

## S1 Appendix

### The model:

For pedagogical reasons we divide the presentation of the model into two phases, one only including cell population dynamics ignoring niche feedback and inflammatory feedback related to cell mortality, in the next step followed by an elaboration of this including the omitted elements. The first phase is essentially the model presented by Dingli & Michor [92] whereas the second phase represents our natural extension.

The basic model without niche feedback and inflammatory feedback related to cell mortality reads,

$$\frac{dx_0}{dt} = (r_x - a_x - d_{x0})x_0 - r_m x_0, \quad x_0(0) = x_{0i}, \quad [x_0=\text{HSC}], \quad (1)$$

$$\frac{dx_1}{dt} = a_x A_x x_0 - d_{x1} x_1, \quad x_1(0) = x_{1i}, \quad [x_1=\text{HMC}], \quad (2)$$

$$\frac{dy_0}{dt} = (r_y - a_y - d_{y0})y_0 + r_m x_0, \quad y_0(0) = y_{0i}, \quad [y_0=\text{MPN SC}], \quad (3)$$

$$\frac{dy_1}{dt} = a_y A_y y_0 - d_{y1} y_1, \quad y_1(0) = y_{1i}, \quad [y_1=\text{MPN MC}], \quad (4)$$

where  $x_{0i}$ ,  $x_{1i}$ ,  $y_{0i}$ , and  $y_{1i}$  denotes the initial number of the respective cells. We emphasize the three possible interpretation of the mutation part of the model: 1) the term  $r_m x_0$  is interpreted as a stochastically variable and  $y_{0i} = 0$ ; 2) it is interpreted as a non-zero continuous variable and  $y_{0i} = 0$ ; or 3) it is interpreted as zero (i.e.  $r_m = 0$ ) and  $y_{0i} = 1$ . In the present work we have mainly adopted the third interpretation.

The full model including the inflammatory feedback related to cell mortality and niche feedback becomes,

$$\frac{dx_0}{dt} = (r_x \varphi_x s - d_{x_0} - a_x)x_0 - r_m s x_0, \quad x_0(0) = x_{0i}, \quad [x_0=\text{HSC}], \quad (5)$$

$$\frac{dx_1}{dt} = a_x A_x x_0 - d_{x_1} x_1, \quad x_1(0) = x_{1i}, \quad [x_1=\text{HMC}], \quad (6)$$

$$\frac{dy_0}{dt} = (r_y \varphi_y s - d_{y_0} - a_y)y_0 + r_m s x_0, \quad y_0(0) = y_{0i}, \quad [y_0=\text{MPN SC}], \quad (7)$$

$$\frac{dy_1}{dt} = a_y A_y y_0 - d_{y_1} y_1, \quad y_1(0) = y_{1i}, \quad [y_1=\text{MPN MC}], \quad (8)$$

$$\frac{da}{dt} = d_{x_0} x_0 + d_{y_0} y_0 + d_{x_1} x_1 + d_{y_1} y_1 - e_a a s, \quad a(0) = a_i, \quad [a=\text{dead cells}], \quad (9)$$

$$\frac{ds}{dt} = r_s a - e_s s + \text{inflammatory load}, \quad s(0) = s_i, \quad [s=\text{inflammatory level}], \quad (10)$$

Here the feedback from the niches in the bone marrow affecting the self-renewal rates are represented by decreasing Hill-functions in the number of stem cells,

$$\varphi_x(x_0, y_0) = \frac{1}{1 + (c_{xx}x_0 + c_{xy}y_0)^2} \quad \text{and} \quad \varphi_y(x_0, y_0) = \frac{1}{1 + (c_{yx}x_0 + c_{yy}y_0)^2}$$

The coupling parameters  $c_{ij}$  represent the strength by which the amount of the  $j$ 'th stem cell type affects the growth of the  $i$ 'th stem cell type. These expressions are in analogy with Gentry et al [86] but differentiate slightly from the Michaelis-Menten expressions used by Michor et al [100] and Stiehl et al. [88,89]. Furthermore, an age dependence may be included in the mutation part of the model as a constant age-dependent function multiplied on the term  $r_m s x_0$  [100]. The numerical values of the model parameters are chosen to fulfil hematopoiesis (steady state) when  $r_m=0$  and no MPN cancer cells are present.

