Two online only appendices

Appendix A

Simulation Modelling based on the emergence of CMV infection and the CMV Doubling Time

The crude data of the probability of developing CMV was plotted against time since transplantation or since completing successful primary prophylaxis with Valganciclovir (Figure 1).

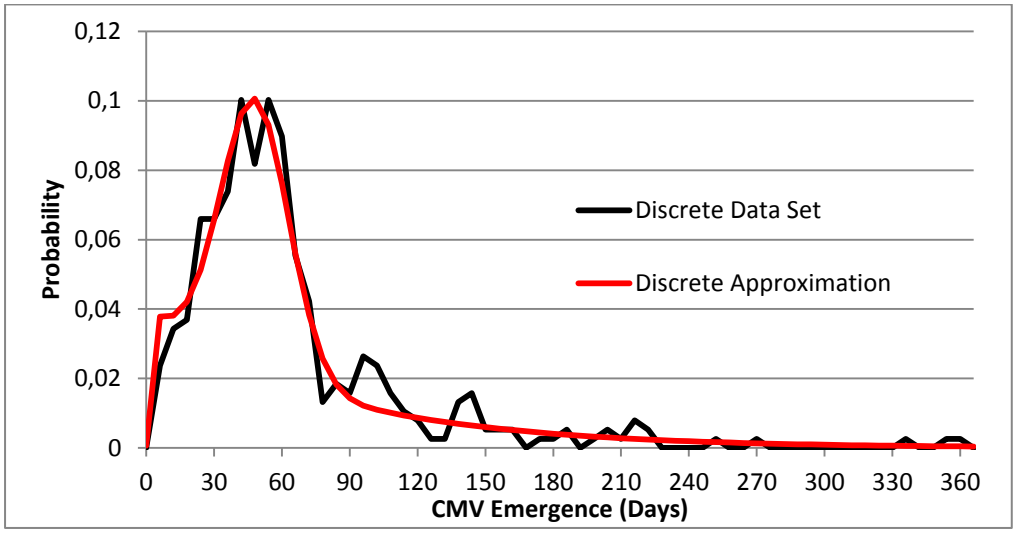

Figure 1: Probability distribution of CMV Emergence and an Approximation to the Data
The y-axis represents the probability of CMV emergence. The x-axis is time in days, where day 0 either represents the day of transplantation (patients treated pre-emptively) or the day that successful primary prophylaxis was stopped.

A combination of a normal distribution and a negative exponential distribution provided a close fit to the data. Different values for the mean and standard deviation for the normal distribution and for a constant for the negative exponential function:

1-exp(-time in days/constant)

were tried. By visual inspection the best fit for the normal distribution had a mean of 45 and a standard deviation of 15 days, and the constant for the negative exponential function was set to 80. This provided an approximation to the crude data (Figure 1).

The combination of these two curves was then used to derive the cumulative probability of CMV with time (Figure 2), where the cumulative approximation shows a close fit to the cumulative data.

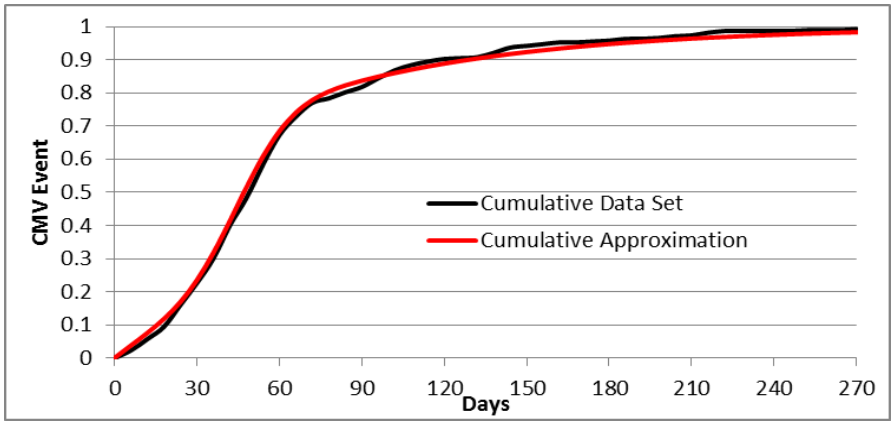

Figure 2: Cumulative Probability of CMV Emergence
The y-axis represents the cumulative probability of the emergence of CMV. The x-axis is time in days, where day 0 either represents the day of transplantation (patients treated pre-emptively) or the day that successful primary prophylaxis was stopped. Within 120 days after transplantation or stopping successful primary Valganciclovir prophylaxis, more than 90% of the CMV infections have emerged in the cohort.

