Supplemental Methods

Computer model. Most of our model parameters are taken from the murine system because more quantitative data is available for it than for the human system. The model simulates a mouse-sized B cell repertoire of 10^7 B cells [1, 2, 3]. B cell and antibody receptors, and antigens, are modeled as strings of symbols that can loosely be thought of as the amino acids of a binding site. The mapping between sequence and shape is not fully understood, so for the purposes of the model sequence and shape are assumed equivalent. The affinity between antibodies and antigens is proportional to the number of symbols that are complementary between their receptors. We define antigenic distance between antigens as the number of symbols in which their receptors differ. The length of the receptor (20 symbols) and the number of distinct symbols at each location (4) were chosen [4] to give the following properties: a potential repertoire of 10^{12} B cells, a 1 in 10^5 chance of a B cell responding to a particular antigen [5, 6, 7], and, with an expressed repertoire of 10^7 B cells [1, 2, 3], two antigens cease being cross-reactive when they have more than about 35% sequence difference [8, 9]. This latter property, combined with the 20-symbol receptor, gives eight degrees of cross-reactivity corresponding to antigenic distances 0 through 7. Due to the very large number of antibody molecules in a real immune response, each antibody in the simulation represents a number of real antibodies. T cell help is not included in the model and thus it is implicitly assumed that it is not a limiting factor in mounting a response. Cell-mediated responses are not included.

The affinity of a B cell or antibody for an antigen is 10^{r-d} where d is the mutation, or Hamming,[1] distance between the strings representing the B cell antibody receptor or the antibody and the antigen epitope, and r is the distance within which a B cell can be stimulated by an antigen. We set r=5, which means about 1 in 10^5 B cells can bind any particular antigen [4, 5, 6, 7]. During each 6 hour time period, each B cell gets a chance to bind with antigen; the antigen is selected at random from all antigens in the simulation.[2] The probability that a B cell will bind an antigen is

[1] The Hamming distance between two receptors is the number of places where the receptors differ. This can also be thought of as the mutation distance. For example, if the four equivalence classes of amino acids represented by the symbols A through D, and receptors are represented by strings of 20 symbols, the Hamming distance between receptors CC$\underline{B}$DD$\underline{B}$BCCCAB$\underline{D}$CCDADAD and CC$\underline{A}$DD$\underline{D}$BCCCAB$\underline{A}$CCDADAD is 3 because they differ in the three underlined locations.

[2] The selection method for B cell-antigen encounters is more complex when more than one specificity of antigen is

$0.1 \times a/\mathrm{av}(a)$, where a is the affinity of the B cell for the antigen and $\mathrm{av}(a)$ is the average affinity of the antibodies bound to the antigen. The probability is not allowed to exceed 1.0. If there are no antibodies bound to the antigen then a random B cell (but one that has some affinity for the antigen) is selected and its affinity is used as $\mathrm{av}(a)$. Comparing the affinity of the binding with the average affinity is an artificial way to model competition between B cells for antigen.

The antibody levels in Figure 3 of the main text were weighted to make them an effective measure of the ability of the antibody to eliminate antigen, and to make them comparable to the antibody levels measured by the hemagglutination inhibition assay. Each antibody in the highest three affinity classes (Hamming distances 0, 1 or 2 from the antigen) counted as 1 towards the antibody level; this was because such high affinity antibody would be expected to remain bound to an antigen for the duration of the test. Each antibody at Hamming distance 3 counted as 0.1 towards the weighted antibody level, Hamming distance 4 counted as 0.01, Hamming distance 5 counted as 0.001 and Hamming distance 6 counted as 0.0001. This 10-fold decrease in contribution to the weighted antibody level corresponds to the 10-fold decrease in affinity of an antibody for an antigen with each unit increase in the Hamming distance between them.

