Dynamics of chronic myeloid leukemia response to dasatinib, nilotinib, and high-dose imatinib

by Adam Olshen, Min Tang, Jorge Cortes, Mithat Gonen, Timothy Hughes, Susan Branford, Alfonso Quintás-Cardama, and Franziska Michor

Haematologica 2014 [Epub ahead of print]

doi:10.3324/haematol.2013.085977

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.
Dynamics of chronic myeloid leukemia response to dasatinib, nilotinib, and high-dose imatinib

Adam Olshen1*, Min Tang2*, Jorge Cortes3, Mithat Gonen4, Timothy Hughes5, Susan Branford5, Alfonso Quintás-Cardama3, and Franziska Michor2**

1University of California at San Francisco, San Francisco, CA, USA. 2Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, and Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA. 3M.D. Anderson Cancer Center, Houston, TX, USA. 4Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. 5Department of Genetics and Molecular Pathology, SA Pathology, and University of Adelaide, Adelaide, Australia. *These authors contributed equally to this work. **Author for correspondence. Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA 02215; e-mail: michor@jimmy.harvard.edu.

Abstract

Treatment with the tyrosine kinase inhibitor imatinib represents the standard of care for newly diagnosed patients with chronic myeloid leukemia. In recent years, several second generation inhibitors – such as dasatinib and nilotinib – have become available that promise to overcome some of the mutations associated with acquired resistance in these patients. Despite eliciting similar clinical responses, the molecular effects of these agents on different subpopulations of leukemic cells remain incompletely understood. Furthermore, the consequences of using high-dose imatinib therapy have not been investigated in detail. Here we utilized clinical data from patients treated with dasatinib, nilotinib, or high-dose imatinib, together with a statistical data analysis and mathematical modeling approach, to investigate the molecular treatment response of leukemic cells to these agents. We found that these drugs elicit very similar responses if administered front-line. However, the patient population displays significantly different kinetics when treated second line, both in terms of differences between front-line and second line treatment for the same drug, and among agents when used second-line. We then utilized a mathematical framework describing the behavior of four differentiation levels of leukemic cells during therapy to predict the treatment response kinetics for the different patient
cohorts. The dynamics of BCR-ABL1 clearance observed in our study suggest that the use of standard or high-dose imatinib or a second generation TKI such as nilotinib or dasatinib elicit similar responses when administered as frontline therapy for patients with CML in chronic phase.

Introduction

The tyrosine kinase inhibitor (TKI) imatinib (STI571, Gleevec; Novartis), administered at 400 mg daily, is the standard front-line therapy for patients with chronic myeloid leukemia in chronic phase (CML-CP)\(^1\). This molecularly targeted therapy leads to a dramatic clinical response: results from the International Randomized Study of Interferon Versus STI571 (IRIS) trial show that 82% of patients achieve a complete cytogenetic response (CCyR)\(^4\). However, a substantial fraction of patients develops acquired resistance to imatinib\(^5\). In order to improve the outcome of patients with early chronic-phase CML, several phase II clinical trials were conducted at the MD Anderson Cancer Center to investigate the treatment effects of new strategies: frontline treatment with high-dose imatinib (800 mg daily) as well as with the second-generation TKIs dasatinib and nilotinib\(^8,9\). Results from these studies suggested the superiority of the use of dasatinib and nilotinib over standard imatinib therapy as frontline treatment. Both dasatinib and nilotinib induced a 98% CCyR rate in patients treated for at least 3 months, with nearly 90% of patients achieving CCyR by 3 months of therapy\(^4,8,9\). Furthermore, randomized phase III trials have demonstrated improved response rates and decreased rates of transformation associated with nilotinib or dasatinib therapy compared to imatinib\(^10,12\). In addition, these TKIs are more potent as inhibitors of the kinase activity of BCR-ABL1 and overcome the resistance imposed by most BCR-ABL1 mutants identified in patients with CML failing imatinib. When nilotinib or dasatinib are used after failure to imatinib therapy, these agents induce CCyR in approximately 50% of patients\(^13\).

Despite the demonstrated superiority of the second-generation TKIs as frontline therapy, several critical questions remain. For instance, the differential effects of dasatinib, nilotinib and high-dose imatinib on different subpopulations of CML cells are incompletely understood. We have previously addressed this question utilizing data from front-line low-dose imatinib-treated patients\(^17,18\). Furthermore, the ability of these agents to eliminate
minimal residual disease remains to be demonstrated. Here we utilized datasets of patients with CML treated in several phase II studies to discern potential differences in the dynamics of molecular response to second-generation TKIs or high-dose imatinib in the context of front-line and post-imatinib failure settings. Our efforts include developing a statistical model that gives a good, low-dimensional representation of the time courses and a mathematical model that estimates biologically interesting parameters. The parameters of these models were then compared between treatments strategies. This approach is useful for understanding the effects of different TKIs on individual cell populations and can also be applied to other cancer types treated with targeted agents.

Methods

Study population and data collection. A total of 290 patients with CML receiving front-line or second-line TKIs were investigated. These included 92 patients treated with dasatinib, 75 treated with nilotinib, and 123 patients treated with high-dose imatinib. All imatinib-treated patients were treated front-line, i.e. right after diagnosis of the disease, while 23 dasatinib–treated and 24 nilotinib-treated patients were administered the respective drug as second-line therapy after progression of disease. All patients received these agents in phase II studies at the MD Anderson Cancer Center. A total of 14 of the 243 frontline patients and 16 of the 47 second-line patients had accelerated phase disease based exclusively on the presence of clonal evolution; the remaining patients were in chronic phase. The assignment of treatment was not done in a randomized fashion; the studies with imatinib and high-dose imatinib were conducted in a sequential fashion. Subsequently, both phase 2 studies using either nilotinib or dasatinib as frontline therapy were conducted in parallel and superseded the study with high-dose imatinib. While patients were not assigned to the latter phase 2 studies in a randomized fashion, they were distributed between both studies so as to sustain a similar rate of accrual.

The levels of BCR-ABL1 transcripts in peripheral blood samples were measured by quantitative real-time PCR (RQ-PCR). The data investigated was the ratio of BCR-ABL1 to the control gene ABL1. BCR-ABL1/ABL1 ratios were expressed in accord with the International Scale. Patients’ samples were obtained to determine the BCR-ABL1/ABL1 ratio ideally at times 0, 3, 6, 9, 12, 18, and 24 months post initiation of therapy. All 290 patients
had measurements at the first five time points. Overall, 32 (11%) patients had five measurements, 41 (14%) patients had six measurements, and 217 (75%) patients had all seven measurements.

A second dataset of 13 patients treated with standard dose imatinib (400mg daily) from the IRIS trial was included as a control. It was quantitated using the same methodology. We only utilized measurements made at 0, 3, 6, 9, 12, 18, and 24 months post initiation of therapy for this analysis. The 15 and 21 month measurements as well as measurements after the 24th month (if any) were neglected to make the data sets comparable. Importantly, only patients with measurements at the first five time points were included.

Data were analyzed using two types of models. The first was a statistical model described in the following part of the manuscript. The second was a mathematical model described partly after the statistical model and partly in the supplement. All analyses were conducted in R. Statistical tests with corresponding p-values less than 0.05 were called significant.

