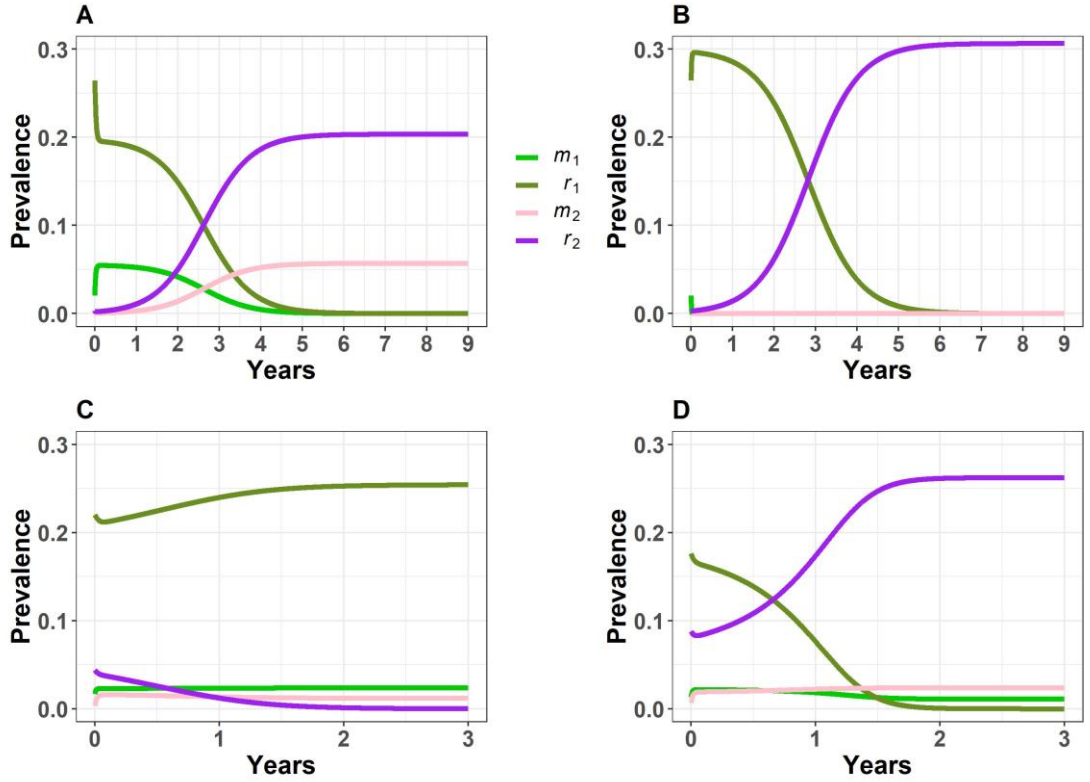
Appendix, with additional figures, to:

Understanding MRSA clonal competition within a UK hospital; the possible importance of density dependence

Anneke S. de Vos, Sake J. de Vlas, Jodi A. Lindsay, Mirjam E.E. Kretzschmar, Gwenan M. Knight

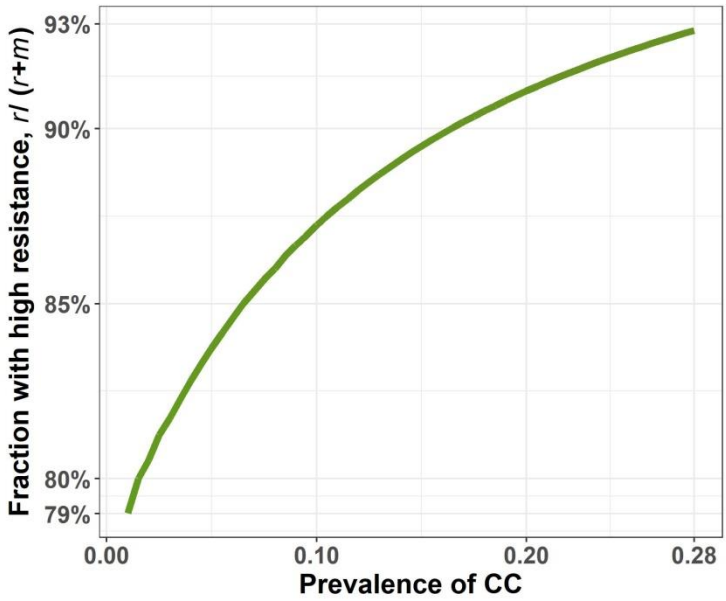
Examination of two competing CCs



Appendix Figure 1. Example dynamics for two competing CCs. Prevalence over time of the basic MRSA resistant m and higher resistant *r*-strains of two CCs, as dependent on starting densities. For both CC, parameters are as at baseline for CC22 (see Table 1) except $b_2 = b_1 * 1.01$, and inflow $i_1 = i_2 = 0$ (for panels A and B), and in **A** $g = 0$ and in **B** $l = 0$. In both these panels CC_2 starts with 1/100th the starting values of CC_1 (which would be in equilibrium if g were 1 and l were 0.03) but ends up as the dominant clone. In panels **C** and **D**, inflow $i_{m1} = i_{m2} = 0.0015$. In **C**, CC_1 starts at 5/6th of its panel A value, with CC_2 at 1/6th, while in **D** CC_1 starts at 4/6th of its panel A value, with CC_2 at 2/6th.