Because we are interested in the generation of B cell memory, we simulate events known to occur within germinal centers. Antigen stimulated B cells divide every 6 hours [10]. On cell division the B cell antibody receptor has a probability of 0.1 of somatically mutating its antibody receptor (i.e. mutation rate of 0.005 per symbol per generation). B cell differentiation is approximated by each daughter cell having a 0.05 probability of differentiating into a plasma cell and a 0.05 probability of differentiating into a memory cell per cell division. Memory B cells have a half-life longer than the duration of the simulation. When memory B cells bind antigen and divide, both daughter cells revert to short-lived B cells. Plasma cells have a 75% chance of having a half-life of 3 days, and a 25% chance of being long lived with an expected half-life of 200 days. (Plasma cell half lives of up to 100 days have been measured [11, 12]; however, we set the half-life to 200 days to keep titers closer to experimentally observed levels.) Plasma cells secrete 2 units of antibody every 6 hours. Free antibodies have a chance to bind antigens every 6 hours. For each unit of antibody, an antigen is selected at random. If the antigen is not bound to a B cell then the probability that the antibody

present in the simulation at the same time.

binds with the antigen is $a \times 0.01$, where a is the affinity of the antibody for the antigen (affinity calculated as described above). The probability is not allowed to exceed 1.0. Unbound antibodies have a half-life of 10 days. No more than 2 units of antibodies can to bind to any particular antigen. When an antigen has 2 units of antibodies bound it is removed from the system. Unbound antigen replicate with a 6hr expected time between divisions, this models an exponential growth in viral load. The expected lifetime for a naive unstimulated B cell is 4.5 days. New B cells, with randomly generated antibody receptors, are generated at a rate of 5.6×10^5 per day, on average, to maintain the B cell repertoire in steady state. Each B cell specificity is represented by one cell. In reality there may be more than one B cell and this is modeled by scaling the amount of antibody produced per B cell.

An example of the dynamics of the simulated immune response is shown in Figure 1.

The HI assay measures the ability of the antibodies raised to one antigen to bind to another antigen. According to our model, the antibodies that will bind are those in the intersection of the balls of stimulation of both antigens (Figure 2b of the main text). As the antigenic distance between two antigens increases, the volume of the intersection of their balls of stimulation decreases. The volume of the intersection in the model approximately halves for each unit increase in antigenic distance (the formula for the exact volume is derived in [4]). The HI assay is also exponential in nature, with concentrations being halved on each titration. In the model there are effectively no antibodies in the intersection (and thus no cross-reactivity between antigens) at antigenic distance 7. Similarly the HI assay is calibrated to indicate maximal cross-reactivity after approximately 7 or 8 halvings of antibody concentration. Thus, we equate a unit increase in antigenic distance between antigens in the model, to a halving of HI titer.

The immune response and clearance of two vaccinations and an epidemic challenge, that takes over 450 simulated days, takes approximately 2 minutes, running in Franz's Allegro Common Lisp (version 4.3), on a 400MHz Pentium II running Linux (Debian version 2.2). The simulator is written in Common Lisp, and the user interface is written in Tcl/TK and Blt. An implementation technique called *lazy evaluation* is used for efficient simulation the mouse-size B cell repertoire [13]. The model is available via *http://www.cs.unm.edu/dsmith/software/PNAS-model.html*.

