Combating Pertussis Resurgence Supplementary Material

Maria A Riolo, Pejman Rohani

S1 Calculating strategy fitness

For simplicity, we used what the WHO refers to as "no-frills DALYs" (1), calculated as the expected duration of illness plus, in the case of a death, the expected years of life lost (average lifespan minus the age at death). Although our transmission model does not account for mortality, for the purposes of assessing costs we assume a case fatality rate of 0.2% in infants one year or younger, of 0.04% in children 1-4 years of age, and no mortality in older age groups (2). In all age groups, we assume that both symptoms and infectiousness last, on average, for 15 days. Thus, under our parameterization, a case in a six month old costs $\frac{15}{365} + 0.002 * 74.5 \approx 0.19$ DALYs, while a case in an adult costs only $\frac{15}{365} \approx 0.041$ DALYs.

Throughout the study, we used a cost of \$33 per dose of vaccine and \$50000 per DALY. However, it is worth noting that in our current model changing the cost per DALY is equivalent to changing the cost of vaccine (e.g. an assumption of \$66 per dose of vaccine and \$100000 per DALY would produce identical results).

S2 GA details

Point mutations are performed by incrementing the value of a single gene by ϵ where ϵ is drawn uniformly at random from the interval $(-0.1, 0.1)$.

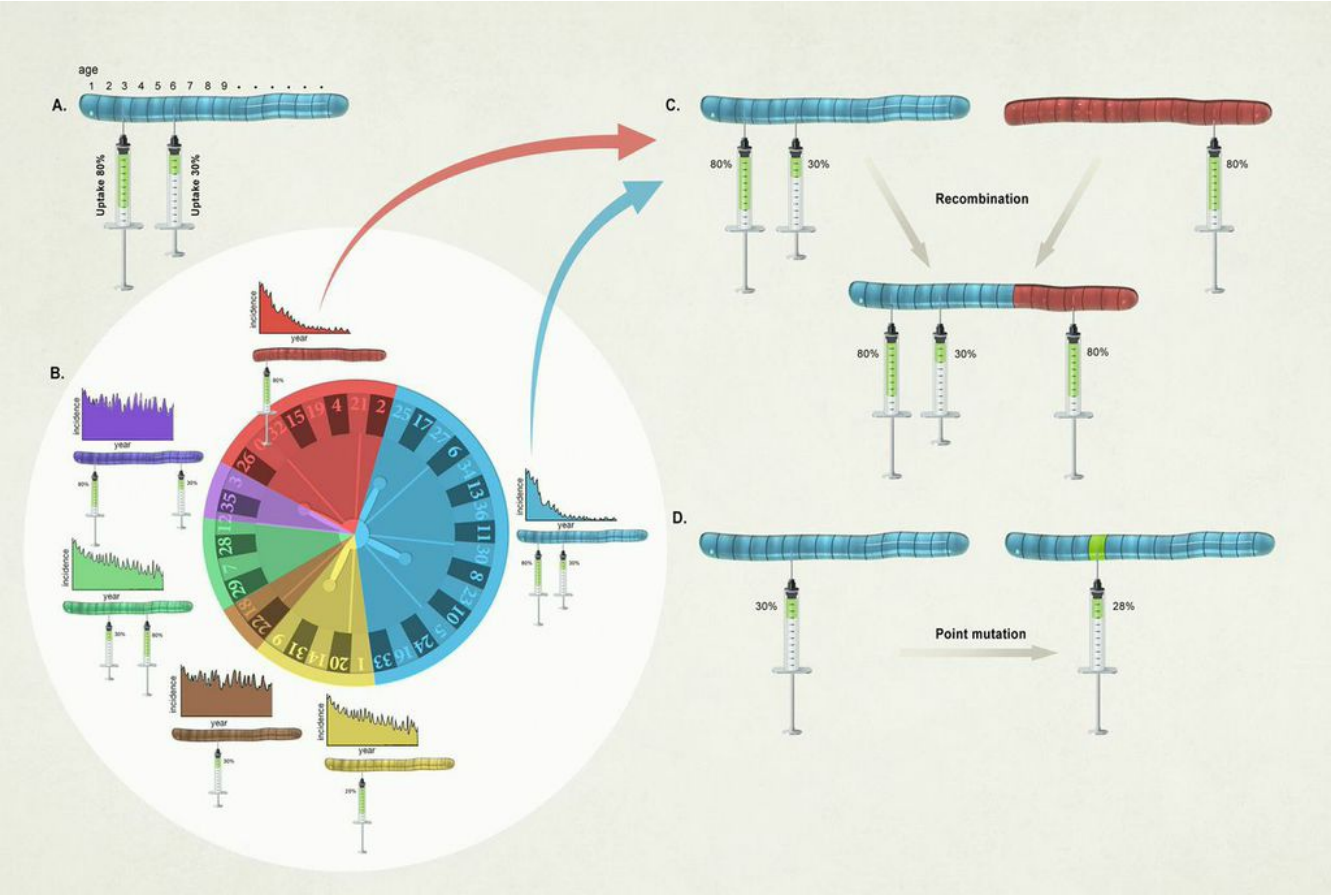


Figure S-1: Schematic showing the operation of the genetic algorithm on booster strategies. Panel A illustrates a single chromosome representation of a booster strategy, with each site on the chromosome determining the vaccine uptake of an age group. Panel B illustrates competing strategies of varying fitness (as determined by a simulation of transmission dynamics under the booster schedule). Note that although the relative fitness of strategies is illustrated as a roulette wheel, we are not using fitness proportional or rank proportional selection but rather tournament selection with a 10% chance of letting the less fit contender win. Panel C illustrates a recombination of two parent schedules. Panel D illustrates a point mutation changing the scheduled vaccine uptake in a single age group.

S3 Transmission details

Individuals are categorised by yearly age groups up to seventy five, with an additional category for infants under five months of age (i.e. too young to have received at least two doses of pertussis vaccine under the pre-1990 vaccine schedule in the UK). For convenience, these age categories are labeled with indices starting from zero, so that N_0 designates the number of 0–5 month olds, N_1 is the number of 6 month–1 year