### Model calibration and validation:

The model is inspired by Dingli & Michor [92]. However, we have adjusted some of these to obtain more appropriate saturation levels in agreement with data and clinical experiences. First, the model is calibrated to the situation of no MPN cancer cells, i.e.  $y_{0i} = 0$  and  $y_{1i} = 0$ . In general at steady state,

$$x_1 = \frac{a_x A_x}{dx_1} x_0,$$

$$y_1 = \frac{a_y A_y}{dy_1} y_0,$$

$$a = \frac{-\frac{e_a}{e_s} BIS + \sqrt{\left(\frac{e_a}{e_s} BIS\right)^2 + 4r_s \frac{e_a}{e_s} D}}{2r_s \frac{e_a}{e_s}},$$

and

$$s = \frac{k(r_s a + BIS)}{e_s}$$

where  $BIS$  denotes the baseline exogenous inflammation load and

$D = dx_0 x_0 + dx_1 x_1 + dy_0 y_0 + dy_1 y_1$  is an abbreviation for the total of dying cells per time.

From the above, the parameter values have to fulfil these equations at  $t=0$  if we assume start in a

healthy steady state, e.g.  $\frac{a_x A_x}{dx_1} = \frac{x_1}{x_0} \approx 10^7$ . Thus, the involved parameters have been chosen

accordingly. We expect the number of HSC to be approximately  $10^4$  and that of HMC to be

approximately  $10^{10}$  in the healthy steady state. These choices are compromises between reported

values: Gentry & Jackson [86] use  $10^4$ , respectively  $10^{10}$ , Stiehl et al. [88,89] use  $8 \cdot 10^7$  per kg,

respectively  $2 \cdot 10^9$  per kg, Dingli & Michor [92] use  $2 \cdot 10^4$ , respectively  $10^{12}$ ; and Haeno et. al. [78]

use  $5 \cdot 10^5$  for the number of HSC, see Table 1 below. From the steady state condition we also have the

number of dead cells to be  $a = \frac{dx_0 x_0 + dx_1 x_1}{e_a s} \approx 700$ .

Regarding the inflammatory level ( $s$ ) we have chosen the value 3.61 pg/mL. The inflammatory level  $s$

is an abstraction. Hence, the values are not known and only a relative change in  $s$  matters in the model.

## Tables

**Table A. Inflammatory cytokines /C-Reactive Protein**

Disease/Cytokine	Normal	ET	PV	PMF	Reference
IL-1beta (pg/ml)	4.0	ND	5.9	10.8	44,45 (PV,PMF)
IL-1RA (pg/ml)	203.2	ND	342.4	552.0	44,45 (PV,PMF)
IL-2R (pg/ml)	216.9	ND	250.4	555.9	44,45 (PV,PMF)
IL-6 (pg/ml)	0.6	ND	11.8	6.2	44,45 (PV,PMF)
IL-8 (pg/ml)	3.2	ND	10.2	14.3	44,45 (PV,PMF)
IL-10 (pg/ml)	4.8	ND	6.5	12.5	44,45 (PV,PMF)
IL-12 (pg/ml)	100.2	ND	156.0	191.8	44,45 (PV,PMF)
C-Reactive Protein (CRP)	< 2.9	5.6	8.5	12.1	Unpublished series of CRP at the time of diagnosis in 27 ET patients , 38 PV patients and 42 PMF patients
LDH (U/L)	152	ND	226	935	19

