H. Hoffmann, C. Thiede, I. Glauche, I. Roeder
"Differential response to cytotoxic therapy explains treatment dynamics of AML patients: insights from a mathematical modelling approach"
*Journal of the Royal Society Interface* (2020)

# Supplementary Material

August 6, 2020

## 1 Derivation of the correction factor

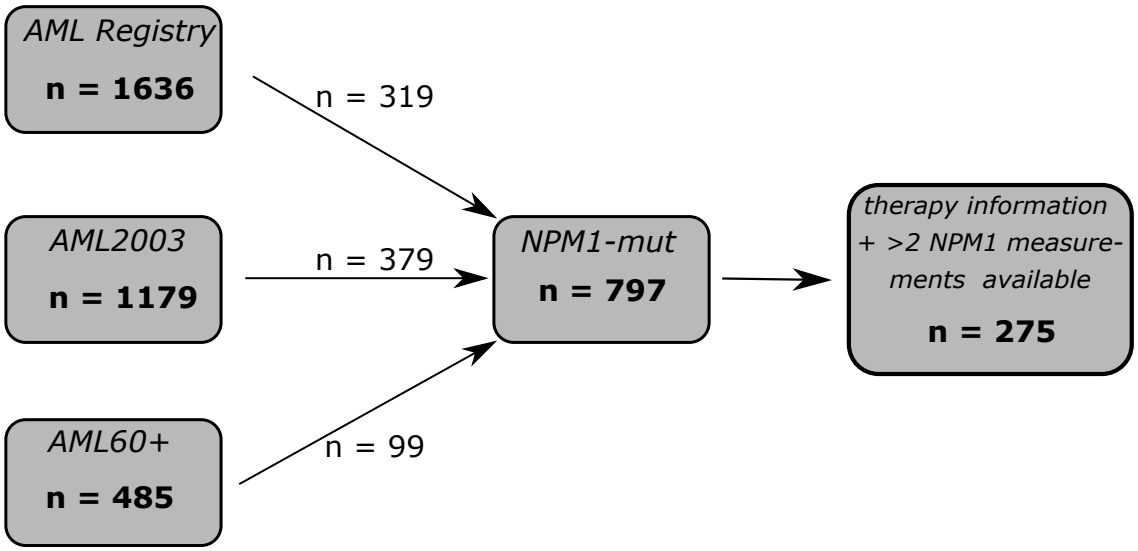
All qPCR data, reported in this study, was measured on either of two different PCR machines, namely a TaqMan 7500 and a Lightcycler480. Given the slightly different setups, we observed a shift between the log10 NPM1 values measured with the two machines (with the Lightcycler-values tending to be higher). For this reason we introduced a correction factor to make the measured values comparable. To derive this factor we used all n=440 samples for which measurements with both machines were available (including blood measurements and patients that were not part of this study). The factor was then calculated as follows:

$$factor = 10^{mean(log_{10}(v_{Lc}) - log10(v_{TM}))}$$

with $v_{Lc}$ being the Lightcycler values and $v_{TM}$ the TaqMan values. This factor of 2.55 was than used to correct measuremnts obtained with TaqMan PCR.

## 2 Patient selection

Model fitting to time course data requires a minimal set of quality criteria. For this reason we excluded patients with obvious discrepancies in the data, such as a missing chemo therapy cycle information, or too few data points. 137 of the 275 patients received stem cell transplantation and all measurements *after* the transplantation were censored. For the model analysis we only considered those patients for which at least 3 measurements were left after censoring. An overview of the data aggregation and selection is provided in the following flow chart:




# 3 Response characteristics

In our previous paper we derived characteristic parameters to describe the dynamics of *NPM1* time courses of AML patients [1]. They characterise the patient-specific treatment response as well as the remission and relapse behaviour:

The **elimination slope ($\alpha$)** was defined as the average slope from the initial measurement until the first measurement below 1% during the first six months of primary treatment (induction + consolidation therapy), using a linear regression of *NPM1/ABL* values. At least two eligible data points were necessary to calculate the elimination slope $\alpha$.

The **NPM1 level after primary treatment ($n$)** was defined as the lowest measurement within the first 9 months after treatment start.

The **relapse slope ($\beta$)** quantifies the speed of relapse occurrence. It is defined as the maximum slope between any two consecutive measurements obtained from all sequentially increasing measurements around the relapse threshold.

The **molecular relapse time ($d$)** was defined as the approximated time point when the *NPM1/ABL* value exceeds the relapse threshold of 1 % using a linear regression between the last point below and the first point above the threshold.