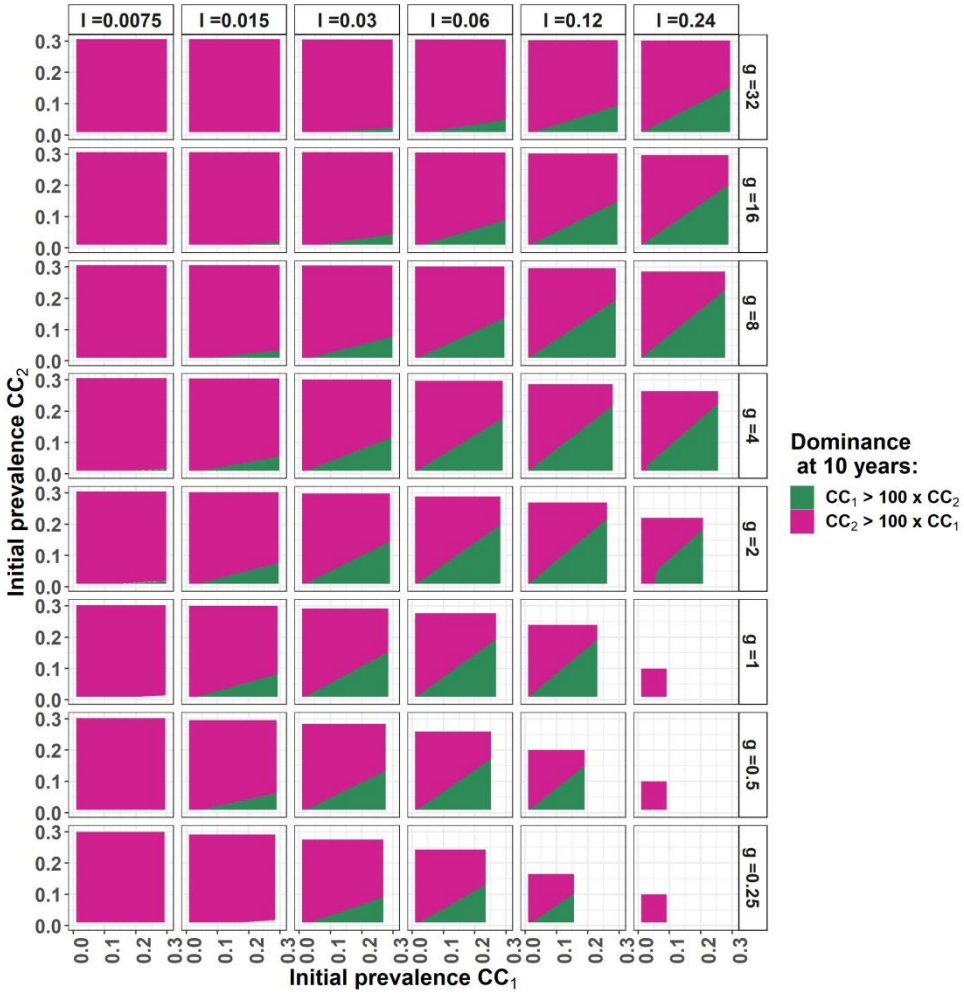
Without gain and loss of resistance, the faster growing CC_2 (purple in Appendix Figure 1) always out-competes CC_1 in the end. Even with external inflow, competition can be density dependent; neither complex will be absent completely in the end, but which one dominates may depend on initial densities.

Density dependence characteristics



Appendix Figure 2. Fraction of the population of bacteria of one CC that carry an element conferring resistance, as dependent on the density of this CC. As stated in the Methods, we solve for (initial) resistance levels using $(r_j + dr_j/dt)/(r_j + dr_j/dt + m_j + dm_j/dt) = r_j/(m_j + r_j)$ with $m_j + r_j = x$, where x is the prevalence of the CC. Parameters are as at baseline for CC22 (see Table 1) except inflow $i = 0$.

The value of $r/(m + r)$ as calculated in this manner was used to set the starting conditions for main text Figure 3 and appendix Figure 3. If the resistant fraction is not initially assumed in equilibrium with prevalence, but for example at a certain equal percentage for both CC, these Figures would be slightly different, showing somewhat less density dependence in CC dominance. However, since gain and loss are relatively fast processes, the semi-steady state like assumption for resistance level is reasonable.



Appendix Figure 3. Eventual outcome of competition between two MRSA CCs as determined by initial densities of the two CCs, and on loss-rate *l* and gain-rate *g*. Here CC_2 (dominance in pink) has a slight (1%) growth advantage over CC_1 (dominance in green). For each CC, we consider only starting densities below or at the equilibrium density of this CC (as achieved without other CCs present). For example, at high loss- and low gain-rate (lower three extreme right panels), both CC populations lose all antibiotic resistance, which significantly lowers the maximum densities they can achieve. Initial resistance level within each CC is assumed at equilibrium with CC density (see Methods and Appendix Figure 1). For both CC, parameters are as at baseline for CC22 (see Table 1) except inflow $i_1 = i_2 = 0$ and $b_2 = b_1 * 1.01$ (i.e., as in main text Figure 3B).

Density dependence in competition outcome occurs when greater density allows for a greater fraction of bacteria to carry the resistant element, giving an advantage to the complex with greatest initial density. Yet with a very low loss rate (left panels), almost all bacteria carry the element even at low density of a CC. At a very high loss rate on the other hand, the element will be lost from the population, unless the gain rate is also very high. Without the resistant element, the MRSA are rendered less fit, so that the maximum density that can be reached by the CCs is then much lowered (bottom right panels).

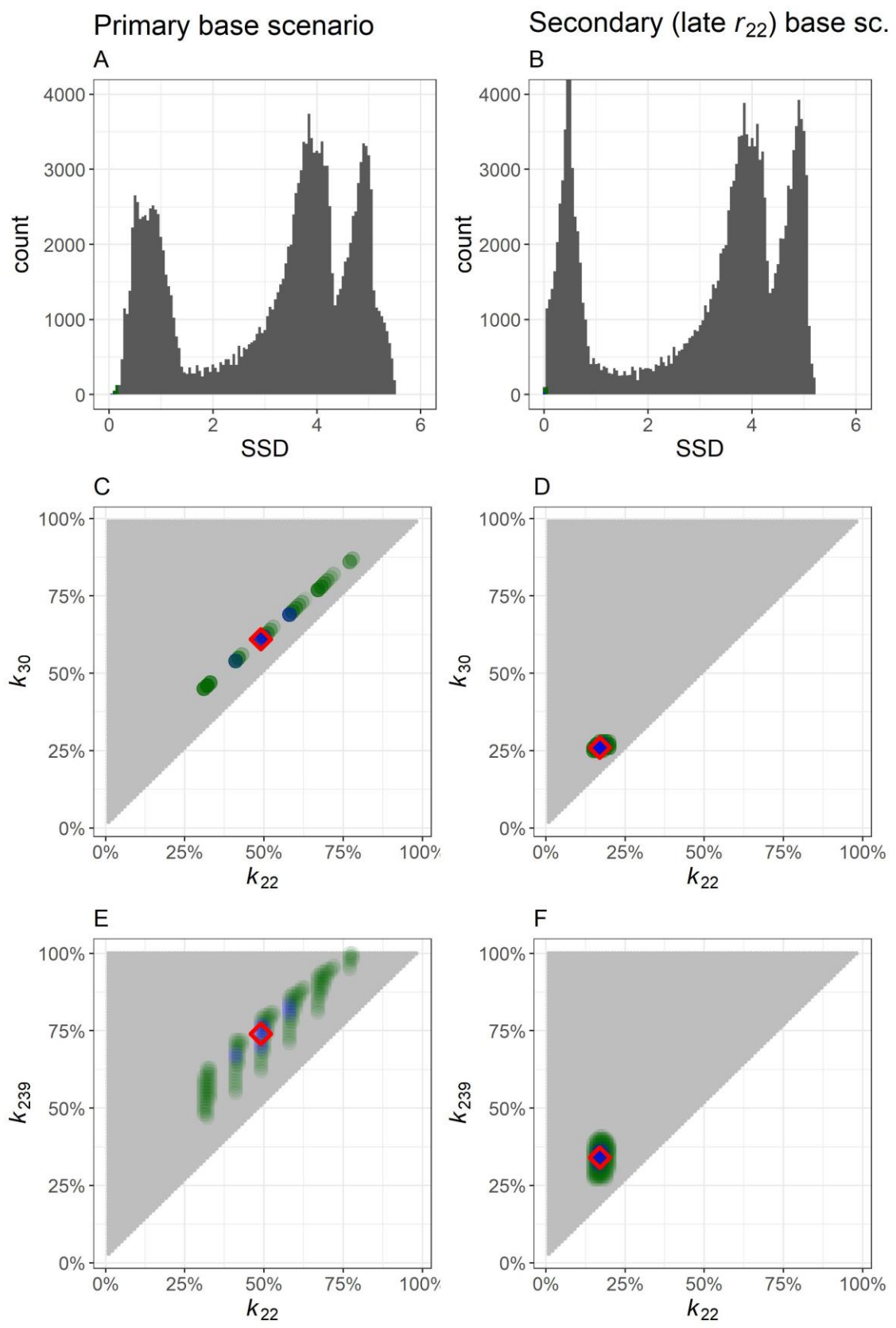
Only at medium loss rate, or if a high loss rate is combined with a high gain rate (upper right panels), will the population resistance level depend on the abundance of the CC; at greater