**Results**

To investigate the time courses of BCR-ABL1 transcript levels, we first designed a statistical model with the primary goal of finding the best fit to the data with minimal assumptions. We chose a low-dimensional model in our statistical approach because we sought to obtain simple summaries of the time courses. In addition, since there were at most seven measurements per patient, the data supported only a low-dimensional model. As a first step, we used the natural logarithm to transform our data, since the measurements were on vastly different scales. We added 1 to the observed data before transformation since some of the measured transcript levels were 0. Our initial model on the transformed data was bi-exponential, which is the sum of two scaled exponentials, and is traditional for modeling exponentially-decaying data\(^9\). Such a model contains four parameters per subject, which were considered to be too many. As shown in Supplementary Figure S1, we found that one of the exponentials could be replaced with a constant, which saved a parameter with little reduction in goodness of fit. To save additional degrees of freedom, we examined models that represented aspects of the times courses that were common across subjects. Our
resulting model for the $i$th subject at time $t$ was given by $y_{it} = \alpha_0 + \alpha_1 \exp[-\beta t] + \varepsilon_{it}$, where $\alpha_0$ was the subject-specific intercept, $\alpha_1$ was the subject-specific slope, $\beta$ was the population-based shape parameter and $\varepsilon_{it}$ were independent, identically distributed mean 0 errors. In other words, $\alpha_0$ was the patient-specific amount by which the exponential curve was offset from 0, $\alpha_1$ was the patient-specific magnitude by which the exponential curve was multiplied, and $\beta$ was the common value across all patients that determined the shape of the exponential curve, with larger values implying faster decreases. We did not consider patient-specific $\beta$s because we wanted a low-dimensional model and because trying to estimate them would introduce identifiability issues. We then developed an iterative procedure to fit the model to the data. We obtained an initial estimate of $\beta$ after merging all data from all subjects using non-linear least squares. With $\beta$ fixed, we then estimated the subject-specific parameters ($\alpha_0$ and $\alpha_1$) using traditional least squares. We iterated between estimating the population parameter and the subject-specific parameters until convergence. A z-test was used to test differences in the population parameters between treatment groups.

Estimates of $\alpha_0$ and $\alpha_1$ for subjects were treated as data for further analysis. These estimates were correlated with each other, so joint modeling was necessary. Specifically, multivariate analysis of variance (MANOVA)\textsuperscript{20}, an extension of univariate analysis of variance (ANOVA), was used with treatment drug, front-line or second-line status of the drug, and stage as independent variables and $\alpha_0$ and $\alpha_1$ as dependent variables. Significance was assessed using the Wilks test. Because the dependent variables were not jointly Gaussian, inference was based on 10,000 permutations of the independent variables. The variable that was permuted (treatment, line, or stage) corresponded to the variable on which inference was made.

Once we decided on our model, the primary analysis tested whether the subject-specific parameters were equal, first three-way and then pairwise. Subsequent analyses compared front-line to second-line treatment and chronic to accelerated phase within the dasatinib and nilotinib cohorts and the whole set of hypotheses on only the set of typical curves.

We fit the exponential model described above to the transcript levels of 290 CML patients. The means for the ratios for time points 0, 3, 6, 9, 12, 18, and 24 months post initiation of
therapy were 28.4, 4.3, 2.1, 0.7, 2.6, 1.8, and 0.8. We first fit the model separately for patients treated with dasatinib-, nilotinib- or high-dose imatinib. Note that this investigation utilizes data of patients treated first-line as well as patients treated second-line; we present the analysis for subcohorts later. We estimated the population parameter $\beta$ as 0.69 for dasatinib-treated patients, 0.65 for nilotinib-treated patients and 0.62 for high-dose imatinib-treated patients. Recall that $\beta$ represents the common value that determined the shape of the exponential curve, with larger values implying faster decreases. These parameter estimates were not significantly different in any pairwise comparison between the treatment groups ($p=0.84$ for high-dose imatinib versus nilotinib; $p=0.58$ for high-dose imatinib versus dasatinib; $p=0.85$ nilotinib versus dasatinib). Therefore, we concluded a single population parameter $\beta$, which represents the rate of exponential depletion of BCR-ABL1 values over time, could accurately represent the curves across treatment groups. To estimate this single $\beta$, we combined all data across groups and fit the exponential model again. The resulting $\beta$ was 0.65. The model fit the data well for all three treatment groups, with a median $R^2$ of 0.96 for each. The fits, shown in Figures 1-3, were the basis for further analyses. There are cases shown in Figure 3 that do not follow our exponential model; repeating the analysis without such cases is discussed later.

The subject-specific intercept ($\alpha_0$) and slope ($\alpha_1$) values from the combined fit were dependent variables in MANOVA models; these models examined whether there were independent variables with which the dependent variables were associated. Results are displayed in Table 1. Recall that $\alpha_0_i$ represents the patient-specific amount by which the exponential curve was offset from 0, and $\alpha_1_i$ was the patient-specific magnitude by which the exponential curve was multiplied. A three-way comparison among the treatment groups was significant ($p=0.004$), so the hypothesis that the subject-specific parameters were the same across groups was rejected. Pairwise comparisons between treatment groups showed no significant difference between high-dose imatinib and dasatinib ($p=0.14$). However, it did show a significant difference between imatinib and nilotinib ($p=0.001$) and between nilotinib and dasatinib ($p<0.001$). The average intercept for nilotinib patients was higher (0.52 versus 0.29) and the average slope was lower (2.34 versus 2.51) relative to dasatinib patients. Note that these results were obtained when analyzing both front- and second-line patients. When utilizing only front-line patients, the three-way comparison among treatments was not significant ($p=0.72$). However, front-line versus second-line status led
to significantly different results (p<0.001). Here front-line treated patients had a lower average intercept (0.14 versus 1.04) and a higher average slope (2.70 versus 1.44) than second-line treated patients. Within the dasatinib and nilotinib cohorts, front-line was significantly different than second-line (p<0.001 for both). The directions of slope and intercept were the same as in the merged front-line and second-line groups.

Overall, there were no significant pairwise differences between treatment groups within the front-line or within the second-line cohorts. While there was a significant difference between chronic and accelerated phase (p<0.001), there was no significant difference between stages when analyzing only front-line patients. A bivariate model with both treatment and front-line status was significant for both variables, indicating treatment and front-line status were independent prognostic factors.

Based on the preceding analysis, we drew a few conclusions. First, the shape of the curves was similar across treatment groups. Second, the driver of the difference between curves was front-line versus second-line status. The front-line curves tended to have lower intercepts but higher scale factors. Third, after taking into account front-line versus second-line status, there was still an impact of treatment.

We also conducted similar analysis on the 400mg imatinib patient cohort. Here the population parameter \( \beta \) was taken from the earlier three-group analysis and the subject-specific parameters were compared. The four-way comparison among groups was again significant (p<0.001). Pairwise comparisons between each of the three groups and standard dose imatinib were all significant.