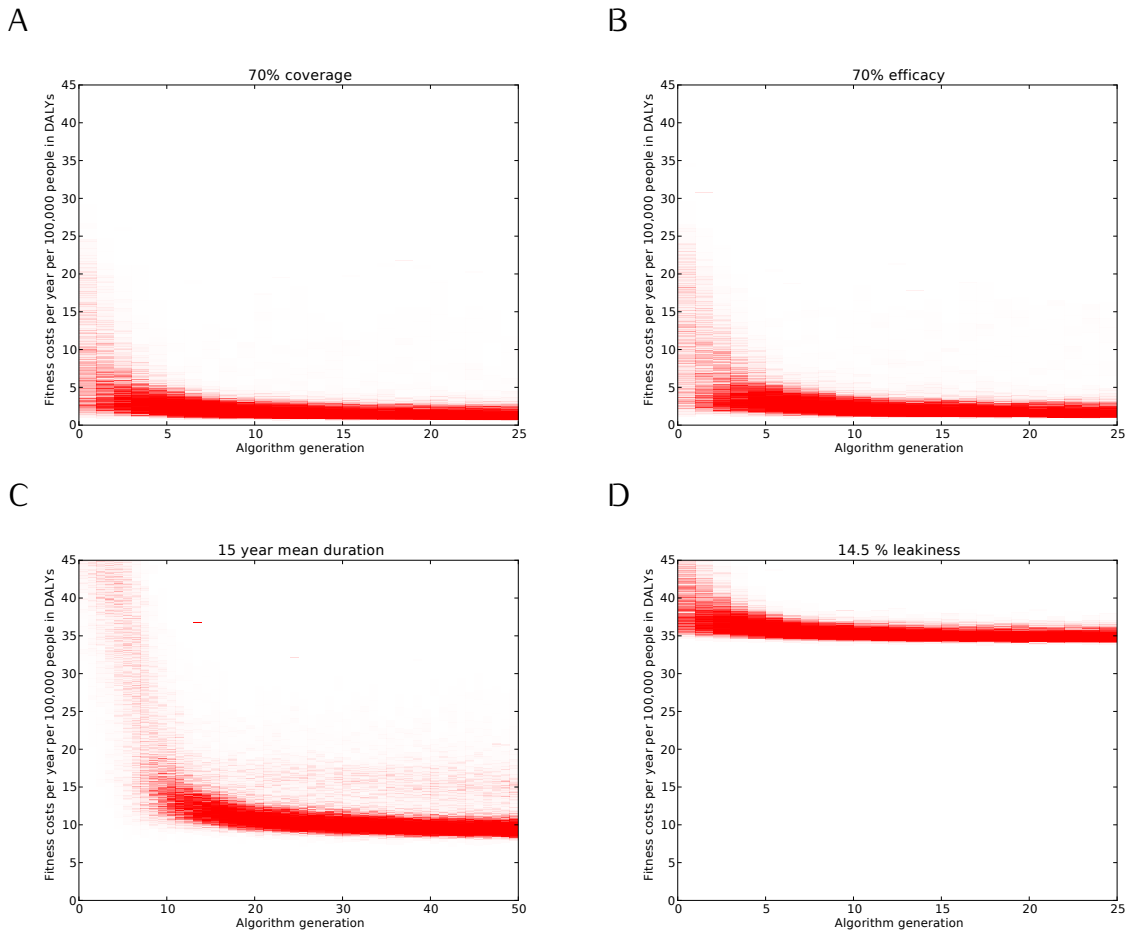


Figure S-2: Fitness costs by algorithm generation. Density maps of the fitness costs (combined vaccination effort and disease burden) of the 2000 strategies in each generation of the algorithm for (A) Scenario I: 70% infant coverage, (B) Scenario II: 70% vaccine efficacy (C) Scenario III: 45 year mean duration of vaccine derived immunity in (G,H,I), and (D) Scenario IV: 14.5% leakiness. In each case, darker color indicates a higher density of strategies.

olds, N_2 is the number of 1–2 year olds, and so on up to N_{75}. The total population is designated by N with no subscript.

All ages except for 0–5 mo olds are tracked as yearly cohorts, the cutoff occurring at the start of each school year. Newborns are assumed to age continuously at rate $a = \frac{12}{5}$ yr^{-1}, corresponding to the assumption that a newborn spends on average 5 mo in the 0–5 mo age category. Susceptible newborns aging at time t have probability

$u(t)\,e$ of being protected by vaccination, where $u(t)$ is the vaccine uptake at time t and e is the vaccine efficacy.

The stochastic model is initialized with conditions from the pre-vaccine era and proceeds by updating the numbers of individuals in each age category who are susceptible, latently infected, infectious, recovered, or vaccinated, respectively. Those in the recovered class are protected from infection for their remaining lifespan while those in the vaccinated