abundance, bacteria with the element are more likely to come into contact with bacteria without the element, giving the opportunity for horizontal gene transfer.

Interpreting the gain rate is not straight-forward, in that it depends on both the movement of bacteria (how many potential meetings take place) as well as the per encounter chance of actual horizontal transfer. Bacteria that carry the element will transfer this element on average to $g * s$ other bacteria per day, so in case of $g = 1$, the baseline value in our study, this means that less than 1 other bacterium is transferred to per day. (If transfer takes much time, then the gain rate might actually saturate with density of the m bacteria, yet it is not likely that this process takes long enough to impact dynamics.)

Loss of the element might occur in several ways, one of which is at cell-division, when all copies might end up in one of the two daughter cells. Assuming only this manner, the loss rate can be roughly re-quantified to the fraction of cell-divisions at which only one daughter cell retains copies, as $l/(b * z_t)$. For example, for CC_1, at equilibrium density, $0.66 * 0.72 = 48\%$ of cells divide each day, then with $l = 0.03$ per day (the baseline value) the element is lost in about 3% of daughter cells.

Reproducing observed dynamics



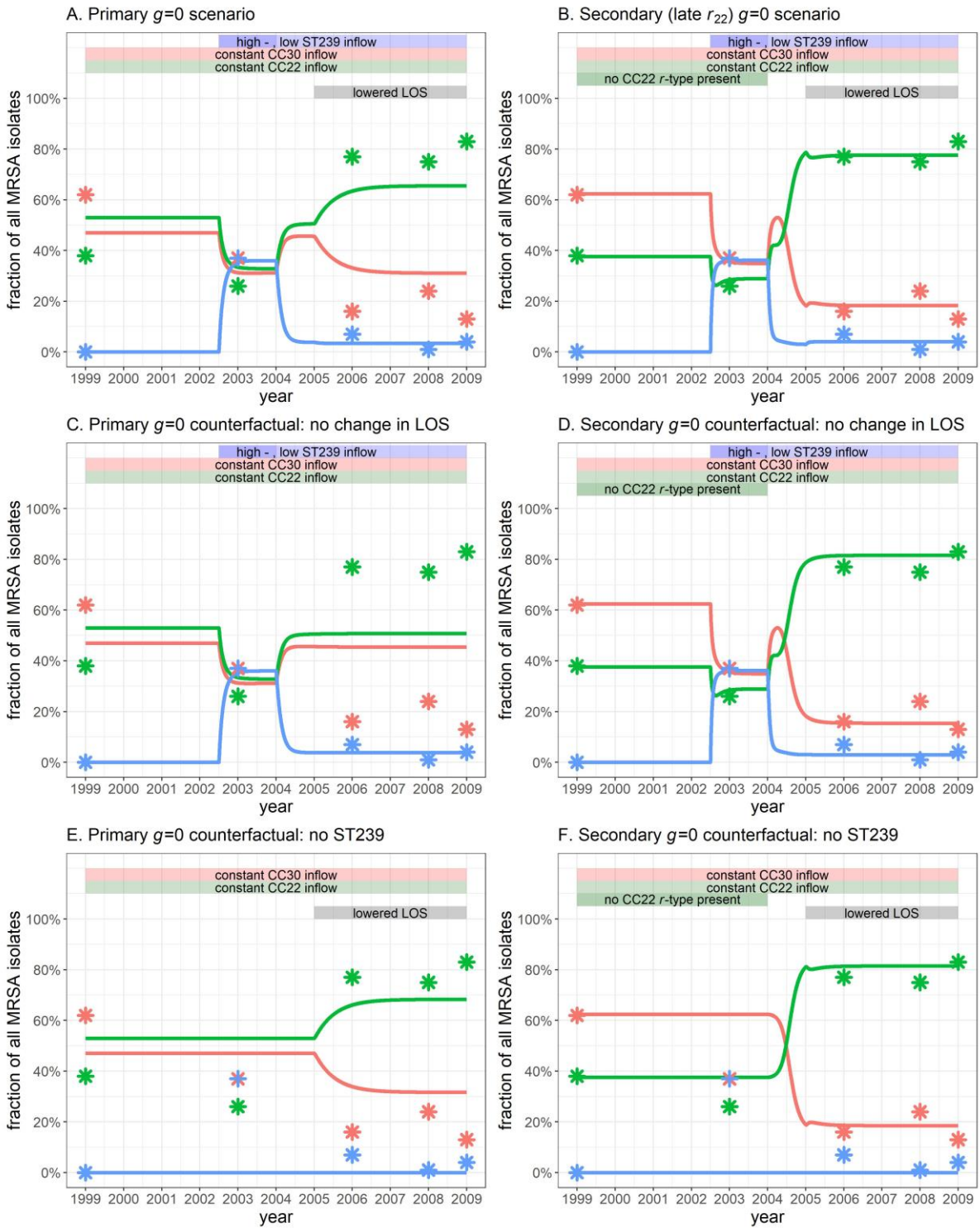
Appendix Figure 4. Exploration of the fitting space. For each scenario, restricting $k_{239} > k_{30} > k_{22}$, and with step sizes of 1% for each of the k_j, all combinations of these three parameters were

explored. The left column shows the fitting space for our primary model version, in the right column for the secondary model version (assuming late CC22 *r*-strain inception), with baseline parameter settings for both. In panels **A** & **B**: histograms of the sums of squared differences (SSD) between model and relative prevalence data for these scenarios. Lower SSD indicates a better model fit. In panels **C & E** for the primary and in **D** & **F** for the secondary scenario, the considered parameter combinations (grey dots). Per scenario, the best fitting 200 combinations (in green), the best fitting 50 combinations (in blue) and the very best fitting combination (red diamond) are indicated in each panel.