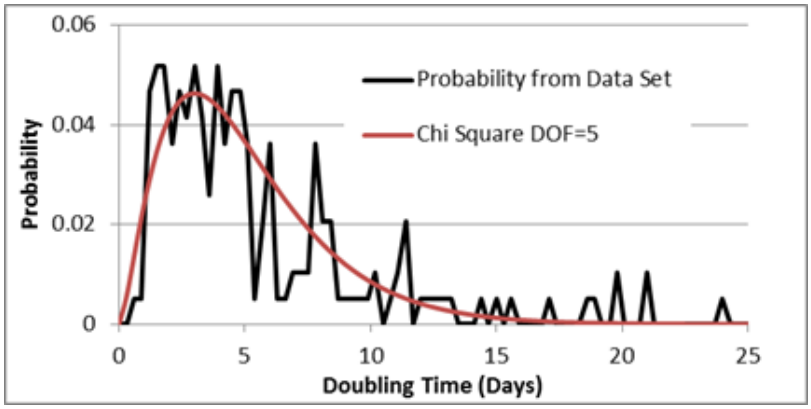

Figure 3: Probability distribution of the observed CMV Doubling Time
The y axis is the probability of doubling times, and the x-axis represents the doubling time in days. The probability distribution of the observed doubling times in the cohort (black line) fits with a Chi-Square distribution with five degrees of freedom (red line).

This process was repeated with CMV Doubling Time data. The crude data was also used to plot the probability of observing the given doubling time (figure 3) and a chi-squared distribution with 5 degrees of freedom that followed the crude data points provided a suitable distribution (Figure 3), which was transformed into the cumulative probability (Figure 4). It should be noted that figure 4 is also included in the manuscript, together with the depiction of the four different quartiles of doubling time.

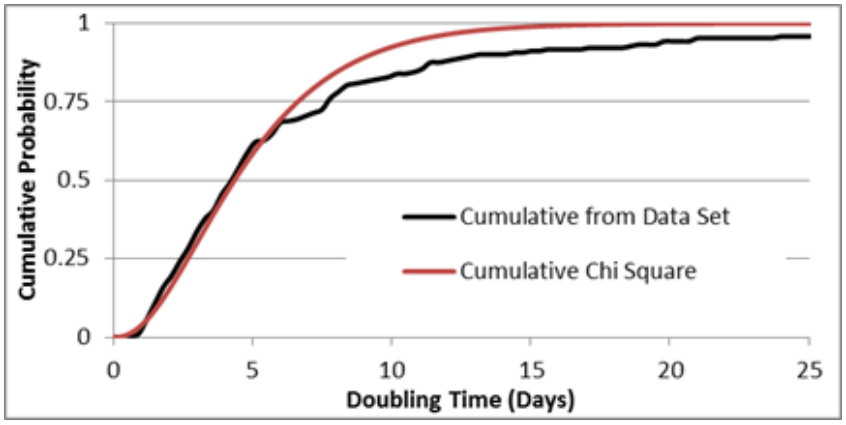

Figure 4: Cumulative Probability of CMV Doubling Time in the cohort
After transformation of the crude data the cumulative probability (y-axis) of the doubling time in days (x-axis) could be plotted. Again the cumulative chi-square distribution (red line) was a good approximation to the observed data on doubling time from the cohort (black line).

Development of a model using simulation techniques

The theoretical distributions, modelled to the crude data, were then used as part of the simulation process. Consider a theoretical patient on the first day post-transplant with an assumed viral load of less than 273 IU/mL (i.e., undetectable and assumed negative). From the crude data, 28.6% are estimated to develop CMV infection within our population (379/1325). A random number is generated from a uniform distribution (U[0,1]) to represent the probability of CMV infection. When the random number is less than 28.6%, they are assumed to develop CMV at some point, a second random number is generated from a uniform distribution (U[0,1]), representing the cumulative probability of developing CMV, given that we know the person will develop CMV. For example, if the random number generated is 0.5, the time at which there is a 50% probability of developing

CMV is 48 days. The viral load at this time is assumed to be 300 and the replication is assumed to start at this time. In the same way, another random number is generated from the uniform distribution (U[0,1]) to select the doubling time from the cumulative distribution in Figure 4. For example, if the random number is 0.75, the doubling time is assumed to be 6 days.