Historical analysis. Antigenic distances between viral strains were derived from the hemagglutination inhibition (HI) assay. Each doubling of an HI titer was equated with an antigenic distance of 1. For example, one of the HI tables used to calculate the distances among strains involved in the 1974 (H3N2) outbreak at Christ's Hospital is show in Table 1 of these supplemental methods. The HI titer difference between the vaccine1 strain A/HongKong/8/68 in vaccine1 ferret antisera (1280) and the vaccine2 strain in vaccine1 ferret antisera (640) is one doubling, so we assign the antigenic distance of 1 between vaccine1 and vaccine2. We use the titer of vaccine2 antigen A/England/42/72 in vaccine1 sera (640), as opposed to vaccine1 antigen in vaccine2 sera (960), because in a vaccinated individual vaccine2 will see vaccine1 sera, and not vice-versa. The difference between the HI titer of vaccine2 antigen in vaccine2 sera (1280) and epidemic antigen in vaccine2 sera (40) is five doublings, so we say the antigenic distance between vaccine2 and the epidemic strain is 5. The HI titer difference between vaccine1 antigen in vaccine1 sera (1280) and epidemic antigen in vaccine1 sera (< 20) was taken as 7. The antigenic distances presented in Table 3 in the main text are the average antigenic distances from four historical HI tables made on different days during the relevant influenza seasons (data kindly provided by the Influenza Branch, CDC). For the 1974 Christ's Hospital outbreak, data from the four tables were averaged with the data in [14]. For the 1976 Christ's Hospital outbreak only one HI table was available that contained all the strains so it was used alone.

The average of antigenic distances derived from four HI tables are not necessarily integer distances, but the model only produces predictions for integer antigenic distances. When all antigenic distances among strains in an outbreak were half way between integer values, the average of vaccine efficacies in the two bounding groups were taken. For example, in the 1985 (B) outbreak, the vaccine2-epidemic strain, vaccine1-epidemic strain, and vaccine2-vaccine1 antigenic distances were 0.6, 0.6, and 0.0 respectively; the vaccine efficacies were calculated as the average of efficacies at antigenic distances 1, 1, and 0 respectively, and 0, 0, and 0 respectively. In situations such as the 1974 (H3N2) and 1987 (H3N2) outbreaks, all distances were scaled by the same amount to round them (as opposed to rounding each distance independently) to maintain the ratios among the distances. In 1974 (H3N2) real-valued distances were 3.2, 4.4, and 1.2, the integer antigenic distances used were 3, 4, and 1. In 1987 (H3N2) the real-valued distances were 3.3, 3.8, and 1.3; when scaled uniformly these give approximately distances 3.0, 3.5, 1.0, and so two sets of integer distances used (3, 4, 1, and 3, 3, 1) and the vaccine efficacies predicted from each were averaged.

Some of the attack rates at antigenic distances required to predict the outcomes of the Hoskins and Keitel outbreaks were not part of the results reported in the main text and are reported here. The timing and dose of vaccinations and epidemic challenge, and sizes of groups where that same as in the main text, the only differences were the antigenic distances among the vaccine and epidemic strains. The attack rates computed from our model for first-time vaccinees when the vaccine2-epidemic strain antigenic distances were 0, 1, and 3 were 0.00, 0.02, and 0.83, respectively. The attacks rate for repeat vaccinees were (the antigenic distances are given in the order of vaccine2-epidemic strain, vaccine1-epidemic strain, and vaccine1-vaccine2, and are followed by the attack rate) 0, 0, 0, 0.00; 1, 1, 0, 0.29; 3, 3, 0, 0.90; 3, 3, 1, 0.79; and 3, 4, 1, 0.95.

Two H3N2 vaccine strains were given concomitantly during the Hoskins study in the two years prior to the 1976 (H3N2) outbreak, and two H1N1 strains were given during the Keitel study prior to the 1986 (H1N1) outbreak. In these cases two vaccines were given concomitantly in the model. The antigenic distance between concomitant strains a and b was calculated as the average of the antigenic distance between strain a serum and strain b virus, and strain b serum and strain a virus. The timing and dose of vaccinations and epidemic challenge, and sizes of groups where that same as in the main text, the only differences were the number of strains in each vaccine and the antigenic distances among the vaccine and epidemic strains. The attack rates computed from our model for first-time vaccinees in 1976 was 0.33 and in 1986 was 0.00. The attack rates computed from our model for repeat vaccinees in 1976 was 0.43 and in 1986 was 0.00.