**Table B. Default Parameters.**

Parameter	Value	Units	Parameter	Value	Units
$r_x$	$8.7 \cdot 10^{-4}$	day <sup>-1</sup>	$r_y$	$1.3 \cdot 10^{-3}$	day <sup>-1</sup>
$a_x$	$1.1 \cdot 10^{-5}$	day <sup>-1</sup>	$a_y$	$a_x$	day <sup>-1</sup>
$A_x$	$4.7 \cdot 10^{13}$	-	$A_y$	$A_x$	-
$d_{x0}$	$2 \cdot 10^{-3}$	day <sup>-1</sup>	$d_{y0}$	$d_{x0}$	day <sup>-1</sup>
$d_{x1}$	129	day <sup>-1</sup>	$d_{y1}$	$d_{x1}$	day <sup>-1</sup>
$c_{xx}$	$7.5 \cdot 10^{-5}$	-	$c_{yx}$	$c_{xx}$	-
$c_{xy}$	$c_{xx}$	-	$c_{yy}$	$c_{xx}$	-
$e_s$	2	day <sup>-1</sup>	$r_s$	$3 \cdot 10^{-4}$	day <sup>-1</sup>
$r_m$	0 or $2 \cdot 10^{-8}$	day <sup>-1</sup>	$e_a$	$2 \cdot 10^9$	day <sup>-1</sup>
Inflammation (pg/ml)	7	day			

$r_x$  = self renewal rate of HSC,

$a_x$  = proliferation rate of HSC,

$A_x$  = multiplication factor which describe the ration of the rate of produced HMC and  $a_x$  lumping the hematopoietic progenitor cell dynamics into one factor  $A$ ,

$d_{x0}$  = dead rate of HSC,

$d_{x1}$  = dead rate of HMC,

$c_{xx}$  = factor describing the strength by which the HSC inhibits their own self renewal,

$c_{xy}$  = factor describing the strength by which the MPN SC inhibits self renewal of the HSC,

$r_y$  = self renewal rate of MPN SC,

$a_y$  = proliferation rate of MPN SC,

$A_y$  = multiplication factor which describe the ration of the rate of produced MPN MC and  $a_y$  lumping the MPN progenitor cell dynamics into one factor A,

$d_{y0}$  = dead rate of MPN SC,

$d_{y1}$  = dead rate of MPN MC,

$c_{yx}$  = factor describing the strength by which the HSC inhibits self renewal of the MPN SC,

$c_{yy}$  = factor describing the strength by which the MPN SC inhibits their own self renewal,

$e_s$  = elimination rate of cytokines (more precisely the abstract quantity  $s$ ),

$r_s$  = rate by which the dead cells up-regulate the amount of phagocytic cells per dead cell,

$e_a$  = elimination rate of dead cells per cytokine (dead cells is down-regulated as a second order elimination process given by  $-e_a \cdot a \cdot s$ ),

$r_m$  = effective MPN mutation rate

**Table C: Hematopoietic Stem Cell (HSC), MPN Stem Cell (MPN SC), Hematopoietic Mature Cell (HMC), and MPN Mature Cell (MPN MC ) counts under various circumstances with references.** Values are our calculated values from the values given in [88,89] in counts per kg assuming a weight of 80 kg [89]. The two last rows represent our expected values, which have been used for validation in the Figures. Following Figure 1 illustrates model prediction and the validation by data.