This process is repeated for each of N simulations; 71.4% will not develop CMV and for the 28.6% who do develop CMV, the random numbers and combinations of distributions determine both the time at which CMV develops and the time to viral load doubling as described above.

In the example given below, the simulation was run for 5000 theoretical patients. For a given day at which testing for CMV starts, and any given interval between testing (periodicity), the simulation model will estimate the proportion of persons whose CMV viral load is estimated to exceed 20,000 copies/ml before their next visit, i.e., the CMV infection has been missed.

Each of the parameters in the simulation shown in Table 1 can be changed. In the first row, the probability that CMV emerges is 0.286, as in the crude data. It should be noted, that if another cohort wishes to determine the optimal screening interval based on its doubling time, the observed doubling time must then first be transformed into a fitting distribution (in the same manner that we choose the Chi-square distribution with 5 degrees of freedom for our doubling time estimate). The failure level set to ≥18,200 IU/mL (corresponding to 20,000 copies/mL) is an arbitrary decision based on the current knowledge of virus load and CMV disease.

Probability that CMV Occurs	0.286
Viral load at CMV Emergence	300
Chi Square Degrees of Freedom or Doubling Time (Da	5
First Test (Day)	21
Test Periodicity (Days)	10
Background (Null Test) Level (<)	300
Failure Level (>)	20000

Run Risk Calc

Table 1: Parameters for mathematical simulation on CMV screening intervals
The probability that CMV infection emerges in the cohort is 28.6%. Viral load at emergence is set to 300 copies/mL; the lower limit of quantification of the COBAS Amplicor CMV PCR kit. The doubling time has been transformed to the Chi square distribution with 5 degrees of freedom. The first test day represent the first day after day 0 (either corresponding to the day of transplantation in case of a solely preemptive approach, or the day of stopping successful primary prophylaxis) that screening should commence from. The testing periodicity represents the screening interval in days. The failure level is set to >20,000 copies/mL.

The viral load at which CMV emerges is set to 300, and the number of degrees of freedom for the chi-squared distribution set to 5, as this provided the best fit to the crude data. In this example, the first test for CMV was at 21 days with 10 days between testing, and the level at which failure occur is set to 20,000 copies/ml. Running this calculation via the blue button gives the results shown in Table 2.

% Failure at first test then any later	1st Test	Later	Total
Percentage Test Schedue Failures	0.50	0.94	1.44

Table 2: Results of mathematical simulation
The output of the mathematical simulation. The proportion to fail the first test (e.g. have a CMV virus load > 20,000 copies/mL, 21 days after day 0) is 0.5%. The proportion at risk of failing any later test with the given test periodicity is 0.94%. The total proportion is 1.44%.

From this simulation, repeated 5000 times, the results show that 0.5% of the population had failed by the time of the first test (21 days), and 0.94% failed over the rest of the follow-up period by only testing every 10 days, giving a total percentage who failed with this testing strategy of 1.44%.

Appendix B

Table. Characteristics of included and excluded cohort of patients for the calculation of the CMV doubling-time.		
Characteristics	**Included recipients N=193**	**Not included recipients N=136**
Age, years (IQR)	46 (32-58)	47 (36-58)
Gender	133/193 (69%) males	78/136 (57%) males
Type of transplantation (tx)		
SOT	74/193 (38%)	56%
HSCT	119/193 (62%)	44%
Risk of CMV		
High	96/193 (50%)	71/136 (52%)
Intermediate/low	97/193 (50%)	65/136 (48%)
Details on CMV infection		
Median time-span from tx to the V_1 sample (IQR), days	55 (42-105)	86 (49-146)
Median V1 (IQR), IU/mL	1,183 (546-3,913)	3,185 (956-19,110)
Median Vmax (IQR), IU/mL	5,460 (1,820-26,390)	4,573 (2,048-21,931)
Table comparing the characteristics of the included and excluded cohorts of recipients for the calculation of the CMV doubling-time.		