classes are partially protected for a period, the duration of which is a random variable, as we detail below. The dynamics of susceptible (S_i), exposed (E_i), infectious (I_i), recovered (R_i), and vaccinated (V_i and V_i') individuals in age group i are given by:

$$\begin{aligned}
\frac{dS_i}{dt} &= w_V V_i' - \lambda_i(t) S_i + && \text{(S1)}\\
&\quad (bN - aS_0)\delta_{i,0} + a(1 - eu(t))S_0\delta_{i,1} && \text{(S2)}\\
\frac{dE_i}{dt} &= \lambda_i(t)S_i + \epsilon\lambda_i(t)(V_i + V_i') - \gamma E_i + aE_0(\delta_{i,1} - \delta_{i,0}) && \text{(S3)}\\
\frac{dI_i}{dt} &= \gamma E_i - rI_i + aI_0(\delta_{i,1} - \delta_{i,0}) && \text{(S4)}\\
\frac{dR_i}{dt} &= rI_i + aR_0(\delta_{i,1} - \delta_{i,0}) && \text{(S5)}\\
\frac{dV_i}{dt} &= eu(t)aS_0\delta_{i,1} - \epsilon\lambda_i(t)V_i - w_V V_i + aV_0(\delta_{i,1} - \delta_{i,0}) && \text{(S6)}\\
\frac{dV_i'}{dt} &= w_V V_i - w_V V_i' - \epsilon\lambda_i(t)V_i' + aV_0'(\delta_{i,1} - \delta_{i,0}) && \text{(S7)}
\end{aligned}$$

where $\delta_{i,j}$ is the Kronecker delta, which is one if i and j are equal and zero otherwise.

Age group i gains susceptible members through immune waning and, in the first two age categories, births and the aging of susceptible newborns, respectively. The birth rate $b = \frac{1}{75}$ yr^{-1} is chosen to keep the population steady given the 75 year lifespan. Individuals leave the susceptible category by becoming exposed or, in the

infant category, aging.