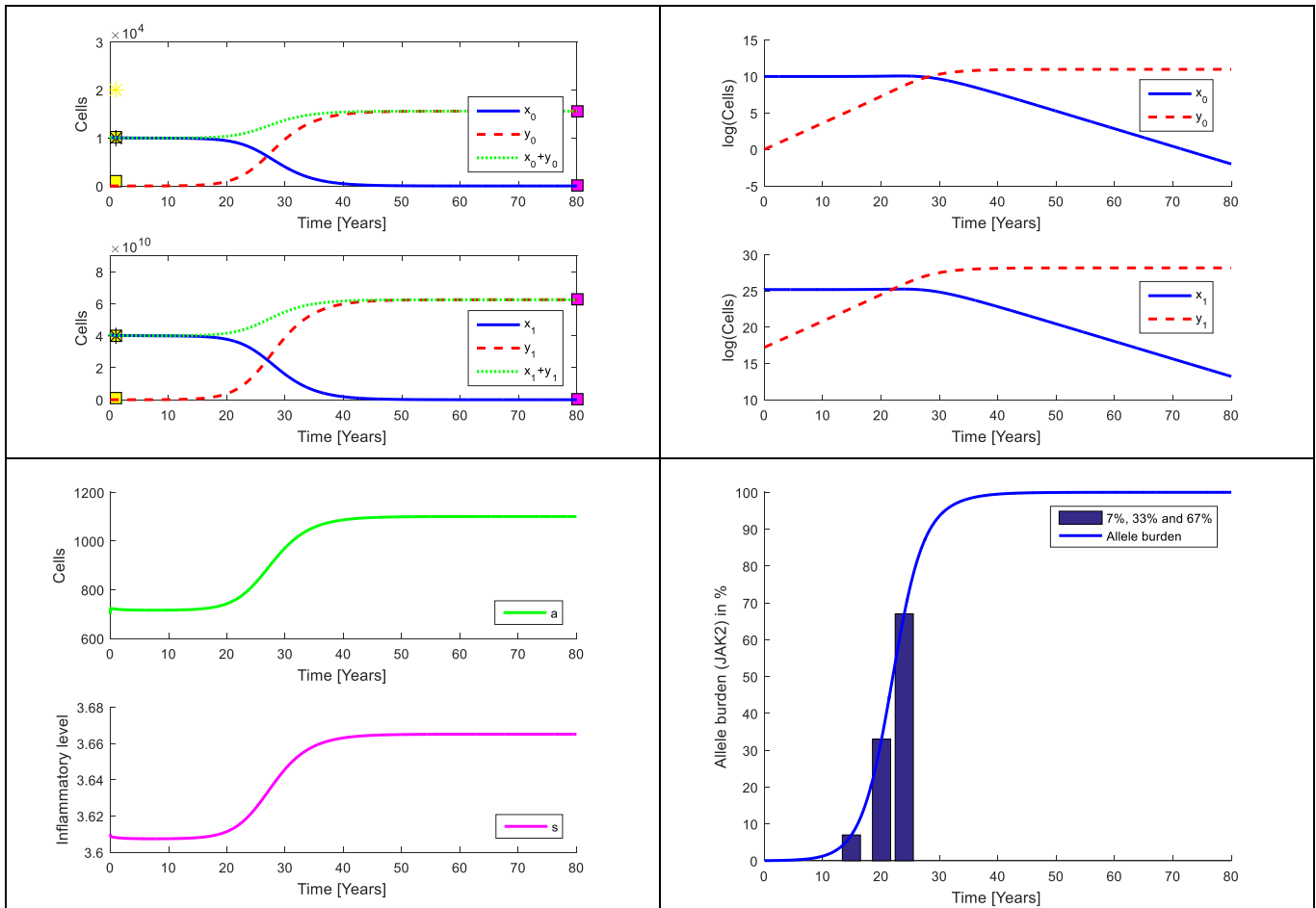
Ref.	HSC	MPN SC	HMC	MPN MC	a	s
[86]	$10^4$	$>10^{10}$	$10^{10}$			
[88,89]	$6.4 \cdot 10^9$		$1.6 \cdot 10^{11}$			
[78]	$5 \cdot 10^5$					
[92]	$2 \cdot 10^4$	$>10^5$	$10^{12}$	$>10^{13}$		
Our “no cancer values”	$1.01 \cdot 10^4$	0	$4 \cdot 10^4$	0	699.9	3.61
Our MPN values	0	$1.56 \cdot 10^4$	0	$6.25 \cdot 10^{10}$		

**Table D. Initial values.**

Variable	$x_0$ (HSC)	$x_1$ (HMC)	$y_0$ (MPN SC)	$y_1$ (MPN MC)	a (Dead Cells)	s (Inflammatory level)
Initial value	$1.01 \cdot 10^4$	$3.84 \cdot 10^{10}$	0	0	699	3.61 pg/ml