The A/HongKong/X31/68 (H3N2) and A/HongKong/1/68X (H3N2) vaccine strains used prior to the 1972 outbreak in the Hoskins study are assumed to be antigenically equivalent to the A/HongKong/8/68 reference strain. We make this assumption because HI tables containing A/HongKong/1/68X were not available. If there are antigenic differences between the vaccine strains this would reduce negative interference and likely increase the efficacy in repeat vaccinees. The assumed antigenic distances give an overestimate, compared to observation, of the vaccine efficacy in repeat vaccinees relative to first-time vaccinees for the 1972 outbreak; thus, if the assumption is not valid, and there are antigenic differences, this would likely increase the overall correlation.

The vaccine strains given before the Keitel study started were not reported; however, it is likely that the repeat vaccinees received A/Brazil/11/78 (H1N1) and B/Singapore/222/79 as these were the recommended strains for the 4 years prior to the Keitel study. Thus, we used these as vaccine1 strains when calculating antigenic distances for the 1983 H1N1 and B outbreaks.

Vaccine efficacy. Vaccine efficacies were calculated as ((attack rate in no-vaccine group) - (attack rate in vaccine group)) / (attack rate in no-vaccine group). Vaccine efficacies for the Keitel study were taken from attack rates determined by infection (as opposed to virus isolation). The Christ's Hospital study did not have a randomized no-vaccine group, and the usual caveats apply regarding non-randomized trials. The no-vaccine attack rate used for the 1974 vaccine efficacy calculations were extracted from Figure 2 in [15].

References

[1] Köhler, G. (1976) *Eur. J. Immunol.* **6**, 340–347.

[2] Klinman, N. R, Press, J. L, Sigal, N. H, & Gerhart, P. J. (1976) in *The Generation of Antibody Diversity*, ed. Cunningham, A. J. (Academic Press, New York), pp. 127–150.

[3] Klinman, N. R, Sigal, N. H, Metcalf, E. S, Pierce, S. K, & Gerhart, P. J. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **41**, 165–173.

[4] Smith, D. J, Forrest, S, Hightower, R. R, & Perelson, A. S. (1997) *J. Theor. Biol.* **189**, 141–150.

[5] Edelman, G. M. (1974) in *Cellular Selection and Regulation in the Immune System*, ed. Edelman, G. M. (Raven Press, New York), pp. 1–38.

[6] Nossal, C. J. V & Ada, G. L. (1971) *Antigens, Lymphoid Cells and The Immune Response*. (Academic Press, New York).

[7] Jerne, N. K. (1974) in *Cellular Selection and Regulation in the Immune System*, ed. Edelman, G. M. (Raven Press, New York), pp. 39–48.

[8] Champion, A. B, Soderberg, K. L, Wilson, A. C, & Ambler, R. P. (1975) *J. Mol. Evol.* **5**, 291–305.

[9] East, I. J, Todd, P. E, & Leach, S. J. (1980) *Mol. Immunol.* **17**, 1539–1544.

[10] Zhang, J, MacLennan, I. C. M, Liu, Y. J, & Lane, P. J. L. (1988) *Immunol. Letters* **18**, 297–300.

[11] Manz, R. A, Thiel, A, & Radbruch, A. (1997) *Nature* **388**, 133–134.

[12] Slifka, M. K, Antia, R, Whitmore, J. K, & Ahmed, R. (1998) *Immunity* **8**, 363–372.

[13] Smith, D. J, Forrest, S, Ackley, D. H, & Perelson, A. S. (1998) *Bull. Math. Biol.* **60**, 647–658.

[14] Both, G. W, Sleigh, M. J, Cox, N. J, & Kendal, A. P. (1983) *J. Virol.* **48**, 52–60.

[15] Hoskins, T. W, Davis, J. R, Smith, A. J, Miller, C. L, & Allchin, A. (1979) *Lancet* **i**, 33–35.

Figure Captions

Figure 1. An enlarged region of Figure 3(a) of the main text; showing finer details of the dynamics of the simulated immune response.