We next explored the robustness of our modeling. There were two main ways in which a subset of curves differed from the typical exponential pattern. Some curves increased over a portion of the time interval, rather than decreasing or staying level. Other curves showed low values throughout. To identify the increasing curves, we used the heuristic of two measurements in the log-transformed data that were more than one unit above the previously observed minimum. The reason for choosing this criterion was that small increases were possible due to random variation in small values. In addition, occasional faulty measurements were possible. There were 22 (7.6%) such increasing curves. The
percentages by treatment groups were 3.3% for high-dose imatinib-, 8.7% for nilotinib-, and 13.3% for dasatinib-treated patients. Pairwise, the only significant difference in percentage was between high-dose imatinib- and dasatinib-treated patients (p=0.01). Among the front-line treated patients, 3.3% had increasing values, while among second-line treated patients, 29.8% had increasing values. This difference was statistically significant (p<0.001).

Low value curves were those with maximum values less than one in the log-transformed data. There were 22 samples (7.6%) of this type. The percentages were 6.5% for imatinib-, 8.0% for nilotinib-, and 8.7% for dasatinib-treated patients. None of these percentages were significantly different from each other by pairwise comparison. These percentages were 7.4% for front-line treatment and 8.5% for second-line treatment, which was again not significantly different.

The analyses above were repeated on just the patients treated at the MD Anderson Cancer Center with the increasing and low value curves eliminated. This approach left 246 out of the original 290 patients. The population parameter \( \beta \) was again not significantly different among treatment groups based on pairwise comparison, so a single model was fit to all curves. For the MANOVA model, the three-way comparison among the treatment groups was only borderline significant (p=0.06). Pairwise differences this time were all not significant. The three-way comparison for just front-line patients was again not significant (p=0.38). Front-line versus second-line status was again significant when comparing all patients or only within nilotinib or dasatinib patients (p<0.001 for all three). The bivariate model with both treatment and front-line status was again significant for both variables, indicating that treatment and front-line status were independent prognostic factors. Overall, the results were robust to inclusion or exclusion of atypical curves.

We then utilized a previously designed mathematical model describing the differentiation hierarchy of hematopoietic cells\textsuperscript{17,18} to analyze the data. This model was created to relate the available data on \( BCR-ABL1 \) transcript levels in peripheral blood to the kinetics of other, unobservable differentiation levels of leukemic cells. Through this model, we thus aimed to investigate the underlying mechanism of treatment response by studying the different response kinetics of separate leukemic subpopulations. In the context of this model, stem
cells on top of the hierarchy give rise to progenitor cells, which produce differentiated cells, which in turn produce terminally differentiated cells. This differentiation hierarchy applies to normal and leukemic cells. The model assumes that the BCR-ABL1 oncogene increases the rate at which progenitors and differentiated cells are being produced. Molecularily targeted therapy counteracts the effects of BCR-ABL1 by reducing the differentiation rates and possibly reducing the growth rate of leukemic stem cells. The basic model together with a description of the parameters is displayed in Supplementary Table S1.

This model represents the four kinetically dominant subpopulations in the hematopoietic differentiation hierarchy; in reality, this hierarchy includes a larger number of distinct differentiation levels. However, for the purposes of explaining the TKI response dynamics, only four populations are necessary to include in a mathematical model since those are the kinetically dominant populations. Similarly, if each subpopulation consists of many clones of the same differentiation stage, which may have distinct growth, differentiation, and death kinetics, then the predictions of the model with regard to the question addressed in this paper does not change. The model then describes the dominant clone within the respective subpopulation at any given time.

To investigate the parameters of the mathematical model, we again utilized the data of newly diagnosed patients who were treated with front-line dasatinib, nilotinib or imatinib (800mg or 400mg daily) outlined above. Any value of zero, i.e. any value below the detection baseline of the PCR assay, was replaced with 0.00001. The replacement value of 0.00001 was chosen because the minimum value in the data was 0.00004. This replacement was needed for the logarithmic transformation and is consistent with the previous use of this model. We later conducted sensitivity analyses on the choice of this replacement value using 0.00002 and 0.000005 and obtained consistent conclusions. For each treatment cohort, we investigated two models and selected the one that best fit the data using $R^2$ and a permutation test. The two models analyzed were (i) an exponential model, which predicts that the leukemic cell burden declines at a single exponential rate; and (ii) a bi-phasic exponential model, which predicts that the BCR-ABL1/ABL1 ratio declines at two exponential slopes with a turning point. The estimated slopes were then incorporated into our mathematical framework to predict the dynamics of treatment responses of the four cohorts. Note that the model fitting and $R^2$ calculation used the logarithmically
transformed data (Supplementary Information). Once we decided on our model, the primary analysis tested whether the estimated slopes were equal, first three-way and then pairwise.

When applying this approach to our CML patient cohorts, we found that the bi-phasic exponential model was the better fitting of our two models for all four patient cohorts (Supplementary Table S2). Besides the analysis of the entire cohort, we also performed individual model fitting to compare, for each individual patient, the fit of the single-phasic and bi-phasic exponential models. Based on the permutation test procedure (using the joinpoint software, http://surveillance.cancer.gov/joinpoint/), we obtained a better fit of the bi-phasic model for 35 out of the 69 patients in the dasatinib cohort, 26 out of the 51 patients in the nilotinib cohort, 54 out of the 123 patients in the high-dose imatinib cohort, and for all 13 patients in the standard dose imatinib cohort (Supplementary Table S3). The first and second slopes of depletion as well as the turning point for those bi-phasic patients in the four cohorts are summarized in Table 2.

There was no significant difference in the number of patients exhibiting 1-phasic versus 2-phasic trends among the dasatinib, nilotinib and high-dose imatinib cohorts (Fisher’s exact test, p-value = 0.55 for three-way comparison, p-values for pairwise comparison were 1, 0.37, and 0.41, respectively (Supplementary Table S3). The differences in the trend were significant for pairwise comparison of the low-dose imatinib cohort with dasatinib, or nilotinib or high-dose imatinib cohorts (Fisher’s exact test, p=5x10^{-4}, 9x10^{-4}, 5x10^{-5} respectively).

For patients with the 2-phasic model as better fitting in the dasatinib, nilotinib and 800mg imatinib cohorts, there was no significant difference in the number of positive/negative/zero second slopes (Fisher’s exact test, p=0.85 for three-way comparison, p-values for pairwise comparison were 0.79, 0.58, 0.94 respectively (Supplementary Table S3). The differences were significant when performing pairwise comparisons between the 400mg imatinib and dasatinib cohorts, or the nilotinib or 800mg imatinib cohorts (Fisher’s exact test, p=0.025, 0.028, 0.009 respectively).
There were 35 patients in the dasatinib cohort, 26 in the nilotinib cohort, 54 in the high-dose imatinib cohort and 13 patients in the low-dose imatinib cohort or whom the 2-phasic model was the better fitting model and for whom the first slopes were negative (Supplementary Table S3). There was no significant difference in the first slope among the four cohorts when applying the Wilcoxon test in a pairwise manner. Regarding the second slopes, the number of 2-phasic patients with both negative first and second slopes was 18 for the dasatinib cohort, 13 for the nilotinib cohort, 24 for the high-dose imatinib cohort and 12 for the 400mg imatinib cohort (Supplementary Table S3). For the three cohorts (dasatinib, nilotinib and 800mg imatinib) from MD Anderson, there was no significant difference in second slopes among them. However, there was a significant difference in second slopes between the 400mg imatinib cohort and all other three cohorts. There was a significant difference in turning points only between the nilotinib cohort and the 800mg imatinib cohort.