## Additional Graphs

**Figure S1**

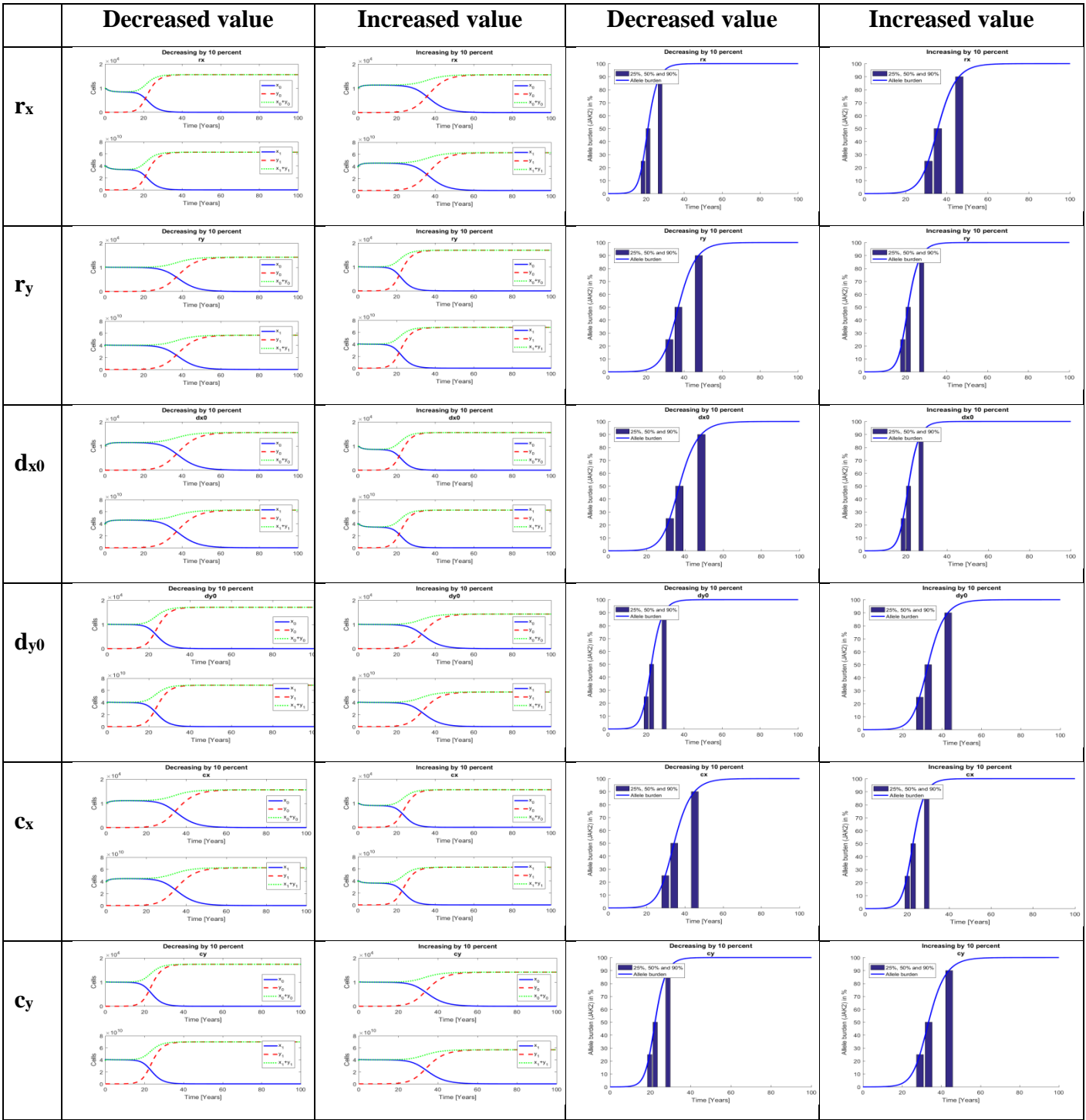


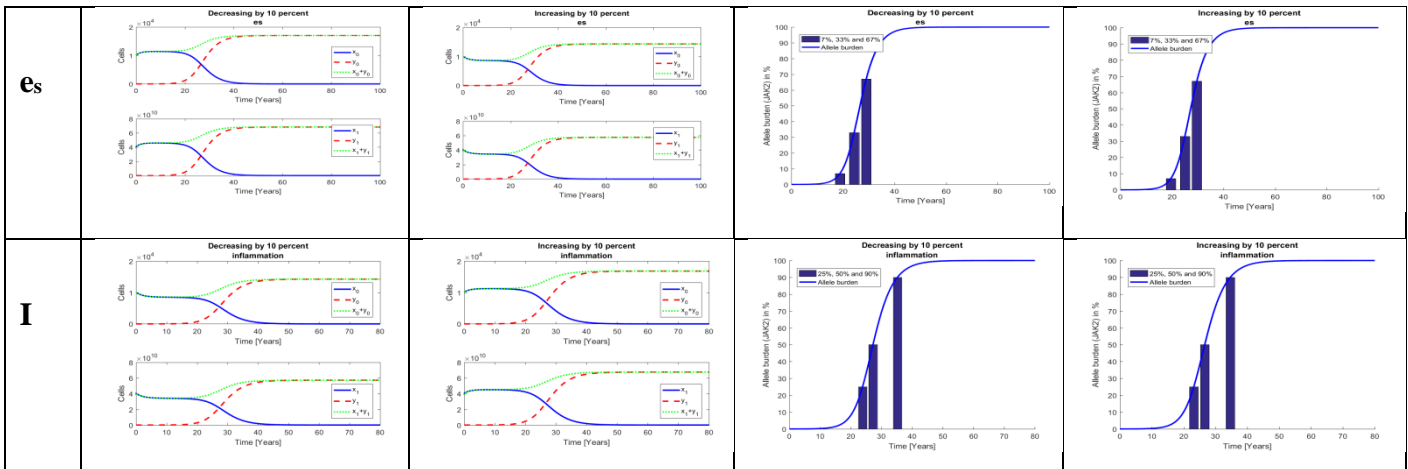
**Figure S1. Upper left:** HSC ( $x_0$ ), MPN SC ( $y_0$ ), HMC ( $x_1$ ), and MPN MC ( $y_1$ ) predicted curves versus time. **Upper right:** Shows the logarithm of the corresponding curves shown to the left. Strait sections of curves are when the original curves change exponentially. **Lower left:** Green curves are the amount of dead cells (a) and magenta is the inflammatory level. **Lower right:** The JAK2V617F allele burden curve with 7% (ET), 33% (PV), and 67% (PMF) emphasized by blue columns.

### **Sensitivity Analysis**

The analysis enables to analyze those parameters which affect the output mostly and at the same time addressing how they are affected if they were varied one at a time keeping the rest fixed at their nominal values. We choose 10% variation in the values of parameters and simulate the effect on cell counts, time intercept (i.e. time from 0 until  $x_0=y_0$ ) and acquired percentage of the JAK2V617F allele burden. The parameters which have the greatest impact upon the output of the model are illustrated in Figure 2 below. These include self-renewal rates and death rates of HSC and MPN SC as well as  $c_x$  and  $c_y$ .