Note that for computational reasons, the minimum and maximum of the fraction i_{rj} of i_j were set to 1% and 99% respectively, which at most parameter settings, certainly for the well-fitting parameter settings shown in our other Figures, has no impact, but which does lead to some very small difference in the fit for some of the parameter combinations shown here, compared to results that would be obtained when not setting any limits.

For the secondary model version, a clear optimum in parameter space is seen. At this setting, CC22 can take over from CC30 as soon as it gains the higher resistance conferring genetic element in 2005. For the primary model version, instead, a band of reasonably fitting parameter combinations is seen. Along this band, where CC22 and CC30 high resistance levels are about 10% apart, density dependence in competition occurs; higher prevalence allows for more transmission of a resistant element, increasing competitiveness of the CC-population as a whole. The CC22 take-over from CC30 can then be explained as a switch between two possible stable states of the system. Outside of this band in parameter space, other fitness factors dominate competition between CC30 and CC22, so that density dependence is lost. Above the band in panel C, CC22 is too susceptible to removal by antibiotics compared to CC30, even when it carries high resistance, so that CC22 never takes over. Conversely, below the band in panel C, CC22 has a lesser disadvantage in lower achievable resistance, and in this case CC22's growth-rate advantage causes it to take over from CC30 from the start of model simulation.

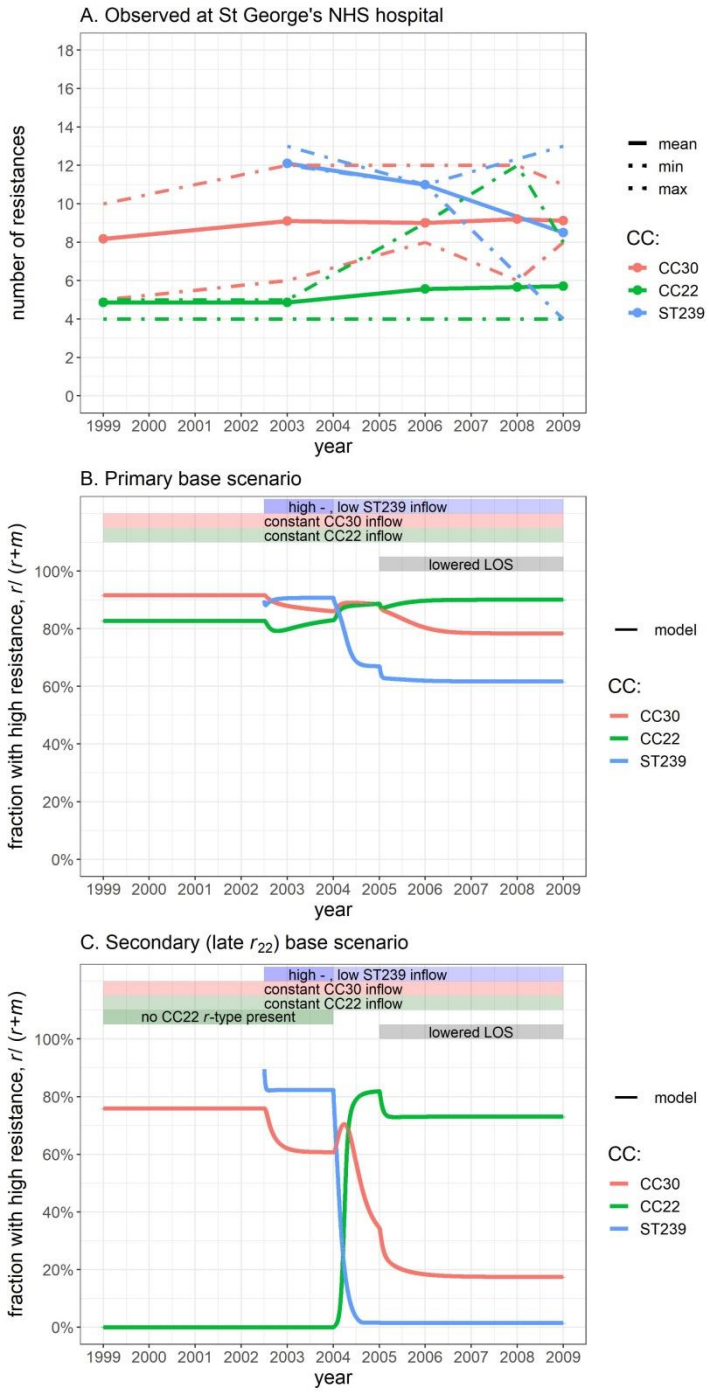
In view of this lack of a single optimum, and also of the high uncertainty in other parameters used in the scenario fits, and the simplicity of only two resistance levels per CC in the model, these fits are not meant to be informative about the actual death-rates averted by antibiotic resistance for these MRSA.



Appendix Figure 5. Model fits to observed clonal dynamics, and counterfactual scenarios, all without density dependence in competition ($g = 0$). Left column, primary model version, right column, secondary model version, with the additional assumption of late r_{22}introuduction in 2004. Panels **A** & **B**) Model fits. **C** & **D**) The same fits as shown in A & B respectively, but re-running the model keeping the mean length of stay in hospital constant ($d = 0.28$ also after 2005). **E** & **F**) The same fits as shown in A & B respectively, but re-running the model without ST239 presence ($i_{239} = 0$ throughout).

Note that without the system bi-stability the disruption by ST239 could have only had a temporary effect, since, with just a single stable equilibrium, the system must return to it when ST239 inflow stops. This can best be seen in panel **B**, where CC30 and CC22 prevalence (almost) returned to their earlier values when ST239 inflow (almost) dried up.