We then incorporated the estimated first and second slopes into our mathematical framework to investigate the dynamics of treatment responses of the four cohorts. We found that the framework accurately predicted the dynamics of treatment responses of all cohorts (Figure 4). These findings suggest that a model based on four distinct subpopulations of cells within the differentiation hierarchy can accurately explain the treatment responses not only to first-line low-dose imatinib, but also to second-line treatment as well as therapy with second-generation TKIs. This model will be useful for analyzing long-term treatment responses as more data becomes available.

Discussion

We have presented a quantitative approach to model the time course of BCR-ABL1 transcripts in patients with CML. We first designed a statistical model using minimal assumptions to provide maximal generalizability and information retrieval from the data. This model provided information on the shape of the treatment response curves as well as differences between and among patient cohorts. We then utilized a mathematical model emphasizing the biological interpretability of the parameters. This model considers four distinct cell populations within the differentiation hierarchy of the leukemia and can be used to predict the treatment response for different patient cohorts.
Our approach demonstrated that dasatinib, nilotinib and high-dose imatinib elicit very similar treatment responses in front-line patients. These patterns include decreases in $BCR-ABL_1$ transcripts in two phases during short-term treatment. The main result from the statistical model was a difference between dasatinib and nilotinib only when including front-line and second-line patients. When comparing the 400mg imatinib cohort from the IRIS trial with the three cohorts from the MD Anderson Cancer Center in a pairwise manner, there were significant differences in phase trends and second slopes.

However, there were a number of limitations to our study. The set of data from the IRIS trial was much smaller than those obtained from the MD Anderson patients and the data were generated at different geographical sites. Furthermore, with the exception of the nilotinib and dasatinib datasets, which were derived from two parallel phase II clinical trials, the data were not contemporaneous with each other nor were they obtained from a randomized study. In addition, the low-dose imatinib cohort contained a small number of patients who tolerated the treatment for up to 10 years, which might lead to the analysis of a very selected subgroup of patients. For the MD Anderson patients, there was not a significant difference among treatment groups by age ($p = 0.53$; Kruskal-Wallis test) or by percentage in the chronic phase ($p = 0.09$; Fisher’s exact test). Another limitation of the analysis was the limited number of data points (between 5 and 7) for each individual subject. A final limitation is that multiple analyses were performed. The nature of these analyses (some of them are multiple treatment comparisons and some are subset analyses) makes it difficult to control the family-wise error rate. However, even in randomized trials it is rare to see an adjustment for subgroup analyses, so we believe that our approach is valid.

These caveats notwithstanding, and in concert with the results herein reported, therapy with the second generation TKIs nilotinib and dasatinib have been shown in randomized phase III studies (i.e. the ENESTnd and DASISION studies) to produce remarkably better response rates in the front-line setting compared to standard dose imatinib$^{10-12}$. However, results from the TOPS studying, a phase III randomized study comparing the efficacy of high-dose imatinib and standard dose imatinib, resulted in similar outcomes in both arms of the study$^{21}$. A potential explanation for this discrepancy is the fact that in the TOPS study, a
higher proportion of patients receiving high-dose imatinib had to discontinue therapy with such an approach due to toxicity, whereas in the ENESTnd and DASISION studies, the toxicity profiles of nilotinib and dasatinib were similar to (if not better than) that of imatinib and the drop-out rates from the study favored the use of the second generation TKIs\textsuperscript{10-12}. It is also worth emphasizing that in the TOPS study, patients treated with high-dose imatinib achieved CCyR and MMR considerably earlier than those receiving standard dose imatinib (6-month CCyR 57% vs 45% with standard dose; p=0.01) and those receiving high-dose imatinib capable of maintaining adequate dose intensity had improved outcomes. A recent study has reported higher MMR rates at 12 months amongst patients receiving high-dose imatinib compared to those receiving standard dose imatinib (59% vs 44%, p<0.001) and improved rates of complete molecular response by 3 years (57% vs 46%, respectively)\textsuperscript{22}. These results are important given the favorable prognostic impact of achieving early deep molecular responses shown by different independent studies\textsuperscript{23-25}. In contrast to these findings, the dynamics of $BCR$-$ABL1$ clearance observed in our study suggest that the use of standard or high-dose imatinib or a second generation TKI such as nilotinib or dasatinib elicit similar responses when administered as frontline therapy for patients with CML in chronic phase. The discrepancy between our results and previous findings might stem from reasons of patient selection, an absence of randomization in the study protocol leading to biases in the M.D. Anderson cohorts, and the use of different statistical methods designed to analyze the shape of the treatment response curves in detail. In particular, the standard dose imatinib cohort was part of the Australasian arm of the IRIS trial and represents a very selected subgroup of patients – those that tolerated imatinib for a prolonged period of time (up to 10 years) without developing resistance or progression of disease. They might thus be examples of the best possible response kinetics that standard dose imatinib can elicit. The blood samples of these patients were also not analyzed at M.D. Anderson like the remainder of the patients included in our study. None-the-less, our results do not provide evidence against using second-generation TKIs as frontline treatment for chronic phase patients.

**Acknowledgements**

We would like to thank the Michor lab for helpful discussion, and also especially thank all the participating patients, their families, the research teams and the clinical investigators in the phase II trial at MD Anderson Cancer Center and the IRIS trial study.
Authorship and Disclosures

Study design: A.O., M.T., M.G., A. Q.-C., F. M.; Analyses: A.O., M.T., M.G; Preparation of manuscript: A.O., M.T., M.G., A. Q.-C., F. M.; Data collection: J. C., T. H., S. B., A. Q.-C. The authors have nothing to disclose.

References


### Tables

**Table 1.** Comparison of intercepts $\alpha_0$ and slopes $\alpha_1$ in the statistical model for treatment groups, frontline versus second line, as well as CML stages.