The force of infection acting on age group i at time t is

$$\lambda_i(t) = q\Sigma_k F_{hk}(t) c_{hk} \frac{\tilde{I}_k}{\tilde{N}_k}$$

where c_{hk} is the average rate (in contacts per year) at which an individual who is $5h$–$(5h+5)$ years old makes contact with $5k$–$(5k+5)$ year olds and q is the probability of infection given exposure. The number of infected individuals and total individuals in the k^{th} five-year age block are denoted by

$$\tilde{I}_k = \Sigma_{5k<i\leq 5k+5}\ I_i \quad \text{and} \quad \tilde{N}_k = \Sigma_{5k<i\leq 5k+5}\ N_i$$

with I_0 and N_0 included in the calculation of $\tilde{I}_0$ and $\tilde{N}_0$, respectively.

Values of c_{ij} and q were adopted from an earlier study. (3) In particular, rates of daily contacts c_{ij} were obtained from the POLYMOD study,(4) (see Figure 1A) and q was fixed at 4% as in Ref. 3, leading to a pre-vaccine era mean age of first infection consistent with historical estimates.

To capture the strong seasonality in children's social contacts,(5) we incorporated an age-dependent seasonal forcing term $F_{hk}(t)$ based on school holidays. For $0 < h < 3$ or $0 < k < 3$ (*i.e.* when either party is 5–15 years old), $F(t) = \kappa(1 \pm 0.2)$, with $+$ when school is in session and $-$ during school holidays. Because there are more school days than holidays, we use the normalization constant κ to ensure that $F(t)$ has a mean of 1.0 over the whole year. The school holidays used in our simulations were July 19 – September 8, October 28 – November 3, December 21 – January 10, and April 10 – 25. If neither party is 5–15 years old, $F_{hk}(t) = 1$, leaving the contact rate unaltered year round.

We assume that vaccination occurs at six months of age. In scenarios where vaccine protection wanes over time, we model two stages of vaccine derived immunity, V_i and

V_i', so that the duration of immunity is gamma distributed with shape parameter two. The waning rate is given by $w_V = \frac{2.0}{d_R}$ yr^{-1} where d_R denotes the mean duration of infection derived immunity. When vaccine protection is leaky, vaccinated individuals have a reduced but nonzero chance of becoming exposed upon risky contact, where the factor by which vaccination reduces the risk of infection is ϵ.

It is worthwhile to note that, although there is some evidence that the duration of pertussis immunity is likely quite variable between individuals (6), pertussis immunity remains poorly understood enough that the precise choice of the distribution of durations of immunity is somewhat arbitrary. However, the simplest "memory free" distribution, an exponential distribution, would result in a very large fraction of individuals with extremely fast waning. We chose gamma-2 distributed immunity in order to avoid this phenomenon with incurring only minimal additional computational costs, and because it is consistent with the evidence of high variation between individuals.

In our simulations, individuals exposed to pertussis become infectious after an average of 8 days ($\gamma = \frac{365}{8}$ yr^{-1}) and the infectious period lasts 15 days on average ($r = \frac{365}{15}$ yr^{-1}), again as in the model of (3).

For the initial conditions of our genetic algorithm runs, we used the population at the end of the 150th year of a run with lifelong immunity and a total population of approximately sixty-three million individuals, where vaccination begins after 100 years at half of its eventual level, ramps up linearly for 20 years, and remains at constant coverage for the final 30 years. All simulations used in the GA were run for 50 years. At the end of each year, for each age category we recorded the population, number of susceptibles, number of successful vaccinations, and number of cases.

Because of the computational cost of using Gillespie's direct algorithm with so many age-categories, we use a multinomial τ–leaping method,(7) in which we move forward

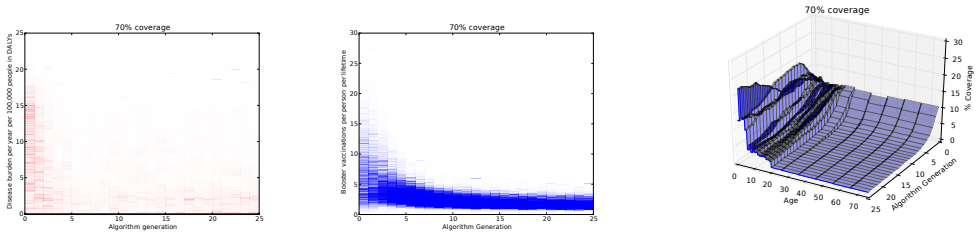
by a fixed time step τ and determine the set of events that occurred during that time step. All simulations presented in this paper use $\tau = \frac{1}{365}$ yr.