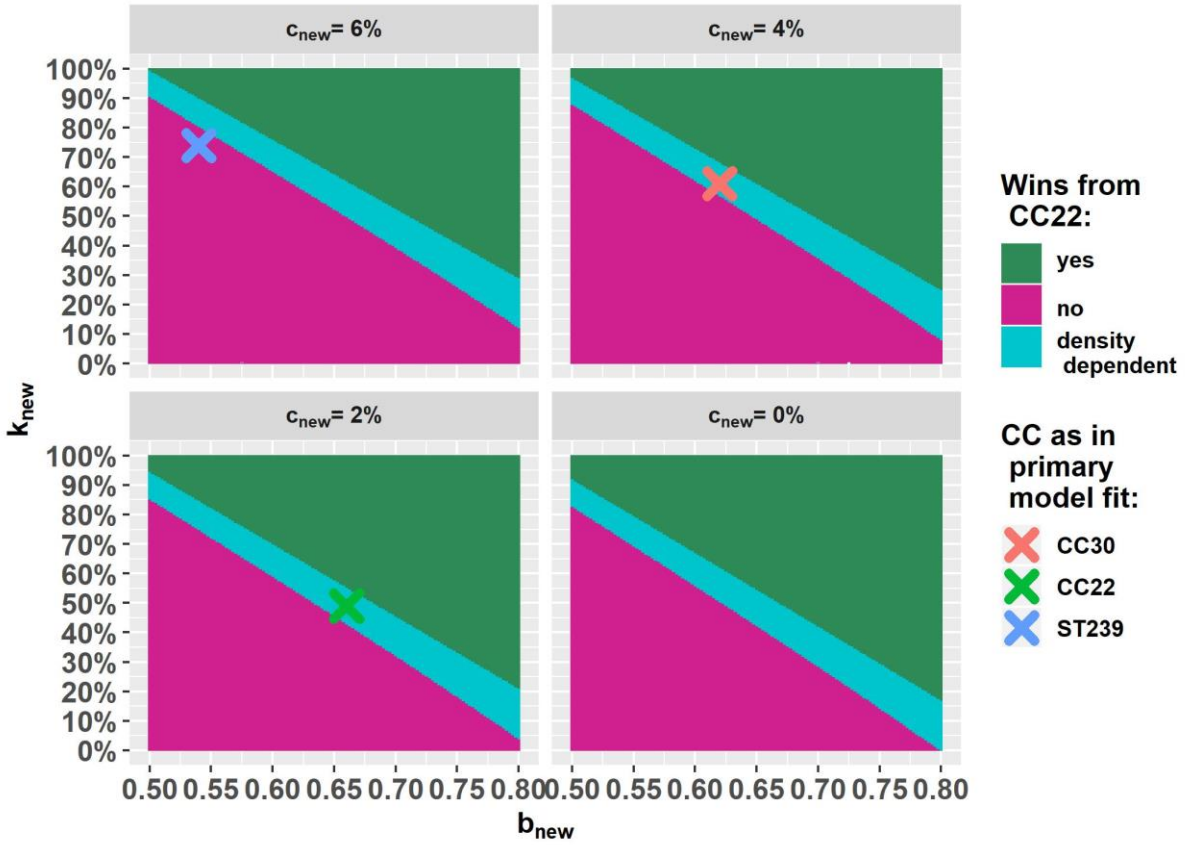
Panel **C** shows that the change in mean length of hospital stay is not sufficient by itself to cause as large an increase in relative CC22 prevalence as was observed in the hospital. A shorter mean length of hospital stay is most disadvantageous for CC30, since it has a lower growth (or infection) rate than CC22; with the more limited time in which individuals can spread MRSA once infected, it is rendered more important to infect them quickly. In panel **F**, in the secondary model version, we find the relative CC22 prevalence level to be slightly negatively impacted by the change in length of stay, despite a shorter mean length of stay being most advantageous for CC22. However, in this alternative model version, since the fit antibiotic resistance levels k_{30} and k_{22} are both low (see Table 2), absolute levels of CC22 and CC30 are low in this scenario, rendered even lower by the decreased length of stay. This renders the relative impact of inflow of CC22 and CC30 high, and since inflow is equal for both, this moves them slightly towards 50% each.



Appendix Figure 6. Resistance data and modelled resistance levels. Panel **A**: Mean, maximum and minimum number of phenotypic resistances to antibiotics that were observed to be carried by the different strains of the CCs at St George's NHS hospital (for details see Knight *et al.* (15)). A minimum of four antibiotic resistances defines MRSA. In panel **B** for our primary model version base scenario fit and in panel **C** for the secondary model version base scenario fit (with assumed CC22 *r*-type introduction in 2004), the modelled fraction of bacteria carrying higher resistance.

Note that the total in modelled CC population resistance is also determined by the (presumed unchanging) k parameters that set how much antibiotic induced death is avoided at high resistance level of a CC (see Table 2); in the primary base scenario the total antibiotic induced bacterial death is always higher for CC30 compared to for CC22.

Hypothetical challenger to CC22



Appendix Figure 7. Outcome of competition between CC22 and a hypothetical challenger. The impact of the growth-rate (b), resistance (k), and resistance cost (c) parameters on competitiveness of a hypothetical new CC is shown. Parameters for CC22 are as in our primary model fit, i.e. $k_{22} = 49\%$, but here we assume inflow of neither CC22 nor of its hypothetical challenger. We do not include presence of CC30 or any other CCs, and set $d = 0.32$ (as at the end in our basic scenario), and we solve for equilibria as explained in our Methods section. Where we find one equilibrium we check which of the CCs has prevalence >0, whereas existence of multiple equilibria indicates density dependence in competition.

With either a significantly higher growth-rate or with much greater resistance to antibiotics (not offset by too high a cost in carrying this resistance), a hypothetical new CC would take over the dominant prevalence position from CC22. However, where fitness is comparable between CCs, we see density dependence in outcome of competition. CC22 would then likely keep its dominance as it starts at high prevalence. Occurrence of this density dependent competition band in parameter space is a robust outcome of our model, as long as the loss rate l of resistance is not too small (see Appendix Figure 3). However, model parameter uncertainty does render the width and breadth of this density dependence band uncertain.

Tables and Figures.

Table 1. Parameter values.

Description	Parameter	Time dependent	Base value	Alternative values
Removal rate of bacteria	d	before 2005	0.28 per day	
		from 2005	0.32 per day	
Additional removal rate due to antibiotics	a		0.3 per day	
Growth-rate	b_{30}		0.62 per day	
	b_{22}		0.66 per day	
	b_{239}		0.54 per day	
Cost to resistance in percentile decrease in growth-rate for resistant bacteria	c_{30}		4 %	
	c_{22}		2 %	
	c_{239}		6 %	
Rate of resistance transfer	g		1 per day	0 per day
Rate of resistance loss	l		0.03 per day	0 per day
Rate of resistance gain by mutation	s		0 per day	0.003, 0.01 per day
Rate of bacterial inflow	$i_{30} = i_{22}$*		0.0015 per day	
	i_{239}*	before 2003	0 per day	
		2003 - 2004	0.004 per day	
		from 2004	0.0008 per day	
Scalar of the dependence of the resistant fraction in inflow of a CC on current hospital resistance level	h		0.5	0, 0.25, 0.75, 0.8, 1
Percentile decrease in the antibiotics induced removal-rate for resistant bacteria	k_{30} & k_{22} & k_{239}		Fit to prevalence data	

Parameter values as used to reproduce the MRSA clonal complex (CC) dynamics as observed at St George's Healthcare NHS Trust hospital between 1999 and 2009. For the more theoretical main Figures 1-3, plus Appendix Figures 1-3, we use parameters as at baseline for CC22, while d is set at

0.3 (the mean over the two time-periods), and $k = 49\%$, as in the primary model version basic scenario fit (see Table 2). For each scenario, i.e. primary or secondary model version, baseline or alternative parameter value, k_{30} & k_{22} & k_{239} are fit to best reproduce the prevalence data. For details, see Methods. See Table 2 for the fit values. * We assume a constant number of patients within the hospital, so that the total patient inflow rate equals the outflow rate of ~0.2 per day (as the mean length of stay is about 5 days). Then infected inflow *i* equals this total inflow rate of 0.2 multiplied by the proportion of individuals infected at hospital entrance, as stated per CC in the Methods section.