<table>
<thead>
<tr>
<th></th>
<th>Number of patients</th>
<th>Intercept mean (SD)</th>
<th>Slope mean (SD)</th>
<th>P-value (comparison)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imatinib</td>
<td>123</td>
<td>0.14(0.42)</td>
<td>2.59(1.24)</td>
<td>0.08 (imatinib vs nilotinib)</td>
</tr>
<tr>
<td>Nilotinib</td>
<td>75</td>
<td>0.52(1.01)</td>
<td>2.34(1.67)</td>
<td>0.23 (imatinib vs dasatinib)</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>92</td>
<td>0.29(0.56)</td>
<td>2.51(1.44)</td>
<td>&lt;0.001 (dasatinib vs nilotinib)</td>
</tr>
<tr>
<td><strong>Previously treated versus naïve patients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontline</td>
<td>243</td>
<td>0.14(0.41)</td>
<td>2.7(1.31)</td>
<td>&lt;0.001 (frontline vs second line)</td>
</tr>
<tr>
<td>Second line</td>
<td>47</td>
<td>1.04(1.14)</td>
<td>1.44(1.53)</td>
<td></td>
</tr>
<tr>
<td><strong>Treatment type for frontline patients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imatinib</td>
<td>123</td>
<td>0.14(0.42)</td>
<td>2.59(1.24)</td>
<td>0.97 (imatinib vs nilotinib)</td>
</tr>
<tr>
<td>Nilotinib</td>
<td>51</td>
<td>0.13(0.38)</td>
<td>2.81(1.44)</td>
<td>0.43 (imatinib vs dasatinib)</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>69</td>
<td>0.14(0.42)</td>
<td>2.82(1.34)</td>
<td>0.60 (dasatinib vs nilotinib)</td>
</tr>
<tr>
<td><strong>Treatment type for second line patients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nilotinib</td>
<td>24</td>
<td>1.35(1.38)</td>
<td>1.34(1.70)</td>
<td></td>
</tr>
<tr>
<td>Dasatinib</td>
<td>23</td>
<td>0.73(0.70)</td>
<td>1.55(1.35)</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>CML stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chronic</td>
<td>260</td>
<td>0.23(0.59)</td>
<td>2.57(1.39)</td>
<td></td>
</tr>
<tr>
<td>Accelerated</td>
<td>30</td>
<td>0.74(1.12)</td>
<td>1.85(1.54)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table 2. Summary statistics of slopes and turning points in all four patient cohorts.

The table displays the first ($\beta_1$) and second ($\beta_2$) slopes as well as the turning points ($\tau$) for all patients in all four cohorts whose treatment response data displayed a 2-phasic trend and whose first and second slopes were negative. The unit for slopes is in BCR-ABL1/ABL1% per day and the unit for the turning point is in month.

### Dasatinib cohort: 35 / 69 patients had a bi-phasic trend

<table>
<thead>
<tr>
<th>Slopes and turning points</th>
<th>mean (standard error)</th>
<th>median (number of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First slope, $\beta_1$, for patients with $\beta_1 &lt; \text{zero}^*$</td>
<td>-0.044 (0.027)</td>
<td>-0.038 (35)</td>
</tr>
<tr>
<td>Second slope, $\beta_2$, for all patients with $\beta_1 &lt; \text{zero}^* \text{ and } \beta_2 &lt; \text{zero}^*$</td>
<td>-0.002 (0.002)</td>
<td>-0.001 (18)</td>
</tr>
<tr>
<td>Turning point, $\tau$, for all patients with $\beta_1 &lt; \text{zero}^* \text{ and } \beta_2 &lt; \text{zero}^*$</td>
<td>5.28 (2.01)</td>
<td>4.39 (18)</td>
</tr>
</tbody>
</table>

### Nilotinib cohort: 26 / 51 patients had a bi-phasic trend

<table>
<thead>
<tr>
<th>Slopes and turning points</th>
<th>mean (standard error)</th>
<th>median (number of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First slope, $\beta_1$, for patients with $\beta_1 &lt; \text{zero}^*$</td>
<td>-0.046 (0.016)</td>
<td>-0.041 (26)</td>
</tr>
<tr>
<td>Second slope, $\beta_2$, for all patients with $\beta_1 &lt; \text{zero}^* \text{ and } \beta_2 &lt; \text{zero}^*$</td>
<td>-0.002 (0.002)</td>
<td>-0.001 (13)</td>
</tr>
<tr>
<td>Turning point, $\tau$, for all patients with $\beta_1 &lt; \text{zero}^* \text{ and } \beta_2 &lt; \text{zero}^*$</td>
<td>4.29 (1.15)</td>
<td>3.90 (13)</td>
</tr>
</tbody>
</table>

### High-dose imatinib cohort: 54 / 123 patients had a bi-phasic trend

<table>
<thead>
<tr>
<th>Slopes and turning points</th>
<th>mean</th>
<th>median</th>
</tr>
</thead>
</table>
### Slopes and turning points

<table>
<thead>
<tr>
<th></th>
<th>mean (standard error)</th>
<th>median (number of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First slope, $\beta_1$, for patients with $\beta &lt; \text{zero}^*$</strong></td>
<td>-0.044 (0.018)</td>
<td>-0.040 (54)</td>
</tr>
<tr>
<td><strong>Second slope, $\beta_2$, for all patients with $\beta_1 &lt; \text{zero}^<em>$ and $\beta_2 &lt; \text{zero}^</em>$</strong></td>
<td>-0.002 (0.001)</td>
<td>-0.002 (24)</td>
</tr>
<tr>
<td><strong>Turning point, $\tau$, for all patients with $\beta_1 &lt; \text{zero}^<em>$ and $\beta_2 &lt; \text{zero}^</em>$</strong></td>
<td>5.68 (1.79)</td>
<td>5.12 (24)</td>
</tr>
</tbody>
</table>

---

**Standard dose imatinib cohort: 13 / 13 patients had a bi-phasic trend**

<table>
<thead>
<tr>
<th></th>
<th>mean (standard error)</th>
<th>median (number of patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First slope, $\beta_1$, for patients with $\beta &lt; \text{zero}^*$</strong></td>
<td>-0.037 (0.011)</td>
<td>-0.036 (13)</td>
</tr>
<tr>
<td><strong>Second slope, $\beta_2$, for all patients with $\beta_1 &lt; \text{zero}^<em>$ and $\beta_2 &lt; \text{zero}^</em>$</strong></td>
<td>-0.004 (0.002)</td>
<td>-0.004 (12)</td>
</tr>
<tr>
<td><strong>Turning point, $\tau$, for all patients with $\beta_1 &lt; \text{zero}^<em>$ and $\beta_2 &lt; \text{zero}^</em>$</strong></td>
<td>5.07 (1.46)</td>
<td>4.85 (12)</td>
</tr>
</tbody>
</table>

*zero refers to any number with an absolute value $\leq 10^{-8}$. 
Figure legends

**Figure 1. The high-dose imatinib cohort.** The figure shows four representative patients from the high-dose imatinib cohort, displaying the ratio of $BCR-ABL1$ to $ABL1$ plus one over time measured in months (broken line) as well as the exponential model (solid line).

**Figure 2. The nilotinib cohort.** The figure shows four representative patients from the nilotinib cohort, displaying the ratio of $BCR-ABL1$ to $ABL1$ plus one over time measured in months (broken line) as well as the exponential model (solid line).

**Figure 3. The dasatinib cohort.** The figure shows four representative patients from the dasatinib cohort, displaying the ratio of $BCR-ABL1$ to $ABL1$ plus one over time measured in months (broken line) as well as the exponential model (solid line).

