{max}}(1 - \hat{P}_{\text{max}} + \epsilon) \right]^A - \tilde{R}_{\text{balance}} \left[ \epsilon/(1 - \hat{P}_{\text{max}}) \right] \doteq T(\epsilon)
\]

and

\[
\frac{\partial T}{\partial \epsilon} = -\hat{P}_{\text{max}} A \left[ (\hat{P}_{\text{max}} - \epsilon)(1 - \hat{P}_{\text{max}})/\hat{P}_{\text{max}}(1 - \hat{P}_{\text{max}} + \epsilon) \right]^{A-1} \frac{1 - \hat{P}_{\text{max}}}{\hat{P}_{\text{max}}(1 - \hat{P}_{\text{max}} + \epsilon)^2} - \frac{\tilde{R}_{\text{balance}}}{1 - \hat{P}_{\text{max}}} < 0
\]

and so as \( \tilde{R}_0 \) approaches \( \hat{P}_{\text{max}} \) from the left, the function \( f \) decreases to approach \( \hat{P}_{\text{max}} \). This means that \( f \) must cross the \( \tilde{R}_0 = \tilde{R}_0 \) line an odd number of times.
If \( f \) crosses the \( \tilde{R}_0 = \tilde{R}_0 \) line only once then this proves the claim. In particular the crossing occurs when \( R(0) = R^*(0) \). In other words, \( R^*(0) \) is implicitly defined when \( R(0) = R^*(0) \) and equality holds in Equation (S.8).

Suppose, on the other hand, that \( f \) crosses the \( \tilde{R}_0 = \tilde{R}_0 \) line more than once. Let \( \tilde{R}_1 \) and \( \tilde{R}_2 \) denote the values of \( \tilde{R}_0 \) at the first two crossings. Then we must have that \( \frac{\partial f}{\partial \tilde{R}_0} \bigg|_{\tilde{R}_1} > \frac{\partial f}{\partial \tilde{R}_0} \bigg|_{\tilde{R}_2} \) and equality holds in Equation (S.8).

\[
\frac{\partial f}{\partial \tilde{R}_0} \bigg|_{\tilde{R}_1} = 1 \quad \text{and} \quad \frac{\partial f}{\partial \tilde{R}_0} \bigg|_{\tilde{R}_2} < 1.
\]

Since \( \frac{\partial f}{\partial \tilde{R}_0} \) is continuous this means that there must be a \( \tilde{R}_3 \in (\tilde{R}_1, \tilde{R}_2) \) such that \( \frac{\partial f}{\partial \tilde{R}_0} \bigg|_{\tilde{R}_3} = 1 \).

In other words,

\[
\frac{\partial f}{\partial \tilde{R}_0} \bigg|_{\tilde{R}_3} = \tilde{P}_\text{max} A \left[ \frac{\tilde{R}_3(1 - \tilde{P}_\text{max})}{\tilde{P}_\text{max}(1 - \tilde{R}_3)} \right]^A \left[ \frac{1}{\tilde{R}_3(1 - \tilde{R}_3)} \right] + \frac{\tilde{R}_\text{balance}}{1 - \tilde{P}_\text{max}} = 1. \quad (\text{S.9})
\]

After some simplification Equation (S.9) becomes

\[
\left[ \frac{\tilde{P}_\text{max}}{\tilde{R}_3} \right]^{1-A} = \left[ \frac{1 - \tilde{R}_3}{1 - \tilde{P}_\text{max}} \right]^{1+A}. \quad (\text{S.10})
\]

Substituting \( \tilde{R}_3 = \gamma \tilde{P}_\text{max} \) for some \( \gamma \in (0, 1) \) into Equation (S.10) results in

\[
\tilde{P}_\text{max} = \frac{1 - \left( \frac{1}{\gamma} \right) \frac{1-A}{1+A}}{\gamma - \left( \frac{1}{\gamma} \right) \frac{1-A}{1+A}} \doteq B(\gamma). \quad (\text{S.11})
\]

But

\[
\frac{\partial B}{\partial \gamma} = \frac{\frac{1-A}{1+A} \left( \frac{1}{\gamma} \right) \frac{1-A}{1+A} \left( 1 - \frac{1}{\gamma} \right) \frac{1-A}{1+A} - 1 \left( \gamma - \left( \frac{1}{\gamma} \right) \frac{1-A}{1+A} \right)^2}{< 0,
\]

which implies that there is at most one \( \gamma \) that will satisfy Equation (S.11). Therefore, \( f \) cannot cross the \( \tilde{R}_0 = \tilde{R}_0 \) line more than once.