Table 2. Fitted model scenarios.

Model version	Change from base values	Presence of multiple stable states at *t*=0 (1999) for best fit ks (and in what fraction of all k_{30} and k_{22} combinations)	Best fit:				
			k_{30}	k_{22}	k_{239}	SSD	Figure
Primary	$h = 0$	Yes (16.7%)	32 %	15 %	39 %	0.022	-
	$h = 0.25$	Yes (6.3%)	40 %	25 %	53 %	0.086	-
	All baseline	**Yes (2.6%)**	**61 %**	**49 %**	**74 %**	**0.043**	**4A**
	$h = 0.75$	Yes (0.9%)	81 %	71 %	94 %	0.026	-
	$h = 0.8$	Yes (0.7%)	82 %	72 %	95 %	0.039	-
	$h = 1$	Yes (0.2%)	84 %	74 %	96 %	0.077	-
	$g = 0$	No (0.0%)	85 %	75 %	95 %	0.181	App.A4A
	$l = 0$	No (0.0%)	85 %	76 %	96 %	0.181	-
	$s = 0.003$	Yes (2.4%)	62 %	50 %	76 %	0.059	-
	$s = 0.01$	Yes (1.9%)	71 %	60 %	84 %	0.062	-
Secondary (with evolutionary event, i.e. CC22 *r*-type introduced in 2004*)	$h = 0$	No (0.0%)	28 %	16 %	33 %	0.020	-
	$h = 0.25$	No (0.0%)	27 %	15 %	34 %	0.017	-
	All baseline	**No (0.0%)**	**26 %**	**17 %**	**34 %**	**0.015**	**4B**
	$h = 0.75$	No (0.0%)	26 %	20 %	36 %	0.013	-
	$h = 0.8$	No (0.0%)	26 %	21 %	36 %	0.014	-
	$h = 1$	No (0.0%)	25 %	22 %	36 %	0.013	-
	$g = 0$	No (0.0%)	34 %	27 %	37 %	0.013	App.A4B
	$l = 0$	No (0.0%)	22 %	17 %	32 %	0.012	-

Each scenario was fit to the data by finding the k parameters, denoting the proportional decrease in the antibiotics induced removal-rate for resistant bacteria, for which the sum of squared differences (SSD) between model outcome and data was minimal. For examples of the fitting space, see Appendix Figure 4. The impact of parameters on model fit was explored by changing one at a time; all other parameters were set at their base values (see Table 1; at baseline $h = 0.5,\ g = 1, l = 0.03$ and $s = 0$). *In the secondary model version with CC22 *r*-type introduced only from 2004, we did not run the model with alternative values for the mutation rate *s*, since if $s \neq 0$, a higher resistance carrying element is obtained by CC22 from 1999, and we in effect regain our primary model setting.

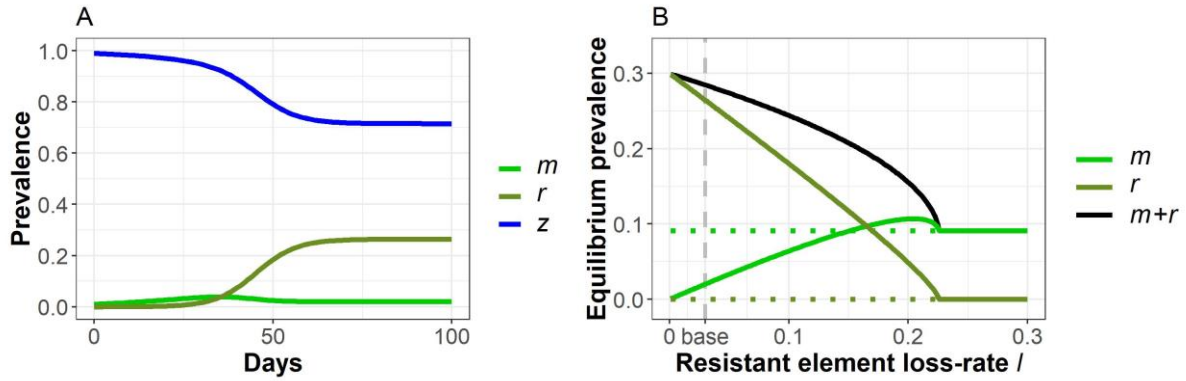


Figure 1. Example dynamics for a single CC. A: Prevalence over time of the basic MRSA resistant m (light-green) and higher resistant *r*-strain (dark-green) of a single CC, which has entered the hospital with low initial prevalence ($m(t=0)=0.01$ and $r(t=0)=0.0001$.) *z* is the density of resource available. Parameters are as at baseline for CC22 (Table 1) except inflow $i=0$. **B:** Equilibrium prevalence of this CC as dependent on the loss-rate l of the resistant element (solid lines). The baseline loss-rate *l* = 0.03 (used for panel A) is indicated here with a vertical dashed line. Total CC prevalence declines with increasing loss-rate, since in this setting (with high antibiotic induced death-rate, a) resistance is fitness enhancing (i.e. outweighs the cost to resistance in diminished growth, *c*). The equilibrium prevalence without the *r*-strain present is also shown (dotted lines). Note that this unstable equilibrium is lost when the mutation rate $s > 0$.

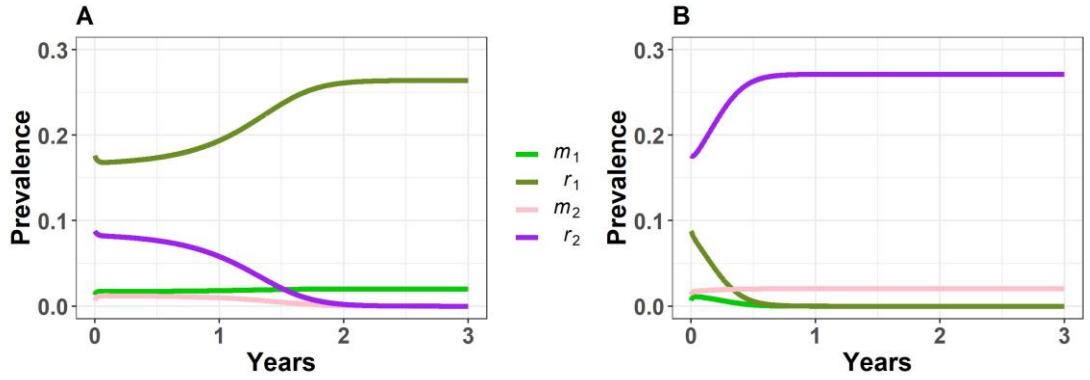


Figure 2. Dynamics for two competing CCs, exemplifying density dependence. Prevalence over time of the basic MRSA resistant m and higher resistant *r*-strains of two CCs, as dependent on starting densities. CC_2 has a 1% higher growth-rate than CC_1, all other parameters are equal (and at baseline for CC22, except inflow $i=0$ (see Table 1)). In **A** both m_1 and r_1 (so CC_1 as whole) (light- and dark-green respectively) start at $2/3^d$ of their equilibrium density, and m_2 and r_2 (light- and dark-purple) start at half these densities. In **B** the starting densities for CC_1 and CC_2 are reversed.