At each step, we consider all the ways an individual can leave each category as a set of competing events. For a sufficiently small time step τ, we can approximate an individual's probability of leaving a category as the total rate at which individuals leave multiplied by the length of the time step. For example, susceptible newborns leave the category by aging at rate a or by becoming exposed at rate $\lambda_i(t)$, so each of the S_0 susceptible newborns has probability $(a+\lambda_0(t))\tau$ of leaving the category. We determine the total number who leave the category drawing from a binomial distribution, $X \sim B(S_0, p_0)$ in our example. The expected fraction of these individuals leaving via each event is proportional to that event's rate. Continuing our example, the X individuals leaving the susceptible newborn category are aging with probability $\frac{a}{a+\lambda_0(t)}$ and have been exposed with probability $\frac{\lambda_0(t)}{a+\lambda_0(t)}$, so we draw the numbers of aging and exposure events (X_a, X_E) from a multinomial distribution with X trials and probabilities ($\frac{a}{a+\lambda_0(t)}$, $\frac{\lambda_0(t)}{a+\lambda_0(t)}$). For aging infants, we perform one more binomial draw with probability $eu(t)$ to determine how many aging infants are successfully vaccinated.

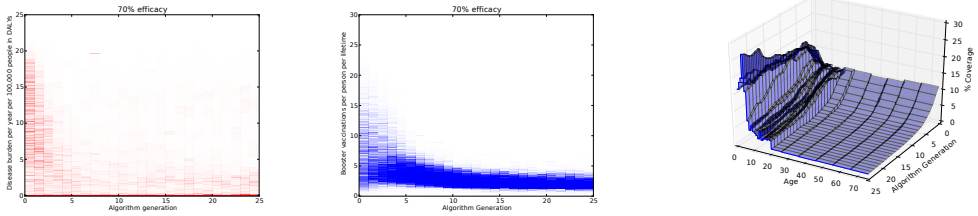
The number of births, which don't deplete any population categories, is determined by drawing from binomial distribution $B(N, b\tau)$, where b is the per capita annual birth rate. Because this is a discrete stochastic model, we also include an immigration rate of one infected individual per year, uniformly distributed among age categories to help distinguish between stable eradication and a chance extinction in an easily re–invaded population. Once the set of events taking place has been determined, the whole population is updated according to those events and the time t is incremented by τ.

Figure S-3: GA results by generation. Each row of panels shows results from a different scenario, with 70% infant coverage in (A, B, C), 70% vaccine efficacy in (D, E, F), 14 year mean duration of vaccine derived immunity in (G, H, I), and 14.5% leakiness in (J, K, M). Panels (A, D, G, J) show the disease burden (measured in DALYs) under strategies in each generation, with the algorithm generation on the horizontal axis, the disease burden on the vertical axis, and the intensity of the color indicating the fraction of strategies in the current generation with the given disease burden. Panels (B, E, H, K) show the number of booster shots administered per person per lifetime under strategies in each generation. Again, with the algorithm generation on the horizontal axis, the number of boosters per person per lifetime on the vertical axis, and the intensity of the color indicating the fraction of strategies in the current generation which use the given number of boosters. Panels (B, D, F, H) show the average schedule during each algorithm generation, plotted with vaccine coverage on the vertical axis, age along the horizontal axis, and algorithm generation on the axis pointing "into" the page. The average schedules were constructed by taking the mean coverage in each age cohort among each generation of booster schedules.

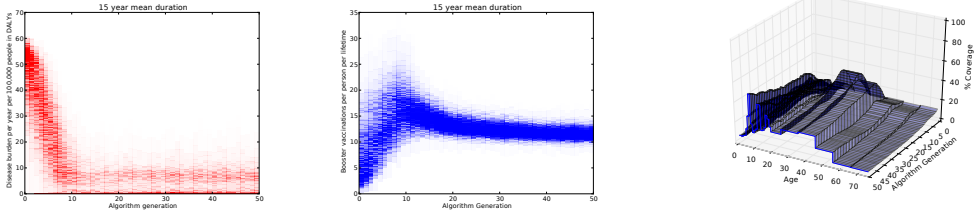
A. Disease Burden B. Vaccination Costs C. Average Strategy



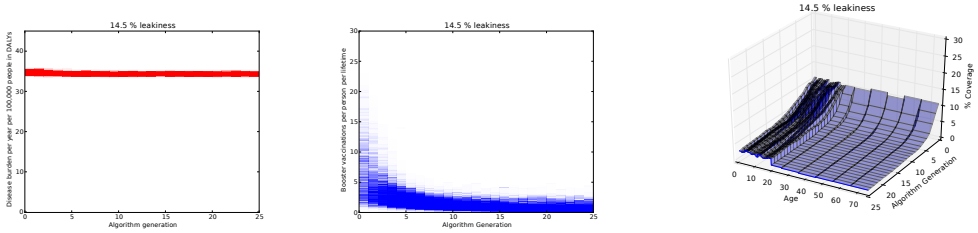
D. E. F.



G. H. I.



J. K. L.



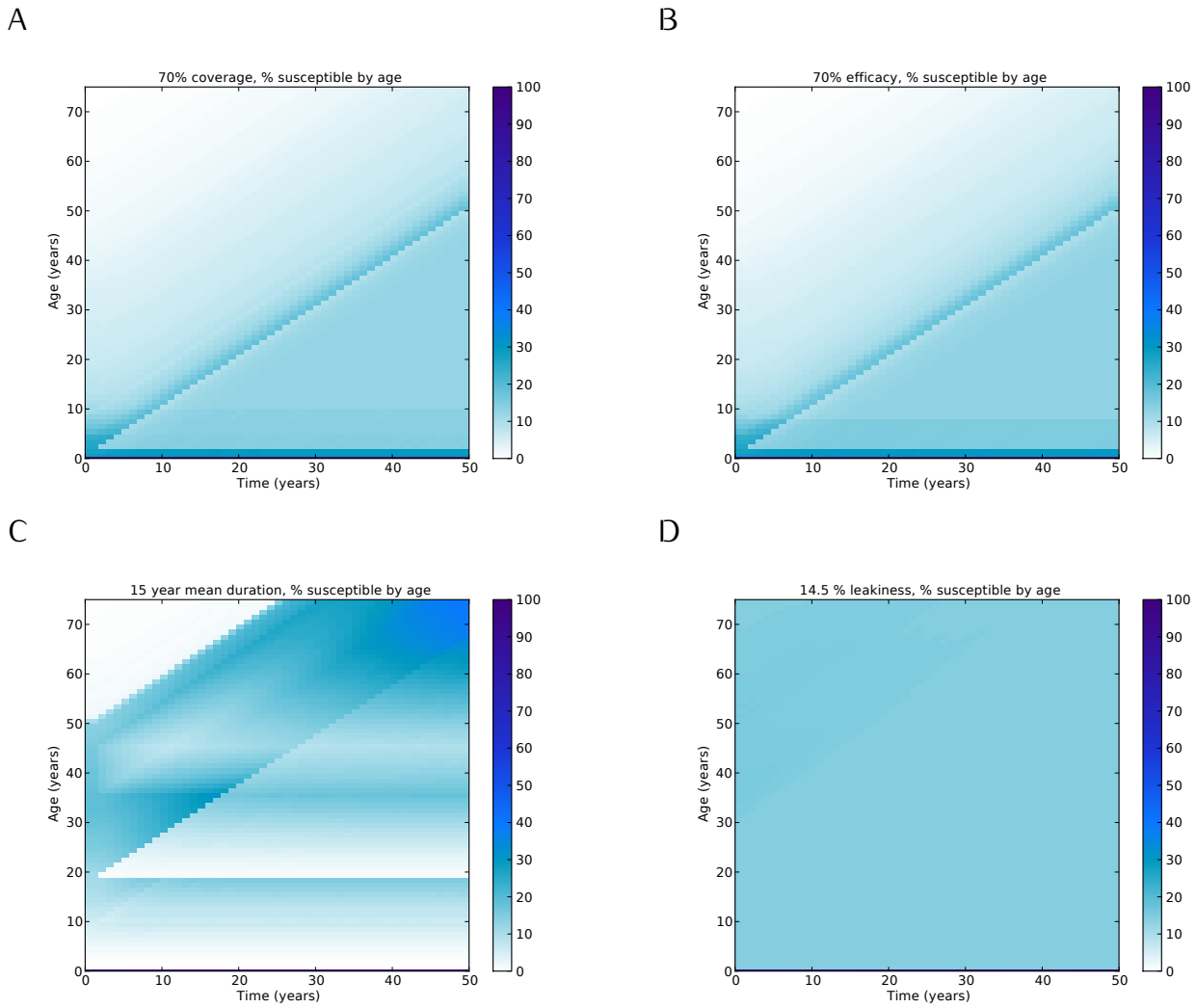
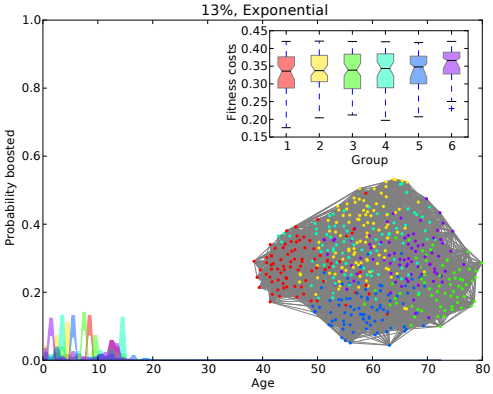


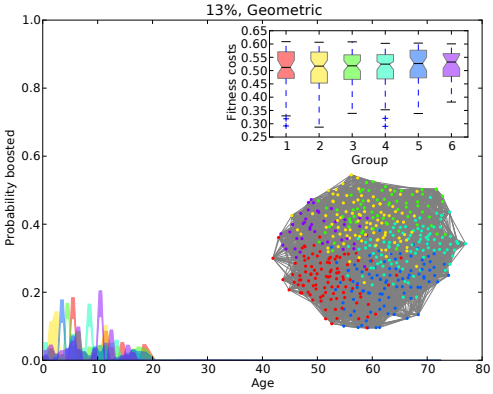
Figure S-4: Susceptibilty by age under the most successful booster schedules for 70% infant coverage (A), 70% efficacy (B), a 45 year mean duration of immunity (C), leaky immunity preventing 85.5% of infections (D), plotted during the first fifty years following the introduction of booster vaccinations. Color intensity indicates the percentage of individuals in each age group who are susceptible at each time. In the leaky immunity case, 14.5% of vaccinated individuals are included in this calculation.

Figure S-5: Leaky vaccines with reduced leakiness. Leakiness decreases either exponentially (left column) or geometrically (right column) with repeated vaccination. In the case of 13% leakiness, an exponential decrease in leakiness would mean that an individual who has been vaccinated n times ($n \geq 1$) would be 0.13^n times as likely as a susceptible individual to become infected in the case of risky contact. With geometrically decreasing leakiness, the same individual would instead be $\frac{0.13}{n}$ times as likely as a susceptible individual to become infected in the case of risky contact.