# 4 Model comparison

In order to motivate and justify our choice of the model setup and especially the choice of the two free parameters used for model fitting, we systematically compared representative models with different free parameters.

A visual summary of the comparison can be found in Supplementary Figure S5A.

- Model 1 was the model used throughout the study with the leukemic proliferation rate ($p_l$) and the leukemic activation rate ($t_l^A$) as the free parameters.

- In model 2 we added an additional free parameter: the chemotherapeutic kill rate ($c$), as varying chemo-sensitivities between patients.
- In model 3 the leukemic inactivation rate ($t_l^Q$) was added to the free parameters (adding up to a number of 4).
- In model 4 we investigated, whether we can improve the quality of the model, when neglect the differentiation of all cells in the model by using model 1 and additionally setting the differentiation rates ($d_{l/h}$) to zero.
- In model 5 we investigated, whether a differential chemotherapeutic effect on healthy and leukemic cells (as different parameters $c_l$ and $c_h$) could improve the fit to the data.
- In model 6 we chose the chemotherapeutic kill rate ($c$) instead of the leukemic activation rate ($t_l^A$) as the second free parameter.
- In model 7 we chose the chemotherapeutic kill rate ($c$) instead of the leukemic activation rate ($t_l^A$) as the second free parameter.

For the quantitative assessment we took a dual approach. First, we used Akaike's information criterion (AIC, calculated within the software Monolix, in which we included the free model parameters as random effects) as a measure to quantify which model optimally describes the average patient behaviour as well as the number of the estimated parameters. The AIC integrates both, the fitting quality (i.e. the residual error) and the number of free model parameters. Second, we evaluated how well the model parameters are structurally identifiable on the level of individual patients. To this end, we used a simulated time course and evaluated the corresponding likelihood landscape. Based on the global optimization using the AIC approach (Supplementary Figure S5A), we could already narrow the set of optimal models to the first three (models 1 - 3). Comparing the likelihood landscapes for these models we reason that the estimation of a third parameter, namely the chemotherapeutic kill rate c severely limits the identifiability (Supplementary Figure S5B). On the basis of these considerations we opt for a minimal model in which the rate $p_l$ (reflecting the aggressiveness of the leukaemic clone) and the leukaemic activation rate $t_l^A$ (reflecting the individual chemosensitivity) are structurally identifiable and suited to also obtain a close to optimal parameterization for the whole population of patients.

# 5 Sensitivity analysis

The overall results of the model fitting depends not only on the individual choice of the free parameters but is also influenced by the choice of the global parameters, which are set equal for all patients. In order to analyse the sensitivity of the model results on the choice of those we systematically varied them around their prespecified values. In particular, we performed a sensitivity analysis for the proliferation rate of the healthy cells ($p_h$), the carrying capacities of both states ($K_{A/Q}$) and the chemotherapeutic kill rate ($c$). For each univariate parameter variation, we again fitted the model to all patients and calculated the mean absolute error (MAE) for this particular setting. The results of this analysis are summarized in Supplementary Figure S6.

For the proliferation rate of the healthy cells $p_h$ we also had to adjust the lower and upper bound of the proliferation rate of leukemic cells ($p_l$), as our model was build in the assumption that the leukemic cells proliferate *faster* than the healthy cells. We observe that the actual choice of the parameter $p_h$ has no major impact on the MAE (Supplementary Figure S6A). Looking at another important model result, namely the estimation of the 1-year relapse probability (see Figure 3F), no qualitative differences appear for different values of $p_h$. This is vidsually shown in Supplementary Figure S7.

When comparing the MAE using different combinations of the carrying capacities $K_{A/Q}$ (see Supplementary Figure S6B), we observe that their absolute values are not important, as the MAE remains the same for all cases $K_A = K_Q$. There is also no major effect, when the active state capacity $K_A$ exceeds the capacity

of the quiescent state $K_Q$. However, when the quiescent state has a higher capacity than the active state ($K_Q > K_A$) the MAE is visibly increased, indicating that this combination is not a good choice for a model fit.

The value of the chemotherapeutic kill rate $c$ has only minor impact on the overall goodness of fit (see Supplementary Figure S6C). It is only important that the value is chosen sufficiently high (above 0.9 1/day).

# References

[1] Hoffmann H., Thiede C., Glauche I., et al. The prognostic potential of monitoring disease dynamics in NPM1-positive acute myeloid leukemia 2019.