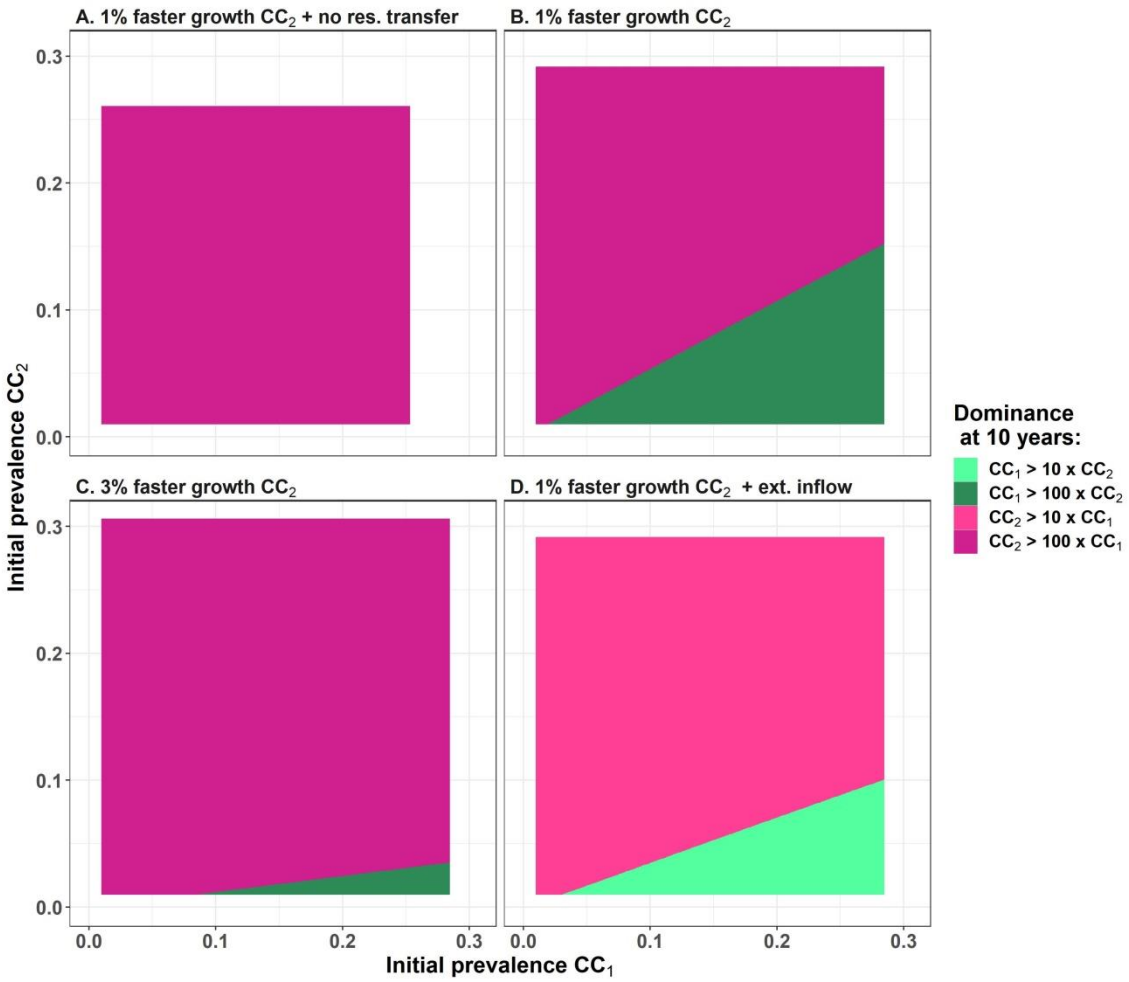


Figure 3. Eventual outcome of competition between two MRSA CCs as determined by initial densities of both CCs. Here CC_2 (dominance in pink) has the growth advantage over CC_1 (dominance in green). For each CC, we consider only starting densities below or at the equilibrium density of this CC (as achieved without other CCs present) (hence the unequal panel sizes). Initial resistance level within each CC is assumed at equilibrium with CC density (see Methods and Appendix Figure 2). For both CCs, parameters are as at baseline for CC22 (see Table 1) except inflow $i_1 = i_2 = 0$ (for panels A, B and C) and $b_2 = b_1 * 1.01$ (for panels A, B and D). For panel **A**, resistance transfer $g = 0$ instead of baseline $g = 1$. For panel **C**, $b_2 = b_1 * 1.03$. For panel **D**, $i_{m1} = i_{m2} = 0.0015$.