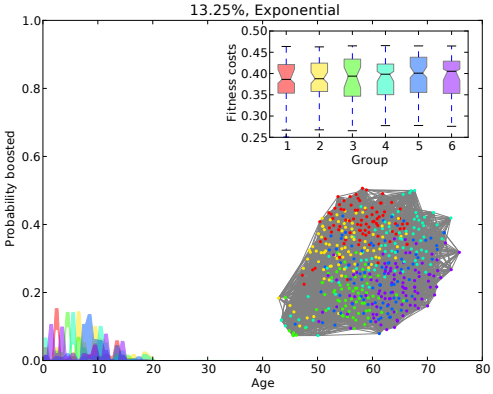
A 13% leaky, exponential decrease



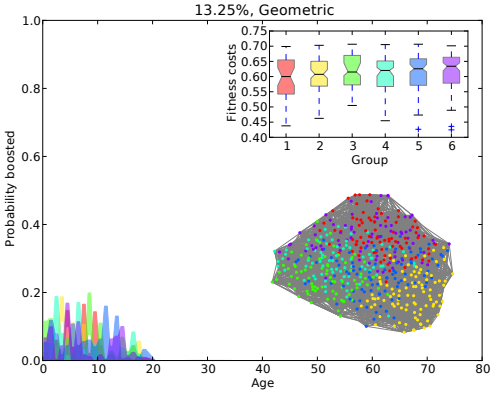
B 13% leaky, geometric decrease



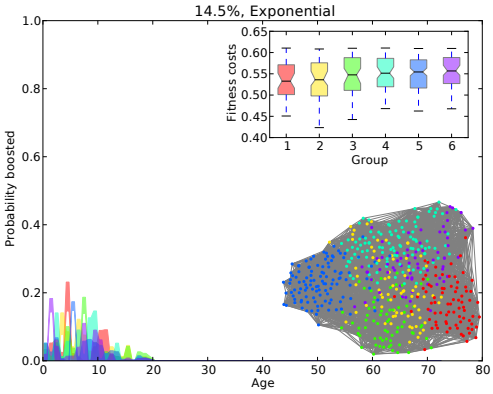
C 13.25% leaky, exponential decrease



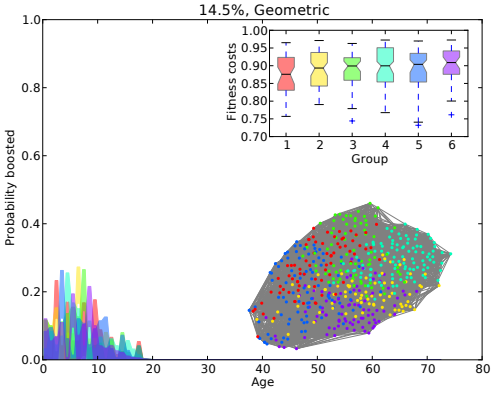
D 13.25% leaky, geometric decrease



E 14.5% leaky, exponential decrease



F 14.5% leaky, geometric decrease



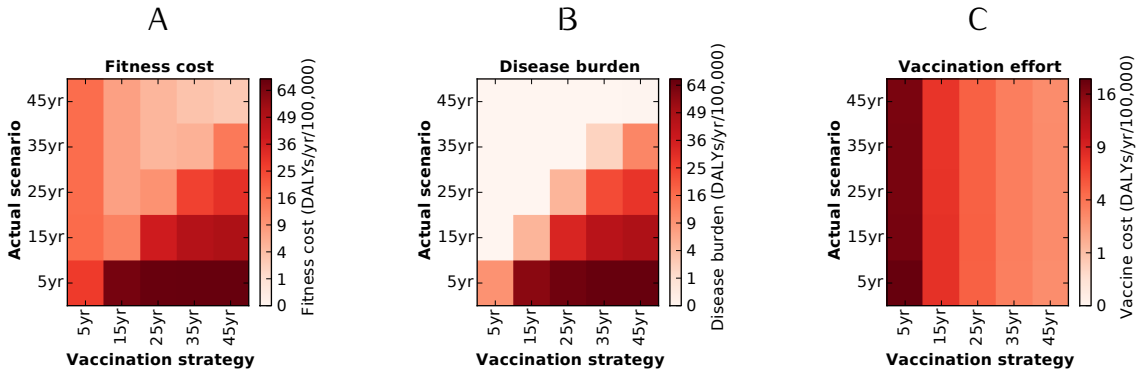


Figure S-6: Costs of mis-application of strategies for waning immunity. (A) The total fitness costs (combined vaccination cost and disease burden) of applying a boosting schedule evolved for each duration of immunity in each other scenario. The representative boosting schedules are constructed by taking the mean coverage at each age in the most fit strategy family (as shown in Fig. 2). These representative schedules (Strategy used) are applied to each scenario (Actual scenario) and the mean fitness cost among 120 replicant runs is plotted. Panels (B) and (C) plot the disease burden and cost of vaccination effort of these same strategies (so that panel (A) shows the sum of (B) and (C)).

References

[1] World Health Organization, Health statistics and information systems. (2014) Disability weights, discounting and age weighting of dalys.

[2] Crowcroft, N, Stein, C, Duclos, P, & Birmingham, M. (2003) *The Lancet Infectious Diseases* **3**, 413 – 418.

[3] Rohani, P, Zhong, X, & King, A. A. (2010) *Science* **330**, 982–985.

[4] Mossong, J. e. l, Hens, N, Jit, M, Beutels, P, Auranen, K, Mikolajczyk, R, Massari, M, Salmaso, S, Tomba, G. S, Wallinga, J, Heijne, J, Sadkowska-Todys, M, Rosinska, M, & Edmunds, W. J. (2008) *PLoS Medicine* **5**, 1.

[5] Eames, K, Tilston, N, Brooks-Pollock, E, & Edmunds, W. (2012) *PLoS Comput Biol* **8**, e1002425. doi:10.1371/journal.pcbi.1002425.

[6] Wearing, H. J & Rohani, P. (2009) *PLoS pathogens* **5**, e1000647.

[7] Pettrgrew, M & Resat, H. (2007) *J Chem Phys* **12**, 084101.