**Figure 4. A mathematical framework accurately predicts the dynamics of treatment responses of patient cohorts treated with dasatinib, nilotinib and high-dose imatinib separately.** The panels display the median (orange circles) and quartiles of the dasatinib, nilotinib and high-dose imatinib response data together with the results of the mathematical framework (blue curves, see Supplementary Table S1). (a) Median plots and results of the mathematical framework for the dasatinib response data. (b) Median plots and results of the mathematical framework for the nilotinib response data. (c) Median plots and results of the mathematical framework for the high-dose imatinib response data. Based on the model presented in the SI, the mathematical model prediction is given by $y_3/(2x_3 + y_3)$. Here $x_3$ and $y_3$ denote the abundance of normal and leukemic terminally differentiated cells. Parameter values are $d_0 = 0.0005$, $d_2 = 1$, $r_x = 0.008$, $r_y = 0.01$, $p_x = 1.5 \times 10^{-5}$, $p_y = 1.9 \times 10^{-6}$, $a_x = 0.35$, $b_x = 5.5$, $c_x = 100$, $a_y = 2a_x$, $b_y = 1.5b_x$, $c_y = c_x$, $c'_y = c_y$, $r'_y = r_y/15$. For dasatinib cohort, $d_1 = 0.0053$, $d_2 = 0.0394$, $a'_y = a_y/200$, $b'_y = b_y/300$; for nilotinib cohort $d_1 = 0.0028$, $d_2 = 0.0442$, $a'_y = a_y/400$, $b'_y = b_y/600$; for the high-dose imatinib cohort, $d_1 = 0.0035$, $d_2 = 0.055$, $a'_y = a_y/400$, $b'_y = b_y/600$. Apart from the dimension-less parameters, all values are given in units per day. Note that these parameter choices represent only one example that can recapitulate the dynamics of the treatment response seen in the clinic; other choices are possible.
Fig. 3
SUPPLEMENTARY INFORMATION

Dynamics of chronic myeloid leukemia response to dasatinib, nilotinib, and high-dose imatinib

Adam Olshen1*, Min Tang2*, Jorge Cortes3, Mithat Gonen4, Timothy Hughes5, Susan Branford5, Alfonso-Quintás-Cardama3, and Franziska Michor2**

1University of California at San Francisco, San Francisco, CA, USA. 2Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, and Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA. 3M.D. Anderson Cancer Center, Houston, TX, USA. 4Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. 5Department of Genetics and Molecular Pathology, SA Pathology, and University of Adelaide, Adelaide, Australia. *These authors contributed equally to this work.

Table of contents

STUDY GROUP .......................................................... page 2
Patient selection .............................................................. page 2
Patient cohorts .............................................................. page 2
DETAILS OF MATHEMATICAL MODELING APPROACH ......................... page 2
Individual model fitting ..................................................... page 2
Whole cohort model fitting ............................................... page 3
Mathematical model ....................................................... page 4
SUPPLEMENTARY REFERENCES ........................................... page 5
SUPPLEMENTARY TABLES ................................................ page 5
SUPPLEMENTARY FIGURE LEGENDS........................................ page 7
STUDY GROUP

Patient selection
We utilized four cohorts in this analysis: 233 patients treated with dasatinib, 215 patients treated with nilotinib and 281 patients treated with high-dose imatinib within the phase II trial at the MD Anderson Cancer Center\(^{1-2}\), as well as 22 patients treated with low-dose imatinib (400mg daily) within the IRIS trial\(^{3-4}\). We employed the following criterion to select patients from the study group for inclusion in our various analyses. Only patients with measurements at the first five time points were included, which were at 0, 3, 6, 9, and 12 months.

Patient cohorts
Based on the above patient selection criteria, we obtained data of 92 patients treated with dasatinib, 75 patients treated with nilotinib and 123 patients treated with high-dose imatinib within the phase II trial at the MD Anderson Cancer Center\(^{1-2}\). We also obtained data of 13 patients treated with low-dose imatinib (400mg daily) within the IRIS trial\(^{3-4}\). For each patient, the disease burden was quantitated using the International Scale standardized methodology over the entire course of the trial.

DETAILS OF MATHEMATICAL MODELING APPROACH
To investigate the parameters of the mathematical model, we first conducted statistical model fitting to data of each patient cohort.

Individual model fitting
We fit two models to each individual patient to identify the better fit: a single exponential curve and a bi-phasic exponential curve with a turning point. These models were chosen since the leukemic cell burden is expected to decrease (or increase) at an exponential rate. When fitting these two models, we first performed a logarithmic transformation of the original data and then fit a bi-phasic linear or linear model to the transformed data.

The log-transformed data is of the form \((t_{ij}, y_{ij})\), \(i = 1, \ldots, N; j = 1, \ldots, n_i\), where \(i\) is the patient-specific index and \(n_i\) is the total number of BCR-ABL1% value measurements for patient \(i\); here \(y_{ij}\) is logarithmically transformed. Then the linear model for each patient \(i\) is given by

\[
M_0: E(y_{ij}) = \alpha_i + \beta_i t_{ij}, i = 1, \ldots, N; j = 1, \ldots, n_i,
\]

where \(\alpha_i\) and \(\beta_i\) are parameters to be estimated for each individual based on data of each individual patient \(i\); while the bi-phasic linear model is given by

\[
M_1: E(y_{ij}) = (\alpha_i + \beta_i t_{ij} I_{(t_{ij} \leq \tau_i)}) + (\alpha_i' + \beta_i' t_{ij} I_{(t_{ij} > \tau_i)}), \quad \alpha_i + \beta_i \tau_i = \alpha_i' + \beta_i' \tau_i.
\]

where \(\alpha_i', \beta_i', \alpha_i, \beta_i\) and \(\tau_i\) are parameters to be estimated for each individual based on data of each individual patient \(i\).

To determine the model with the best fit for each individual patient's data, choosing among the above linear versus bi-phasic linear models, we utilized the joinpoint software, which is publicly available on the NCI website, [http://surveillance.cancer.gov/joinpoint/](http://surveillance.cancer.gov/joinpoint/). We chose to use this...
software because it was designed to identify segmented models that best fit to longitudinal data. Within this framework, we utilized the permutation test approach since it controls the error probability of selecting the wrong model at a given significance level, for instance 0.05. This option was chosen over other approaches (such as the BIC) since the latter does not provide an estimate of the error probability.

For each model with $k$ turning points, we estimated a total of $2k + 2$ parameters ($k = 0$ or 1 in our study). For the parameter estimation, we utilized Hudson’s Method\(^5\), as it provides more accurate estimates compared to the Grid Search method\(^6\), even though it is computationally more expensive. For details of the parameter estimation and the hypothesis testing see Kim et al.\(^7\). In brief, for any linear model with $k$ turning points, Hudson’s method first partitions the observed data into $k+1$ consecutive subsets. For each subset, an ordinary least squares method is applied to obtain the intercept and slope estimates over the data range of that subset. The turning points are then directly calculated as the intersections of two adjacent linear segments. If these intersection points divide the observed data into the same partitions as chosen originally, then the fit is admissible and its sum of squared error (SSE) is noted. Otherwise, the fit is not admissible and further adjustments are made. The least squares estimates for the linear model with $k$ turning points are then obtained from the fit, which provides the smallest SSE over all feasible partitions.

After the parameter estimation outlined above, the following permutation test procedure is performed to determine the better fit between the linear model $M_0$ and the bi-phasic linear model $M_1$. This approach\(^5\) can be summarized in several steps:

1. Fit the original data set with the null hypothesis model with 0 turning points.
2. Permute the residuals from the null hypothesis model and add them back to the means from the null hypothesis model to obtain a new permutation data set.
3. For this permutation data set, fit both the null model with 0 turning points and the alternative model with 1 turning points and calculate a scalar goodness-of-fit measure. This measure is a ratio, $\text{SSEN}/\text{SSEA}$, where $\text{SSEN}$ is the sum of squared errors (SSE) from the null model $M_0$ and $\text{SSEA}$ is the SSE from the alternative model $M_1$.
4. Repeat steps 2 and 3 $N_p - 1$ times. Denote the ratios, $\text{SSEN}/\text{SSEA}$, from the permutation data sets $p$ as $\{T_p, p = 1, \ldots, N_p - 1\}$. Also calculate this ratio for the original data set and denote it as $T_0$. Values of the ratio close to 1 represent the case in which the alternative is not much better than the null hypothesis model, while larger ratios signify that the alternative is much better.
5. The p-value of testing the hypothesis $M_0$ versus $M_1$ for the original data set is determined from the permutation distribution of the goodness-of-fit statistics. Then p-value $= (\text{number of times that } [T_p \geq T_0] \text{ for } p = 1, \ldots, N_p - 1) / N_p$.

The permutation tests are used to investigate whether there is enough evidence to require a model with a larger number of turning points than the one in the null hypothesis. This approach controls the error probability of selecting the wrong model at a significance level of 0.05.

For the dasatinib treatment response data, 35 out of 69 patients had the bi-phasic model as better fitting. For the nilotinib treatment response data, 26 out of 51 patients had the bi-phasic model as better fitting. For the high-dose imatinib treatment response data, 54 out of 123 patients had the bi-phasic model as better fitting. For the low-dose imatinib treatment response data, all 13 patients had the bi-phasic model as better fitting (Supplementary Table S3).

**Whole cohort model fitting**
We then identified the model with the best fit to the entire patient cohort. When fitting each model to data of each patient individually as stated in the above section, we obtained the corresponding SSE, as well as SST \( i \) (Total Sum of Squares) for each subject \( i \). The total SSE and total SST of the model were then calculated as \( \sum_i SSE_i \) and \( \sum_i SST_i \) separately. We defined the final \( R^2 \) for each model over the whole cohort as \( 1 - \text{total SSE/total SST} = 1 - \sum_i SSE_i / \sum_i SST_i \). Let \( R_0^2 \) be the final \( R^2 \) of the model \( M_0 \) and \( R_1^2 \) be the final \( R^2 \) of the model \( M_1 \).

For the dasatinib treatment response data, 35 out of 69 patients had bi-phasic model as better fitting, \( R_0^2 \) was 0.59 , and \( R_1^2 \) was 0.92. For the nilotinib treatment response data, 26 out of 51 patients had bi-phasic model as better fitting, \( R_0^2 \) was 0.60, and \( R_1^2 \) was 0.93. For the high-dose imatinib treatment response data, 54 out of 123 patients had the bi-phasic model as better fitting, \( R_0^2 \) was 0.52, and \( R_1^2 \) was 0.89. For the low-dose imatinib treatment response data, all 13 patients had the bi-phasic model as better fitting. Also, for the dasatinib, nilotinib and high-dose imatinib cohorts, we obtained \( N \) subject-specific \( R^2 = 1 - \text{SSE}_i / \text{SST}_i \), \( i = 1, ..., N \). The summary information (Minimum, 1st Quartile, Median, Mean, 3rd Quartile, Maximum) of the \( R^2 \) for each model of different analysis is reported in Supplementary Table S2.

We found that the bi-phasic exponential model provided a larger final \( R^2 \) and smaller sum of BICs than the exponential model for all cohorts. The difference between the final \( R^2 \) was large enough to convince us of the better fit of the bi-phasic model in all four cohorts.

**Mathematical model**

We utilized a mathematical model of the treatment response of CML cells to TKI therapy \( 8-9, \) which describes four layers of the differentiation hierarchy of the hematopoietic system. Stem cells give rise to progenitors, which produce differentiated cells, which in turn produce terminally differentiated cells. This hierarchy applies both to normal and leukemic cells. Only stem cells have the potential for indefinite self-renewal, but progenitor and differentiated cells possess the capability to undergo limited reproduction, which, together with differentiation, leads to an expansion of the cell number at each level of the differentiation hierarchy. The BCR-ABL oncogene is present in all leukemic cells, leading to slow clonal growth of leukemic stem cells and accelerating the rate at which leukemic progenitors and differentiated cells are generated. Imatinib therapy reduces the production rates of leukemic progenitors and differentiated cells, and potentially also inhibits the expansion of leukemic stem cells.

Denote by \( x_0, x_1, x_2, \) and \( x_3 \) the abundances of normal hematopoietic stem cells, progenitors, differentiated cells, and terminally differentiated cells. Their respective leukemic abundances are given by \( y_0, y_1, y_2, \) and \( y_3 \). We assume that homeostatic mechanisms maintain the hematopoietic stem cell population at a constant level, and therefore introduce a density dependence term, \( \phi \), in the stem cell production rate. Leukemic stem cells grow at a slow pace until reaching their maximum number, which may be larger than that of normal stem cells; afterwards, their number is also held constant by a density dependence mechanism. Then the system containing stem cells (SC), progenitor cells (PC), differentiated (DC) and terminally differentiated cells (TC) is described by

\[
\begin{align*}
\text{healthy cells} & \quad \text{leukemic cells}
\end{align*}
\]
Here density dependence in the stem cell population is given by $\phi = 1/[1 + p_\phi (x_0 + y_0)]$ and $\varphi = 1/[1 + p_\varphi (x_0 + y_0)]$. The potentially different carrying capacities of normal and leukemic stem cells are represented by the parameters $p_\phi$ and $p_\varphi$. Imatinib dramatically reduces the differentiation rates of cells, $a_y$ to $a'_y$ and $b_y$ to $b'_y$. This change in rates leads to a bi-phasic decline of the leukemic cell burden. The parameters during imatinib therapy are denoted by $r'_y$, $a'_y$, $b'_y$ etc.

SUPPLEMENTARY REFERENCES


SUPPLEMENTARY TABLES

**Supplementary Table S1.** The basic model of the differentiation hierarchy of normal and leukemic cells. The abundances of normal stem cells, progenitors, differentiated, and terminally differentiated cells are given by $x_0$, $x_1$, $x_2$, and $x_3$, while the respective abundances of leukemic cells are given by $z_0$, $z_1$, $z_2$, and $z_3$. Normal and leukemic stem cells divide at rates $r_0$ and $r'_0$ per day, respectively. The rate constants for the production of progenitors, differentiated cells and terminally differentiated cells are given by $a_0$, $b_0$ and $c_0$ for normal and by $a'_0$, $b'_0$, and $c'_0$ for leukemic cells. Stem cells die at rate $d_0$, progenitors at rate $d_1$, differentiated cells at rate $d_2$, and terminally differentiated cells at rate $d_3$ per day. Cells at all levels are assumed to potentially reproduce symmetrically and/or asymmetrically; the limited replication potential of more
differentiated cell types is then considered as part of the differentiation rates. Density dependence is achieved by the functions $q_z = 1/[1 + \rho_z(x_0 + z_0)]$ and $q_z = 1/[1 + \rho_z(x_0 + z_0)]$; these functions take into account crowding, limited resources, and interactions with the microenvironment. We assumed that the BCR-ABL1 oncogene increases the rate at which progenitors and differentiated cells are being produced; $a_z > a_x$ and $b_z > b_x$. Molecurally targeted therapy counteracts the effects of BCR-ABL1 by reducing the differentiation rates to $a'z < a_x$ and $b'z < b_x$ and possibly reducing the growth rate of leukemic stem cells to $r'z < r_x$.

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>Leukemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stem cells</td>
<td>$\dot{x}_0 = [r_0 q_z - d_0]x_0$</td>
<td>$\dot{z}_0 = [r_z q_z - d_0]z_0$</td>
</tr>
<tr>
<td>Progenitors</td>
<td>$\dot{x}_1 = a_x x_0 - d_x x_1$</td>
<td>$\dot{z}_1 = a_z z_0 - d_z z_1$</td>
</tr>
<tr>
<td>Differentiated cells</td>
<td>$\dot{x}_2 = b_z x_1 - d_z x_2$</td>
<td>$\dot{z}_2 = b_z z_1 - d_z z_2$</td>
</tr>
<tr>
<td>Terminally differentiated cells</td>
<td>$\dot{x}_3 = c_z x_2 - d_z x_3$</td>
<td>$\dot{z}_3 = c_z z_2 - d_z z_3$</td>
</tr>
</tbody>
</table>

Supplementary Table S2. Summary statistics of the two statistical models for data in the dasatinib, nilotinib and high-dose imatinib patient cohorts. The two statistical models investigated were a single-phase exponential model (denoted as 1-phase in the following table) and a 2-phase exponential model (denoted as 2-phase in the following table). Note that summary statistics for the low-dose imatinib cohort were not presented here because all 13 patients in the low-dose imatinib cohort had 2-phase model as the best fitting model, thus 2-phase model was obviously the best fitting model for this entire cohort.

<table>
<thead>
<tr>
<th></th>
<th>Dasatinib cohort (69 patients)</th>
<th>Nilotinib cohort (51 patients)</th>
<th>High-dose imatinib cohort (123 patients)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2-phase</td>
<td>1-phase</td>
<td>2-phase</td>
</tr>
<tr>
<td>Min* of $R_i^2$</td>
<td>0.230</td>
<td>0.011</td>
<td>0.575</td>
</tr>
<tr>
<td>1st Quartile* of $R_i^2$</td>
<td>0.901</td>
<td>0.449</td>
<td>0.913</td>
</tr>
<tr>
<td>Median* of $R_i^2$</td>
<td>0.960</td>
<td>0.573</td>
<td>0.964</td>
</tr>
<tr>
<td>Mean* of $R_i^2$</td>
<td>0.920</td>
<td>0.569</td>
<td>0.922</td>
</tr>
<tr>
<td>3rd Quartile* of $R_i^2$</td>
<td>0.984</td>
<td>0.736</td>
<td>0.982</td>
</tr>
<tr>
<td>Max* of $R_i^2$</td>
<td>1.000</td>
<td>0.968</td>
<td>0.999</td>
</tr>
<tr>
<td>**Final $R_i^2$</td>
<td>0.924</td>
<td>0.592</td>
<td>0.934</td>
</tr>
<tr>
<td>***Sum of BICs</td>
<td>-26.1****</td>
<td>96.7****</td>
<td>-9.86</td>
</tr>
</tbody>
</table>

* the Minimum/1st Quartile/Median/mean/3rd Quartile/maximum of the $R_i^2$, $i = 1, ..., N$, calculated from the corresponding fitted model for each patient, where $N$ is the total number of patients and $R_i^2 = 1 - \text{SSE}_i / \text{SST}_i$;

** Final $R_i^2$, calculated as $1 - \sum \text{SSE}_i / \sum \text{SST}_i$, evaluates the overall fit of the corresponding model to the whole time series data with all patients;

*** Sum of BICs is the sum of BICs over all subjects for each model.
**** One patient in this cohort had SSE being exactly zero when fitting the 2-phasic exponential model which resulted in negative infinite BIC. The Sum of BICs here did not include the BIC from this patient.

**Supplementary Table S3.** Summary of statistical analysis of 1-phasic versus 2-phasic model comparison in all four patient cohorts. Note that all 2-phasic patients in each cohort had negative first slopes.

<table>
<thead>
<tr>
<th></th>
<th>Total # of patients</th>
<th># of 1-phasic</th>
<th># of 2-phasic</th>
<th>2-phasic, beta2 &gt;0*</th>
<th>2-phasic, beta2 &lt;0*</th>
<th>2-phasic, beta2 =0*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dasatinib cohort</td>
<td>69</td>
<td>34</td>
<td>35</td>
<td>15</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Nilotinib cohort</td>
<td>51</td>
<td>25</td>
<td>26</td>
<td>10</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>High-dose imatinib cohort</td>
<td>123</td>
<td>69</td>
<td>54</td>
<td>23</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>Standard dose imatinib cohort</td>
<td>13</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

* zero refers to any number with an absolute value ≤ 10^-6.

**SUPPLEMENTARY FIGURE LEGENDS**

**Figure S1.** A bi-exponential and an exponential fit to the log-transformed nilotinib curves. The bi-exponential fit is the sum of two exponentials. The curves labeled “1” (black) and “2” (red) are the components of the bi-exponential and the curve labeled “3” (green) is the result. The curve labeled “4” (blue) is a single scaled exponential with an intercept (our model in Statistical Methods). Note its similarity to the bi-exponential. This is because curve “2” is virtually a straight line.

**Figure S2. Individual fitting for patients in the imatinib 800mg cohort.** The figure displays each individual’s BCR-ABL1/ABL1% data (black circle), together with the bi-phasic fitting curves (red lines). For those patients whose data exhibit a better fitting with a single-phasic model, a dashed green line is shown. The green circles represents value of 0.00001, which is the original zero value.

**Figure S3. Individual fitting for patients in the dasatinib cohort.** The figure displays each individual’s BCR-ABL1/ABL1% data (black circle), together with the bi-phasic fitting curves (red lines). For those patients whose data exhibit a better fitting with a single-phasic model, a dashed green line is shown. The green circles represents value of 0.00001, which is the original zero value.

**Figure S4. Individual fitting for patients in the nilotinib cohort.** The figure displays each individual’s BCR-ABL1/ABL1% data (black circle), together with the bi-phasic fitting curves (red lines). For those patients whose data exhibit a better fitting with a single-phasic model, a dashed green line is shown. The green circles represents value of 0.00001, which is the original zero value.

**Figure S5. Individual fitting for patients in the imatinib 400mg cohort.** The figure displays each individual’s BCR-ABL1/ABL1% data (black circle), together with the bi-phasic fitting curves (red lines). For those patients whose data exhibit a better fitting with a single-phasic model, a dashed green line is shown. The green circles represents value of 0.00001, which is the original zero value.
Dasatinib 2–phase 1

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 2

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 3

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 4

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 5

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 6

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 7

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 8

log10 of BCR–ABL1/BCR%

Months

Dasatinib 2–phase 9

log10 of BCR–ABL1/BCR%

Months