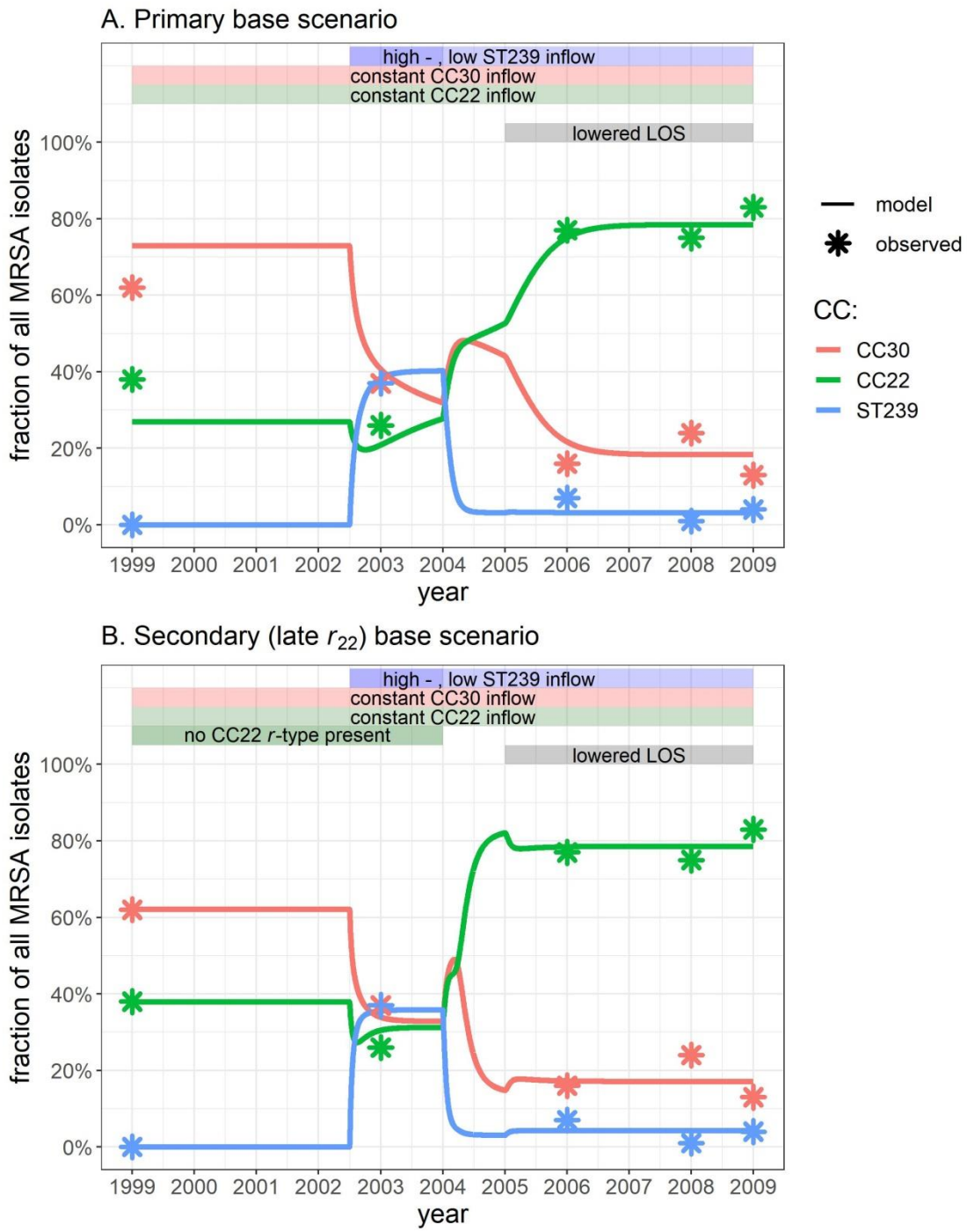


Figure 4. Model fits to observed clonal dynamics. Model output (coloured lines, for CC30 (red), CC22 (green) and ST239 (blue)) compared to the relative CC-prevalences observed at St George's Healthcare NHS Trust (star points). As explained in the Methods section, the hospital system is assumed to be at steady state in 1999, meaning modelled CC levels would not change until something happens. In both scenarios we include two known events: an ST239 outbreak in a nearby hospital around 2004, causing a short-term high inflow of this CC, and a drop in length of hospital stay from ~6 to ~5 days in 2005. The timings of these events are indicated in the top text-bars. Panel **A**: Primary model version fit. Panel **B**: Secondary model version fit, with an additional evolutionary event assumed, causing the CC22 *r*-type to be introduced only in 2004, i.e. no CC22 *r*-type present before. See Table 2 for values of the fit parameters.

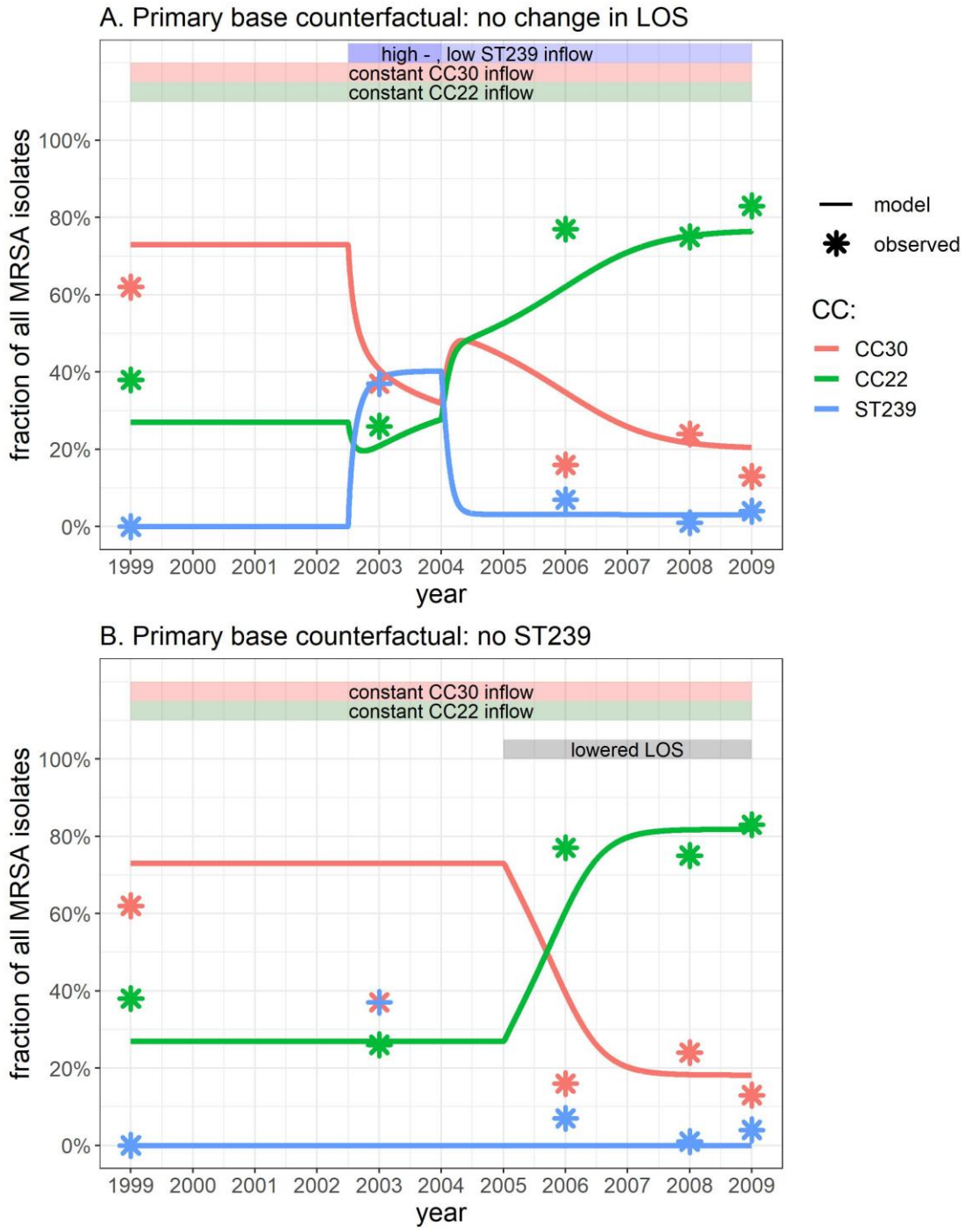


Figure 5. Counterfactual scenarios. Model output is shown for the same primary scenario fit as shown in Figure 4A, but the model is re-run not including either one of the two known disturbance events; In panel **A**, the mean length of stay in hospital is kept constant ($d = 0.28$ also after 2005). In **B**, ST239 presence is not included ($i_{239} = 0$ throughout).