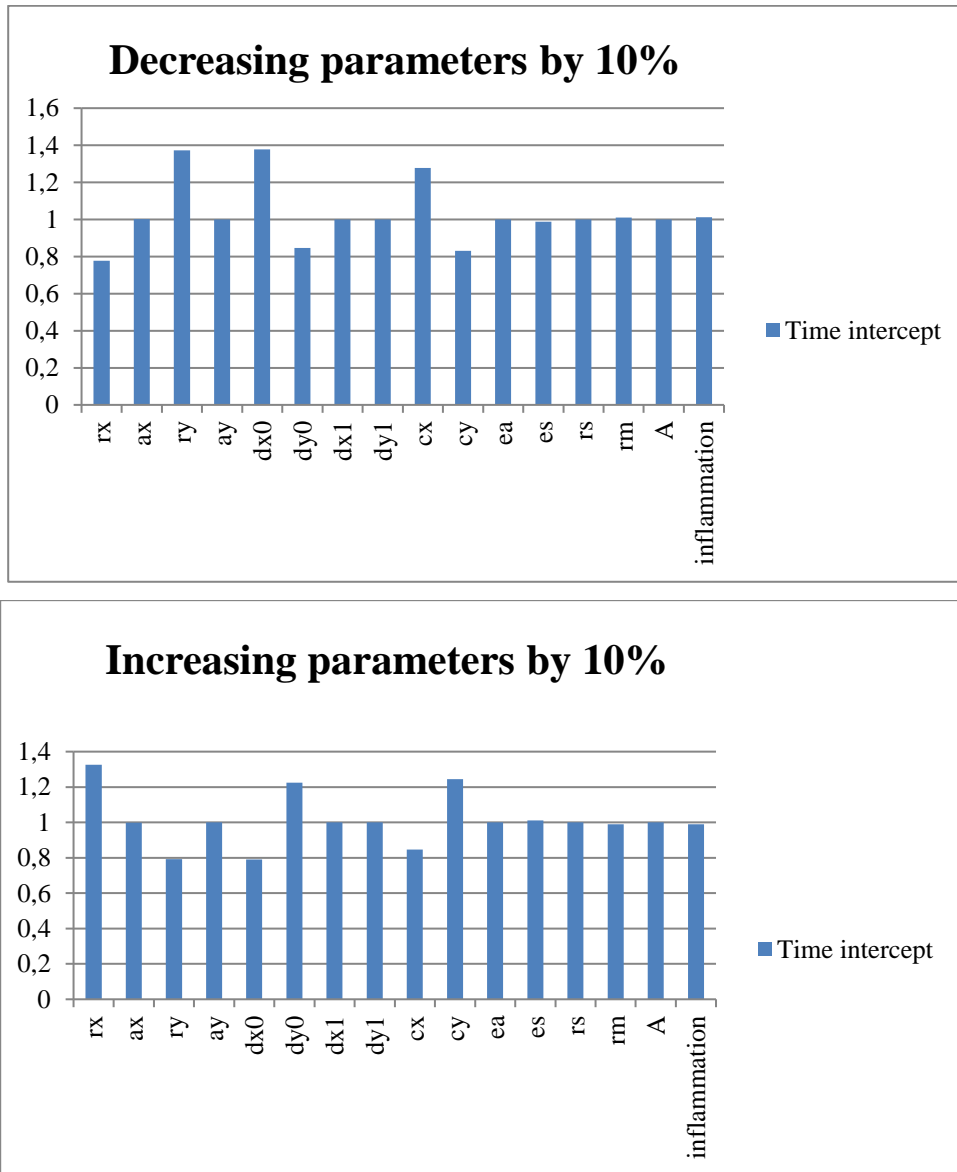


**Figure S2.** Sensitivities for selected parameters. 1<sup>st</sup> column shows the parameters, 2<sup>nd</sup> column shows how the output curves of the cell numbers are changed when the parameter is decreased by 10%, 3<sup>rd</sup> column shows the corresponding curves for a 10% increase in the parameter, and the two last columns the corresponding values for the JAK2V617F allele burden.

From Figure S2 we observe which of the parameters have the most pronounced impact upon the lag-time. The 1<sup>st</sup> and 3<sup>rd</sup> columns show the effect of decreasing the parameter values by 10%. Here  $r_y$  and  $d_{x0}$  increase the time intercept equally well but more markedly than other parameters. In addition, by increasing parameter values (in the 2<sup>nd</sup> and 4<sup>th</sup> column),  $r_x$  has the most pronounced effect on lag-time followed by  $c_y$  and  $d_{y0}$  which influence the output almost similarly.

A more clear picture of change in lag-time can be seen from the following bar charts in Figure S3.

**Figure S3:** Bars above and below 1 show increase and decrease in duration to time-intercept (i.e. until  $x_0=y_0$ ) respectively. Here, parameters labels the x - axis and average values w.r.t time as y - axis



In addition to lag-time some parameters also change cell counts especially number of MPN cells,  $y_0$  and  $y_1$ . This is illustrated in Figure S4.

**Figure S4.** Bars above and below 1 show increase and decrease in cell counts, respectively, when a parameter is varied (values equal to 1 are cell counts for the default parameter values). Parameters are shown along the x-axis and the average cell counts along the y-axis. For each parameter four columns are shown; The first (D:y0 in light blue) shows the effect on  $y_0$  when the parameter is decreased by 10%, the second (I:y0 in brown) shows the effect on  $y_0$  when the parameter is increased by 10%, the third (D:y1 in green) shows the effect on  $y_1$  when the parameter is decreased by 10%, and fourth (I:y1 in dark blue) shows the effect on  $y_1$  when the parameter is increased by 10%.