Antigen	Ferret antisera		
	A/HK/68	A/Eng/72	A/PC/73
A/HongKong/8/68 (vaccine1)	1280	960	80
A/England/42/72 (vaccine2)	640	1280	640
A/PortChalmers/1/73 (epidemic)	<20	40	320

Table 1: Hemagglutination inhibition titers between influenza strains similar to the vaccine and epidemic strains of the influenza H3N2 outbreak at Christ's Hospital in 1974. Titers were extracted from a larger table in [14].

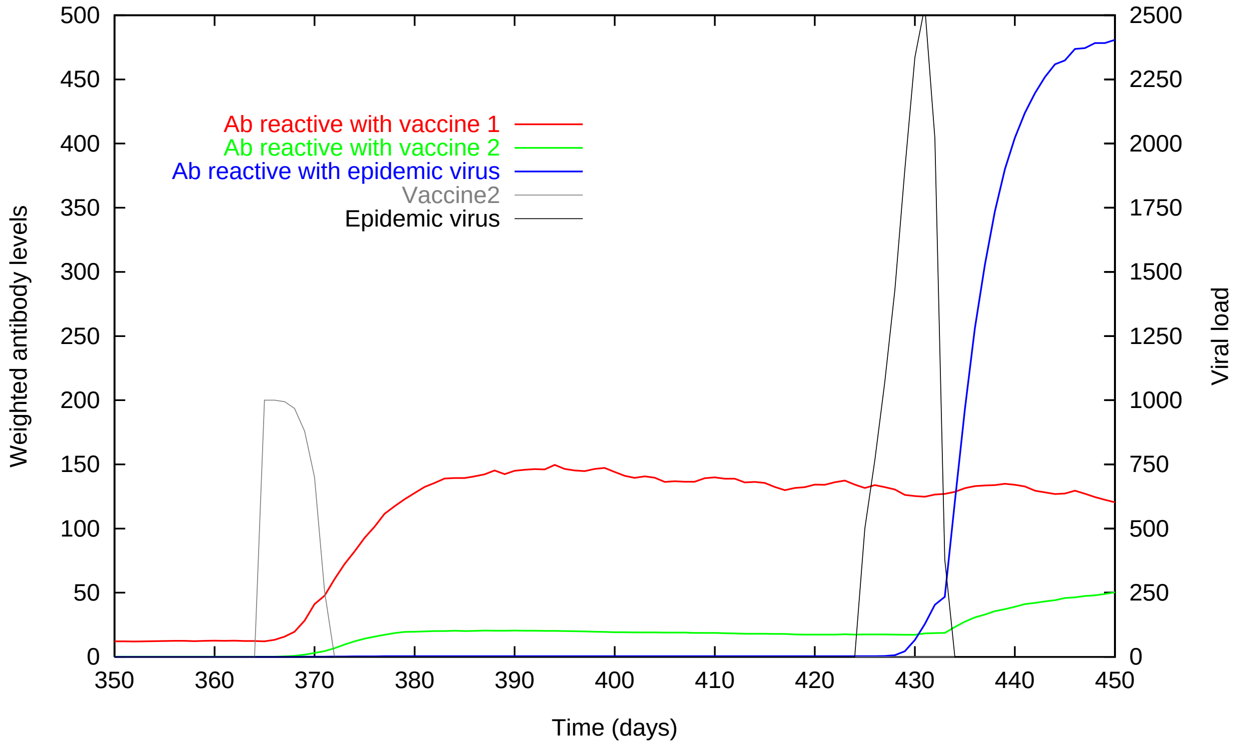
Ab reactive with vaccine 1
Ab reactive with vaccine 2
Ab reactive with epidemic virus
Vaccine2
Epidemic virus
Weighted antibody levels
Viral load
Time (days)
0
50
100
150
200
250
300
350
400
450
500
0
250
500
750
1000
1250
1500
1750
2000
2250
2500
350
360
370
380
390
400
410
420
430
440
450

Figure 1: