ARTICLE

Reduced tyrosine kinase inhibitor dose is predicted to be as effective as standard dose in chronic myeloid leukemia: a simulation study based on phase III trial data

Artur C. Fassoni,[1,2] Christoph Baldow,[2] Ingo Roeder[2,3*] and Ingmar Glauche[2*]

[1]Instituto de Matemática e Computação, Universidade Federal de Itajubá, Brazil; [2]Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Germany and [3]National Center for Tumor Diseases (NCT), Partner Site Dresden, Germany

**IR and IG contributed equally to this work.*

Haematologica 2018
Volume 103(11):1825-1834

ABSTRACT

Continuing tyrosine kinase inhibitor (TKI)-mediated targeting of the BCR-ABL1 oncoprotein is the standard therapy for chronic myeloid leukemia (CML) and allows for a sustained disease control in the majority of patients. While therapy cessation for patients appeared as a safe option for about half of those patients with optimal response, no systematic assessment of long-term TKI dose de-escalation has been made. We use a mathematical model to analyze and consistently describe biphasic treatment responses from TKI-treated patients from two independent clinical phase III trials. Scale estimates reveal that drug efficiency determines the initial response while the long-term behavior is limited by the rare activation of leukemic stem cells. We use this mathematical framework to investigate the influence of different dosing regimens on the treatment outcome. We provide strong evidence to suggest that TKI dose de-escalation (at least 50%) does not lead to a reduction of long-term treatment efficiency for most patients, who have already achieved sustained remission, and maintains the secondary decline of *BCR-ABL1* levels. We demonstrate that continuous *BCR-ABL1* monitoring provides patient-specific predictions of an optimal reduced dose without decreasing the anti-leukemic effect on residual leukemic stem cells. Our results are consistent with the interim results of the DESTINY trial and provide clinically testable predictions. Our results suggest that dose-halving should be considered as a long-term treatment option for CML patients with good response under continuing maintenance therapy with TKIs. We emphasize the clinical potential of this approach to reduce treatment-related side-effects and treatment costs.

Introduction

In tyrosine kinase inhibitors (TKI)-treated chronic myeloid leukemia (CML), the proportion of *BCR-ABL1* mRNA is used to monitor the individual treatment response.[1-3] Most patients show a typical bi-phasic response with a steep, initial decline (slope α), followed by a slower, secondary decline (slope β) of *BCR-ABL1* levels.[4-6] Whereas the initial decline is attributed to the eradication of proliferating leukemic cells (LC), the secondary decline has been suggested to result from a slower eradication of quiescent leukemic stem cells (LSCs).[5-7] About two-thirds of the patients achieve major molecular remission (MMR), i.e. a reduction of three logs from the baseline (MR3), while one-third of these even achieve deep molecular remission (DMR, i.e. MR4.5) within five years of treatment.[3,6,8]

Recently, TKI cessation and, thus, treatment-free remission has been established as an important therapeutic goal.[9,10] However, about 50% of the patients with good response experience a molecular relapse after stopping TKI, pointing towards persisting residual LCs that cannot be controlled by patient-specific immunological mechanisms. As these mechanisms underlying the currently unpredictable individual molecular relapse risk remain controversial, complementary approaches to

Correspondence:
ingmar.glauche@tu-dresden.de

Received: March 29, 2018.
Accepted: June 26, 2018.
Pre-published: June 28, 2018.

doi:10.3324/haematol.2018.194522

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/103/11/1825

minimize side-effects associated with continuous TKI therapy are required. While a number of studies evaluate TKI cessation, strategies to apply long-term dose reductions are currently under-appreciated, although the potential benefits of dose de-escalation are at hand: besides reducing treatment-related side-effects and increasing patients' quality of life, it also reduces the treatment-related costs.[11-14] The DESTINY trial (*clinicaltrials.gov identifier: 01804985*)[15,16] is an ongoing study addressing TKI dose de-escalation. However, with its primary end point (molecular relapse risk after TKI cessation preceded by a one year dose de-escalation) this trial focuses on stopping TKI, rather than on long-term outcomes under continuous but reduced TKI treatment.

Here, we describe a systematic, conceptual analysis of the impact of dose de-escalation on the long-term disease kinetics. Our simulation study relies on a mathematical description of TKI-treatment, which builds on a previously published CML model.[5,17] In contrast to the earlier approach, we use a simplified model which allows for a stringent analytical formulation of the disease dynamics without changing the overall qualitative system properties. The model parameters are estimated from available, patient-specific *BCR-ABL1* kinetics, determined within controlled clinical phase III trials [IRIS (*clinicaltrials.gov identifier: 00006343*18), CML IV (*clinicaltrials.gov identifier: 00055874*19)].

Our results support the rationale for TKI dose de-escalation in patients who have already reached sustained remission. We provide strong evidence that the long-term depletion of residual LSCs in remission phase is not affected by defined TKI-dose reduction. Furthermore, we propose a strategy to determine patient-specific optimal TKI doses, and predict that dose-halving is a safe treatment option for the majority of patients in sustained molecular remission. The suggested dose optimization can contribute to the prevention of severe side-effects (e.g. cardiovascular complications, pleural effusion) and to a reduction of overall treatment costs.

Methods

Patient data and parametrization

The presented results are based on a secondary analysis of previously published data from the IRIS (*clinicaltrials.gov identifier: 00006343*)[18] and CML IV (*clinicaltrials.gov identifier: 00055874*) trials.[19] In particular, we used 69 patients from the German imatinib (400 mg) arm of the IRIS trial and 280 patients from the 400 mg imatinib monotherapy arm of the CML-IV study, for which IS corrected *BCR-ABL1/ABL1* time courses were available at the time of our primary model analysis.[5,17] As described in the original publications,[20,21] both clinical trials were conducted in accordance with the Declaration of Helsinki and applicable regulatory requirements. The protocols were approved by the institutional review board or ethics committee of each participating center. All patients or guardians gave written informed consent before participation.

For each individual patient, the treatment response at time t ($L_{OBS}(t)$), measured in the form 100% × *BCR-ABL1/ABL1*, is further described according to a biphasic characteristic, i.e.

$$L_{OBS}(t) = A\, e^{\alpha t} + B\, e^{\beta t}. \qquad (E1)$$

For each patient, parameters A, B, α and β are determined using maximum likelihood estimation. For the model analysis, we selected patients with (i) sufficient time points for model fitting (>4) which (ii) do not show a long-term increase in *BCR-ABL1/ABL1* ratio (β<0) but are characterized by (iii) a biphasic decline, (α<β<0). We also excluded 2 additional patients with measured *BCR-ABL1* ratios of more than 500%, which indicates a pronounced non-linearity between *BCR-ABL1* abundance and tumor load, resulting in n=55 (IRIS cohort) and n=134 (CML-IV cohort) (*Online Supplementary Table S1*). Median follow up of this patient cohort is 4.3 years [IQR (2.8,6.3)]; 98.4% of the patients achieved *BCR-ABL1* ratio of less than 1%, while 91% achieved MR3 at least once. We also tested the robustness of our model results with respect to the reliability of high *BCR-ABL1* values (*Online Supplementary Table S2*).

For comparisons with the DESTINY trial (*clinicaltrials.gov identifier: 01804985*),[15,16] only patients treated with TKI for at least three years and with *BCR-ABL1* levels below MR3 for at least the last year of treatment were used. Therefore, we excluded from the study 53 patients treated for less than three years (excluding n=4 for IRIS; n=49 for CML-IV) and 14 patients with no MR3 in the entire last year of treatment (excluding n=4 for IRIS; n=10 for CML-IV). The time courses of the remaining 122 patients [(n=47 IRIS, median follow up 6.5 years [IQR(5.9;6.9)]; n=75 CML-IV, median follow up 4.6 years [IQR(3.9;6.1)] are available in *Online Supplementary Figure S1*. Following the DESTINY trial, these patients were further split into an MR4 and an MR3 cohort, depending on whether their *BCR-ABL1* levels in the last year were below MR4 or not. These selection and classification procedures were based on the individual bi-exponential fit $L_{OBS}(t)$ of each patient (also shown in *Online Supplementary Figure S1*).

Mathematical model

We apply a mechanistic model that describes TKI response as a dynamic process resulting from the interplay between tumor growth, activation/deactivation of LSCs, and cytotoxic TKI action on proliferating, but not on quiescent LSCs (Figure 1A). The model is a simplification of our previous computational CML model[5,17] and formally related to a model proposed by Komarova and Wodarz.[22] It considers three leukemic cell types: quiescent LSCs (X), proliferating LSCs (Y), and fully differentiated LCs (W). Mathematically, the model is described by the following set of differential equations:

$$\frac{dX}{dt} = -p_{XY}X + p_{YX}Y \qquad (E2)$$

$$\frac{dY}{dt} = p_{XY}X - p_{YX}Y + p_Y Y - e_{TKI}Y \qquad (E3)$$

$$\frac{dW}{dt} = p_W Y - r_W W \qquad (E4)$$

The activation of dormant LSCs and deactivation of proliferating LSCs are described by rate constants p_{XY} and p_{YX} respectively. LSCs proliferate with a rate constant p_Y. During therapy, a cytotoxic TKI effect acts on proliferating LSCs, described by the rate e_{TKI}>0. Differentiation of proliferating LSCs is quantified by p_W and the limited life-time of differentiated LCs is modeled by a mortality rate r_W. We define a net leukemia reduction q=e_{TKI}−p_Y with q>0 for effective treatment. For illustrating simulations we used parameter values corresponding to the median values of the selected patients (*Online Supplementary Text S1-S3*).

Reimplementation of the model in a stochastic version using a Gillespie algorithm ensured that there are no distinct differences resulting from small cell numbers (*Online Supplementary Text S4*).

In contrast to previous models,[5,17] competition between normal cells and LCs is described only implicitly, by assuming constant total cell numbers, T_Y, T_X, T_W in each cell compartment (Figure 1D, *Online Supplementary Text S1* and *Online Supplementary Figure S2* complementing Figure 1C on the level of absolute cell numbers). The actual tumor load, corresponding to *BCR-ABL1* levels, is modeled as the percentage of LCs with respect to the total cell number. Figure 1C demonstrates that the modeled *BCR-ABL1* levels of proliferating LSCs behave exactly like the *BCR-ABL1* levels in the peripheral blood (PB). Therefore, only the dynamics of proliferating LSCs will be considered.

Results

The long-term effect of TKI is limited by the rare activation of quiescent LSCs

We apply a simple mathematical model that describes the time course of TKI response in CML as a dynamic

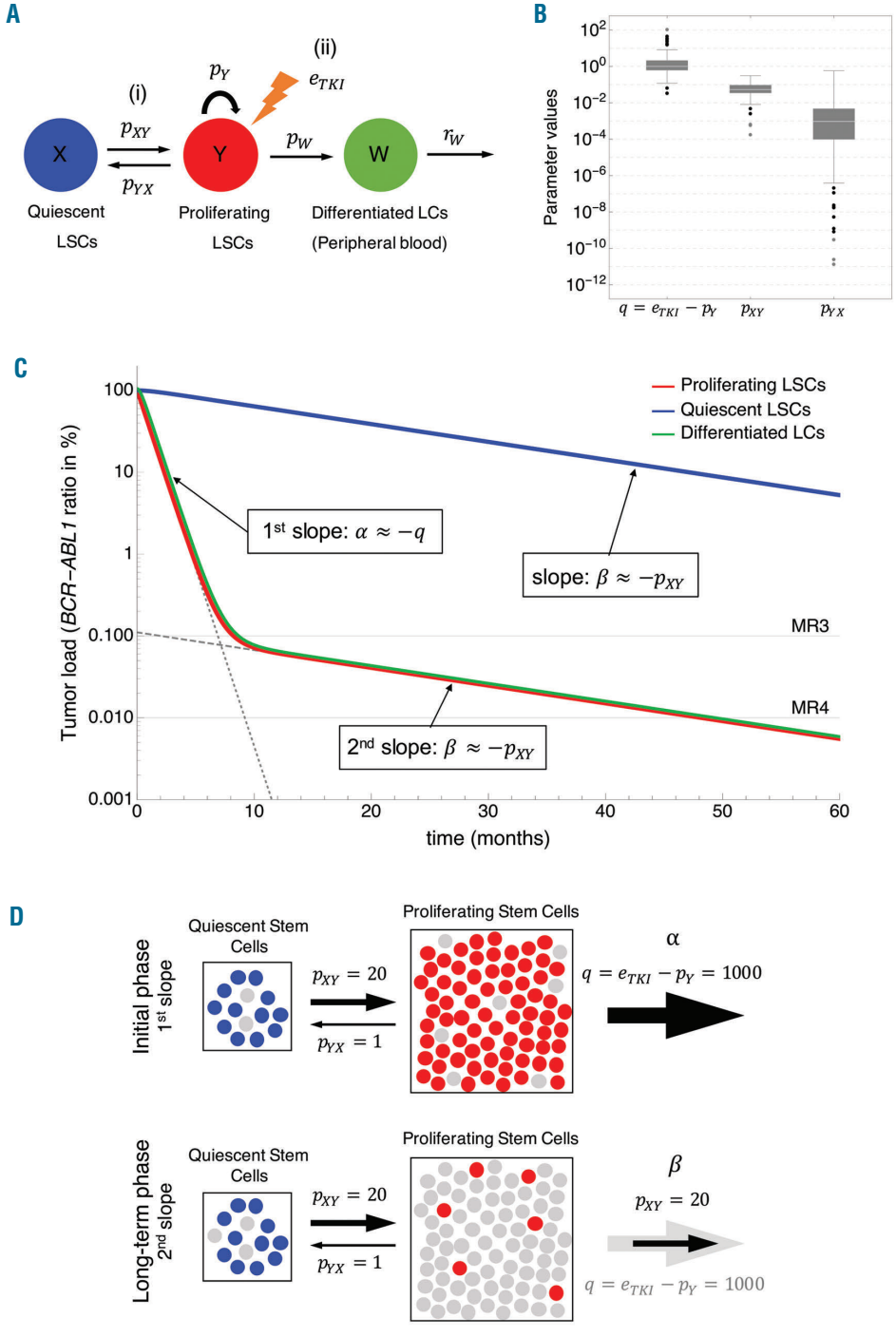

Figure 1. Mathematical model for chronic myeloid leukemia (CML) treatment and mechanistic interpretation of the bi-phasic decline. (A) Schematic model representation with three cell types: quiescent (X, blue) and proliferating (Y, red, turnover with rate p_Y) leukemic stem cells (LSCs), and differentiated leukemic cells (LCs), denoted by W (green, generated with rate p_W, decaying with rate r_W). The model assumes (i) mechanisms of activation/deactivation of quiescent/proliferating LSCs with rates p_{XY} and p_{YX} and (ii) a cytotoxic effect of TKI on proliferating LSCs with intensity e_{TKI}. (B) The mechanistic model parameters [(TKI net effect (q=e_{TKI}−p_Y), activation rate of quiescent LSCs (p_{XY}), deactivation rate of proliferating LSCs (p_{YX})] were fitted to individual patient data from the IRIS and CML-IV trials.[18,19] The resulting distributions reveal an intrinsic scaling between them, which are dispersed over different orders of magnitude. (C) Model simulation with median parameter values obtained from IRIS and CML-IV data illustrating the equivalence between tumor load (in terms of *BCR-ABL1* levels) in the peripheral blood (green) and within the proliferating LSCs (red). Values on the y-axis indicate the relative abundance of *BCR-ABL1* positive cells in each specific cell compartment [see equation (SE1) in *Online Supplementary Text S1*], which corresponds to the tumor load in terms of PCR-based measurements of the *BCR-ABL1/ABL1* ratio. We adopted this scheme for all corresponding figures throughout the manuscript. Using the intrinsic scaling (B), the slopes in the bi-exponential decline of the BCR-ABL1 levels simplify to $\alpha \approx -q$ and $\beta \approx -p_{XY}$. The abundance of quiescent LSCs follows a monophasic decline approximated by $\beta \approx -p_{XY}$. See *Online Supplementary Text S3* for parameter values used in all model simulations. (D) During the initial phase (upper panel, "1st slope"), eradication of the proliferating LSCs (red) with effective rate q is the dominating process (large black arrow). After the strong initial reduction, few proliferating cells remain (lower panel, "2nd slope") and eradication is now limited by the activation rate p_{XY} (small black arrow) of quiescent LSCs (blue). Normal cells are shown in gray. See also *Online Supplementary Figure S2*.

process resulting from the interplay between tumor growth, activation/deactivation of LSCs, and cytotoxic TKI action (Figure 1A, and see Methods). In brief, the model describes three LC types: quiescent LSC (X), proliferating LSC (Y) and fully differentiated LCs (W). The activation of dormant LSCs and deactivation of proliferating LSCs are described by rate constants p_{XY} and p_{YX} while LSCs proliferate with a rate constant p_Y. During therapy, we assume a cytotoxic TKI effect on proliferating LSCs, described by the rate constant $e_{TKI}>0$.

We obtain an exact solution for the model, in which the patient-specific response can be expressed in terms of the mechanistic model parameters [equation (SE5) in *Online Supplementary Text S2*]. In other words, a patient's bi-phasic *BCR-ABL1* decline characterized by the slopes α, β is expressed in terms of the resulting net cytotoxic effect $q=e_{TKI}-p_Y$ (difference between TKI toxicity and LSC proliferation) and the effects of LSC activation/deactivation p_{XY}

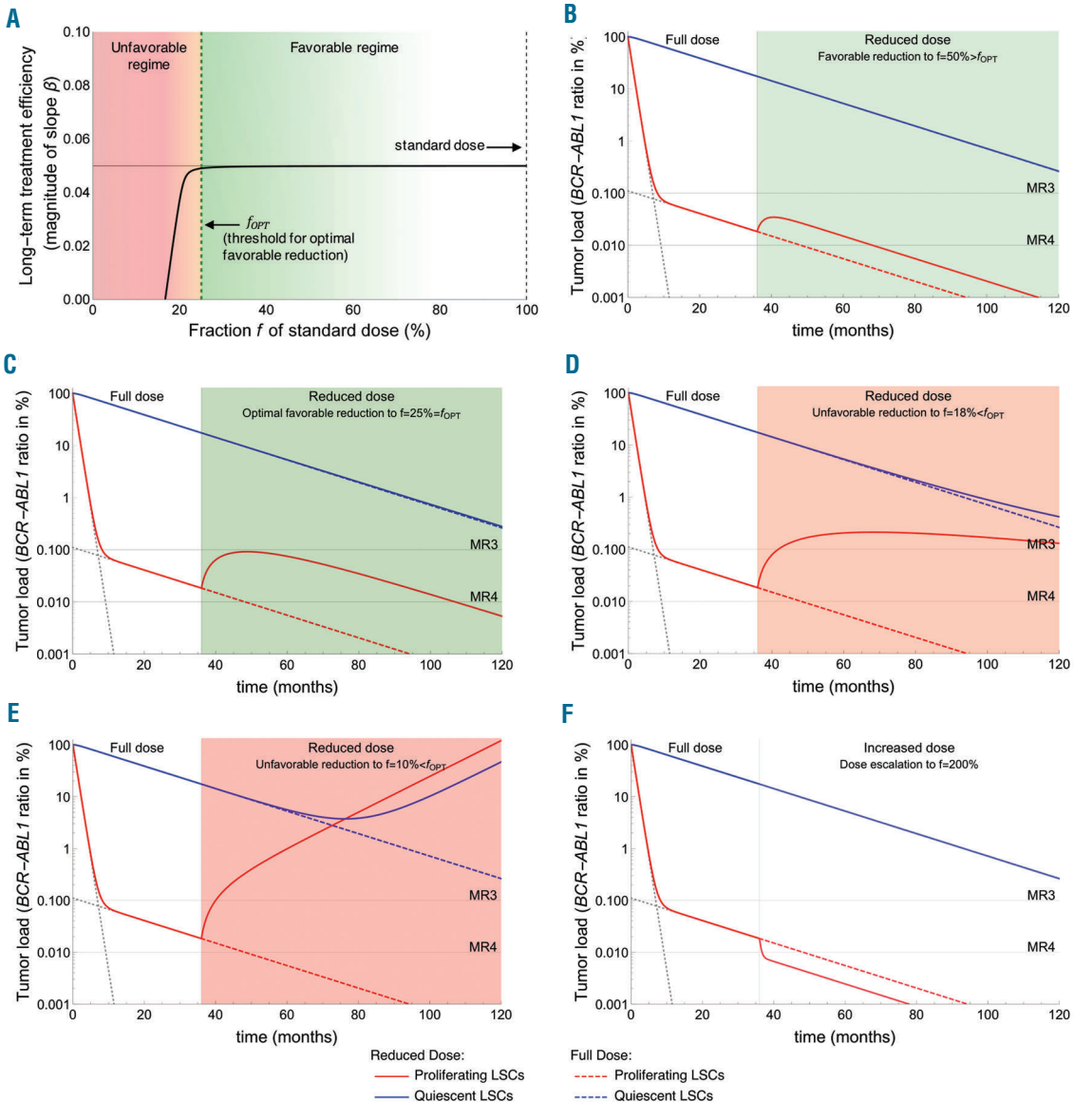

Figure 2. Model predictions on dose de-escalation and dose escalation. (A) The long-term treatment efficiency, defined as the magnitude of the second slope β, is shown as a function of the dose reduction. The threshold for optimal favorable reduction, f_{OPT} (≈25% in this example, i.e. using median parameters as in Figure 1C) indicates how much the standard dose can be reduced without losing treatment efficiency. f_{OPT} can be calculated for each patient (see main text). Any other favorable dose reductions (dose fraction $f>f_{OPT}$, green region) also retain the long-term treatment efficiency, while unfavorable dose reductions (dose fraction $f<f_{OPT}$, red region) are predicted to lead to a severe decrease in the long-term treatment efficiency. (B-E) Simulations of favorable (B), optimal favorable (C) and unfavorable (D and E) dose reductions after 36 months under standard dose. After favorable dose reductions, a transient increase in proliferating leukemic stem cells (LSCs) (red) is followed by a return to the original decrease rate $\beta\approx-p_{XY}$, while the dynamics of quiescent LSCs remains unchanged (blue lines). In the case of unfavorable reduction, an impaired scenario is observed. See also *Online Supplementary Figures S3-S6*. (F) dose escalation to f=200% after three years of treatment; although a deeper level is reached in the *BCR-ABL1* levels of proliferating LSCs, the dynamics of quiescent LSCs remains unchanged.

and p_{YX}. Taking into account the intrinsic scaling between the mechanistic parameters (Figure 1B), which indicates that LSC deactivation, activation, and depletion by TKI occur at different time scales, those expressions for α and β can be further simplified, allowing us to dissect the prominent processes governing each treatment phase (Figure 1C and D, *Online Supplementary Texts S5-S7* and *Online Supplementary Figure S2*).

We found that slope α can be expressed as $\alpha \approx -q = p_Y - e_{TKI}$, thereby confirming that the initial treatment phase is dominated by the cytotoxic TKI effect on proliferating LSCs, which leads to a rapid reduction in *BCR-ABL1* levels. Dose-escalation studies for imatinib substantiate this result by indicating that a higher TKI dose leads to a more rapid response.[23,24] Similarly, slope β can be approximated as $\beta \approx -p_{XY}$, implying that after depletion of initially abundant proliferating LSCs, the treatment response is bounded by the rare activation of quiescent LSCs. This provides a consistent explanation for the slower long-term decrease in proliferating and quiescent LSCs. These analytical conclusions confirm previous findings,[5-7] but also allow further predictions on the effect of dose de-escalation to be made.

A wide range of reduced TKI doses is predicted to induce the same long-term response as standard dose

Due to the bounded activation of quiescent LSCs after the initial therapy response, there is a range of favorable reduced TKI doses where the long-term efficiency (defined by the magnitude of slope β) remains almost constant with the same overall efficiency as that achieved when applying the standard dose (green region in Figure 2A, and *Online Supplementary Figure S3*). In this case, the resulting, although reduced, cytotoxic TKI effect is still sufficient to target the abundant proliferating LSCs once a patient has reached sustained remission. This range of 'favorable' reduced doses spans from the standard full-dose to a certain threshold, i.e. an 'optimal favorable' dose (green dashed line in Figure 2A), below which an accelerated decrease in long-term treatment efficiency is observed. Therefore, dose reductions below this optimal dose are considered as 'unfavorable' (red region in Figure 2A).

We look for a mathematical expression, which allows us to estimate this optimal dose reduction for each patient in terms of the model parameters. Although there is a minimal required plasma concentration for TKI cytotoxicity,

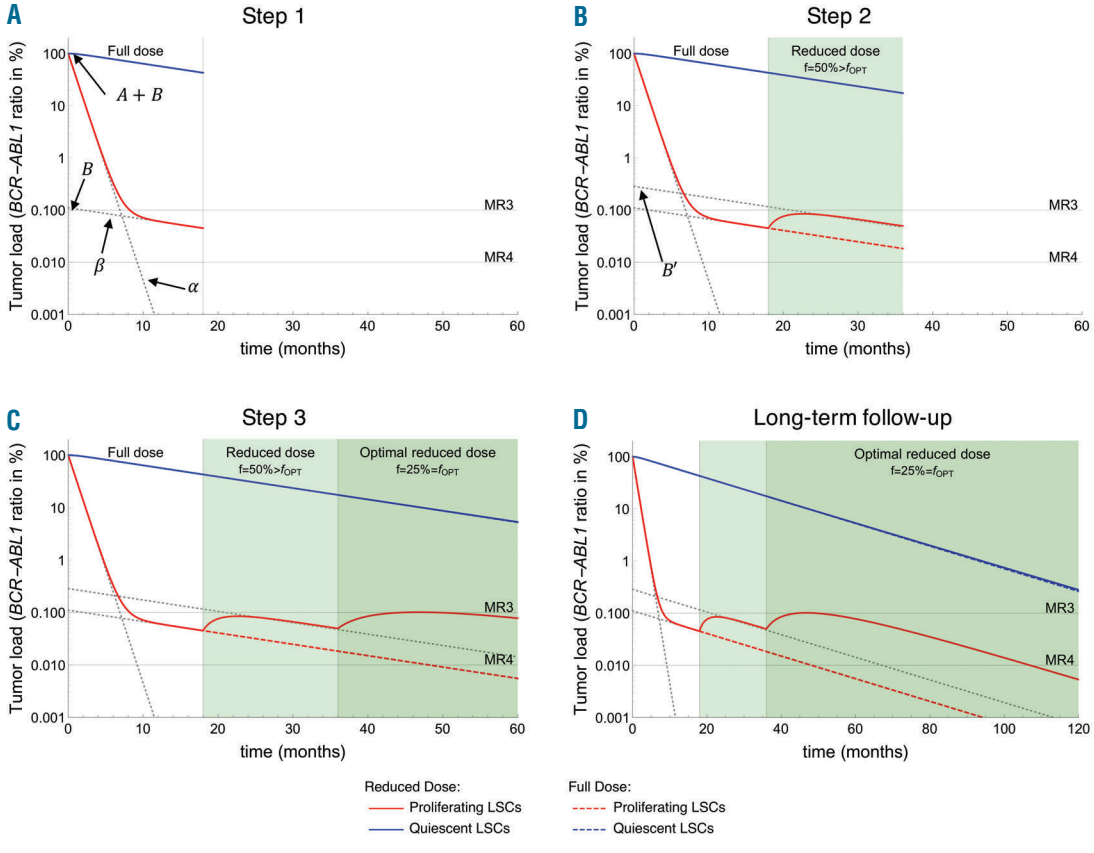

Figure 3. Step-wise treatment optimization. (A) Step 1: initial treatment with standard dose until the patient shows a clearly identifiable second slope (approx. 18 months) and determination of the bi-exponential parameters (*A, B, α, β*). (B) Step 2: reduction of tyrosine kinase inhibitor (TKI) dose by half and continuous monitoring of the treatment response until the new intercept *B′* can be inferred (approx. 18 months). (C) Step 3: reduction to the optimal dose calculated from values of the identified parameters (*A, B, B′, α, β*; see main text). (D) The long-term follow up using optimal dose shows that the response to this adaptive treatment adheres to the original slope β for the eradication of residual leukemic stem cells (LSCs) as standard dose treatment. Although the adapted treatment leads to a delay in the reduction of *BCR-ABL1* levels in proliferating LSCs and, therefore also in the peripheral blood, the treatment dynamics in the residual quiescent LSCs are unaltered while drug intake is drastically reduced.

which sets a lower limit to the dose de-escalation, a linear dose-response relationship is generally accepted over a wide range of treatment-relevant doses above this threshold.[25,26] Thus, we deduced an explicit expression for the patient-specific optimal favorable dose reduction fraction f_{OPT} (*Online Supplementary Text S8*), given by

$$f_{OPT} = \frac{p_Y + 2p_{XY}}{p_Y + q}. \qquad (E5)$$

This optimal fraction f_{OPT} corresponds to the minimal favorable dose which still maintains the original long-term reduction rate of both proliferating and quiescent LSCs. Therefore, as long as the TKI dose is not reduced below this threshold, our model predicts no impaired long-term efficiency, while an over-reduction compromises the overall treatment success.

We conclude from equation (E5) that f_{OPT} is a patient-specific, fixed quantity determined by the proliferation rate of LSCs, p_Y, their activation rate p_{XY}, as well as the toxicity of the TKI, $q=e_{TKI}-p_Y$. For the median parameters of the available dataset, the optimal favorable reduction fraction is f_{OPT}=0.247, and corresponds to a long-term treatment efficiency of 98.4% compared to the standard dose. Therefore, for the 'median patient' in our analysis, a reduction to 24.7% of the original dose would lead to a marginal decrease of only 1.6% in the long-term treatment efficiency, given that a minimal required plasma concentration for TKI cytotoxicity is maintained.

Our model predicts transient increases in *BCR-ABL1* levels of proliferating LSCs when applying different dose reductions after the first decline, i.e. once a substantial reduction in *BCR-ABL1* levels had been achieved (Figure 2). However, for favorable reductions, *BCR-ABL1* levels decrease again with the original long-term treatment efficiency (slope β) after a few months (Figure 2B and C). For the example of a 'median patient', dose reductions at month 36 of treatment will maintain MR3, while *BCR-ABL1* levels are predicted to return to their original values at de-escalation after about 20 months (in case of favorable reduction with f=0.5) or 58 months (for the optimal favorable reduction with f=0.25). Importantly, the transient increase of proliferating LSCs and, therefore, of *BCR-ABL1* levels in the PB, does not lead to either relevant differences in the overall response of quiescent LSCs or in the total

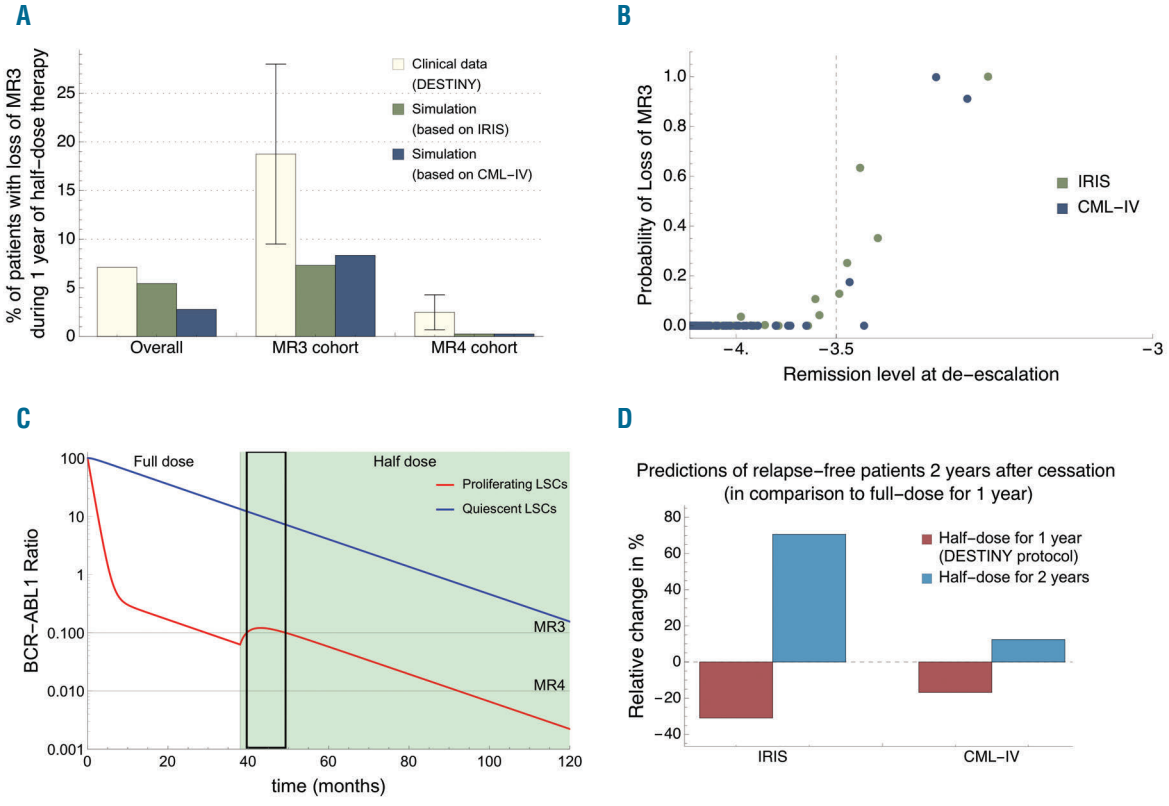

Figure 4. Model predictions on dose de-escalation and comparison with clinical data. (A) Comparison of DESTINY interim results with model simulations of 50% dose de-escalation applied to IRIS and CML-IV patient data (assuming the same protocol and patient selection criteria of DESTINY). We simulated dose de-escalation starting from the individually predicted remission level at the time of the last *BCR-ABL1* measurement of each patient, and evaluated the fraction of patients above MR3 one year after de-escalation. Error bars indicate 90% confidence intervals. (B) Model estimates of the risk of losing MR3 within one year after de-escalation, depending on the patient's individual predicted remission level just before de-escalation. Patients with remission level above MR3.5 are very likely to lose MR3 at least transiently. (C) Model simulation illustrating the transient relapse above MR3 three months after de-escalation (highlighted time interval). De-escalation of 50% was implemented for a hypothetical patient of the DESTINY trial one year after reaching MR3. The simulation of a continuing half-dose regimen predicts that after about nine months the *BCR-ABL1* levels fall below MR3 and the response regains the original slope β. (D) Simulation results showing the predicted relative increase/decrease in the number of patients without molecular relapse two years after cessation. We use the standard treatment scenario (full-dose for one year) as the reference (corresponding to the dashed line at 0%) to compare it with: i) half-dose for one year (the DESTINY protocol; red), and ii) half-dose for two years (blue). Relapse is defined as loss of MR3.

LSCs population when compared to standard dose (*Online Supplementary Figure S4*). Our model also predicts that returning to the full dose regimen at a later point completely restores the original response levels of proliferating LSCs within a few months (*Online Supplementary Figure S5*).

Patient-specific optimal dose can be estimated after an initial dose reduction

The calculation of patient-specific optimal TKI doses requires the knowledge of model parameters p_{YX}, p_{XY}, e_{TKI}, and p_Y. Whereas p_{YX}, p_{XY}, and $q=e_{TKI}-p_Y$ can be estimated from time course data, the LSC proliferation rate p_Y is confounded with the individual TKI-effect e_{TKI} and cannot be deduced directly. Therefore, we propose to estimate p_Y by observing the transient increase in proliferating LSCs occurring after a first favorable dose reduction. Technically, we suggest a moderate initial reduction to fraction f of the standard dose, after the patient's response under standard dose has been sufficiently quantified in terms of the kinetic parameters (α,β,A,B) (Figure 3A). Although dose-halving (i.e. f=0.5) seems to be safe for this first de-escalation (see below), the proposed approach is valid for any reasonable reduction. After a transient increase in the *BCR-ABL1* levels, a new intercept B' can be observed approximately 18 months after the dose reduction (Figure 3B). Based on the difference of intercepts B and B', and the reduction fraction f, the proliferation rate p_Y can be estimated (*Online Supplementary Text S9*) by

$$p_Y = |\alpha| \frac{B}{1-f}\left(\frac{f}{B}-\frac{1}{B'}\right) \qquad (E6)$$

and the individual optimal dose reduction fraction f_{OPT} can be calculated using equation (E5). This dose is predicted to retain the original long-term treatment efficiency (Figure 3C and D).

A population-based estimate predicts that the majority of patients in sustained remission retain the long-term treatment efficiency after dose-halving

As pointed out above, the LSC proliferation rate p_Y, as an indicator of the 'aggressiveness' of the untreated disease, is intrinsically unknown. In order to circumvent this limitation, we also used a population-based estimate derived from CML latency times, i.e. the time between the first leukemic transformation and diagnosis, given that no secondary events change the kinetics of disease emergence (*Online Supplementary Figure S6A*). Sampling from a distribution of CML latency times as reported by Radivoyevitch *et al.*[27] [median latency time = 6.9 years, IQR (5.0,10.1)] (*Online Supplementary Figure S6B*) and taking into account the observed TKI response, we obtained an individual distribution of possible proliferation rates p_Y for each patient (*Online Supplementary Figure S6C*). In other words, we fit our mathematical model several times under different, plausible assumptions about the aggressiveness of the underlying leukemia. For each of those hypothetical but realistic scenarios we calculate the reduction level f_{OPT} (*Online Supplementary Figure S6D*).

Based on these different possible scenarios (i.e. different leukemia growth parameters), we calculated for each patient the fraction of its individual values of f_{OPT} which are below 0.5 (*Online Supplementary Figure S6F*). This fraction indicates whether dose-halving appears as a suitable treatment option or not. With these estimates, our model predicts that 90% of the patients in the German IRIS cohort and 81% of the patients in the CML-IV trial who once achieved MR3, could have safely decreased their TKI dose by at least 50%, while maintaining the overall therapy effect on quiescent LSCs (*Online Supplementary Figure S7A*). Therefore, dose-halving is expected to be safe for the majority of patients in sustained remission and might serve as the initial step to estimate the optimal individual dose.

Our results also suggest that the ratio α/β can be used to identify patients who are likely to benefit from dose reduction. We predict that patients with $\alpha/\beta>15$ are very likely to retain the original long-term treatment efficiency after a 50% dose reduction (*Online Supplementary Figure S7B and C*). Furthermore, we derived a condition to identify patients who do not obtain sufficient TKI dose initially. Specifically, we found that patients with slopes $\alpha/\beta<2$ would benefit from dose escalation, while only patients with $\alpha/\beta\gg2$ benefit from dose de-escalation (*Online Supplementary Figure S8*).

The model predictions are similar to results from the DESTINY trial and support the design of new, informative trials

We compared our model results with findings from the DESTINY trial, which studies dose-halving in 174 TKI-treated patients with CML (being either in MR4 or MR3 for at least 12 months) before TKI cessation. The published DESTINY interim-analysis indicates that 93% of the patients showed no loss of MR3 within 12 months post dose reduction.[15] We simulated this scenario by predicting virtual treatment responses from *BCR-ABL1* measurements in the IRIS/CML-IV trials. In particular, we identified 122 patient time courses fulfilling the inclusion criteria of the DESTINY protocol (> 3 years under TKI, > 1 year in MR3) and simulated a virtual TKI dose reduction according to the DESTINY protocol at the end of the available follow up for each of those patients. Using the same distribution of latency times as above, we calculated for each patient the fraction of values of p_Y, which lead to loss of MR3 within one year after de-escalation (*Online Supplementary Figure S6E*). This fraction can be interpreted as an estimate for the patient-specific risk of a molecular relapse. We also calculated the expected proportion of relapsed patients within the overall population, as well as in the corresponding subcohorts of patients being in either MR3 or MR4 within the last year before dose reduction (Figure 4A). Although a quantitative comparison should be considered with caution due to potential differences in the study populations and patient compliance, the results predicted for the IRIS/CML-IV patients show qualitatively similar relapse rates as observed in the DESTINY trial. Our findings also suggest that the individual relapse probability is related to the remission level before de-escalation, with patients below MR3.5 having a very low probability of relapse (Figure 4B). Furthermore, we predict that most of the observed relapses are transient, i.e. MR3 regain is expected when continuing the half-dose regimen (Figure 4C). Therefore, we argue that the current focus on exceeding MR3 as an indicator for a potential relapse might be reconsidered in the context of dose de-escalation strategies, while closer monitoring of the disease dynamics following dose reductions should be applied to distinguish transient from permanent *BCR-ABL1* regrowth dynamics.

Although our model does not yet reflect immunological effects, which are proposed to be important determinants of disease dynamics post TKI-cessation, we speculated about the relative impact of dose reductions with respect to treatment stop. Specifically, we simulated the DESTINY cessation protocol for the IRIS and CML-IV patients in our data-set and evaluated the primary DESTINY end point, i.e. the proportion of patients who can de-escalate TKI dose by 50% for one year and then stop treatment completely for two years without losing MR3. Our simulations indicate higher rates of molecular relapse in comparison to a structurally identical control group receiving full-dose before TKI stop (Figure 4D) if no additional immunological control mechanisms are considered. However, comparing dose reductions of different duration, we predict a beneficial effect if patients remain at half-dose for longer before stopping TKI. We conclude that the same cumulative dose is more efficient if applied over a longer time period, thereby emphasizing that the full benefit of TKI dose de-escalation appears in the long term.

Discussion

Our study supports the concept that TKI dose reduction in maintenance therapy can be a safe option for many CML patients who have already achieved a sustained remission. In particular, we dissect the typical biphasic response pattern under continuing TKI treatment and conclude that the TKI effect during the secondary treatment phase is limited by the rare activation of quiescent LSCs. Our simulations predict that the overall treatment effect is maintained for most patients, even if the TKI dose is reduced. Based on their initial treatment response under full dose, we identify patients who most likely benefit from a reduction scheme and present a strategy to estimate a patient-specific, optimal dose. Our results suggest a treatment strategy that could considerably reduce cumulative drug intake and, therefore, decrease drug-mediated side-effects. It might also increase compliance of patients to adhere to the prescribed treatment schedule. On the population level, the long overall survival times of continuously treated CML-patients add a distinct economical aspect to the outlined strategy, as the high treatment costs could be substantially reduced.

The proposed strategy is not restricted to a particular TKI. Although our model encompasses all TKI pharmacokinetic parameters and doses within a single parameter e_{TKI}, we have found that this simple model, using a relative reduction compared to a standard dose of the respective TKI, is equivalent to a more elaborate formulation explicitly considering the daily TKI intake (*Online Supplementary Text S10* and *Online Supplementary Figure S9*). Although a linear dose-response relationship appears to be an appropriate model assumption,[25,26] we emphasize that the potential for dose reductions might be limited by a necessary TKI plasma concentration to ensure drug activity. Based on *in vitro* data for imatinib,[28] we estimate this limit to be in the order of 25% of the original dose (*Online Supplementary Text S11*). Furthermore, recent results from clinical trials[15,29] indicate that a 50% dose reduction is therapeutically active.

Several clinical studies have addressed the potential of dose reductions in various settings. While Naqvi *et al.* report good results from a cohort of newly diagnosed CP-CML patients treated with low-dose dasatinib (50 mg daily) first line,[29] Russo *et al.* study the effect of a one-month-on/on-month-off imatinib regimen in elderly patients after at least two years of initial full-dose therapy.[30] In the later study, one-third of the patients lost their previous remission levels. Based on our model, we speculate whether the extended treatment interruptions might have added to this outcome as no therapeutically active TKI concentrations were achieved during this time. Furthermore, we have seen that several clinical studies reported a clear advantage of more potent TKI to achieve molecular response earlier as compared to standard-dose imatinib.[8,31,32] However, the corresponding advantage in long-term survival is less pronounced, and we suggest that it is not the drug potency but the rare activation of LSCs that marginalizes the survival benefit.

In order to test our predictions in a controlled clinical setting, we suggest an approach in which the dynamics of initial treatment response to standard therapy are sufficiently monitored to obtain reliable estimates of the relevant slopes for an informed treatment adaptation (measurements every 3-4 months within the first year and at least every six months thereafter). Only after a clinically relevant remission level (at least MR3) is reached can dose reductions (e.g. < 50% of the initial dose) be considered and these should be accompanied by a detailed follow-up monitoring of *BCR-ABL1* levels to guarantee patient safety and also inform on the validity of our model approach. Further dose reductions might be considered as a second step towards approaching an optimally reduced dose that retains the therapeutic threshold.

Our modeling results predict a transient increase of the number of proliferating LCs after dose reduction. However, because this is only a transient and expected effect, we suggest that the clinical criteria for molecular relapse after dose de-escalation would benefit from considering a follow-up period rather than focusing on fixed thresholds. Although the transient increase in proliferating LCs might increase the chance of acquiring secondary mutations, we reason that this is a marginal effect which needs to be compared with the benefits of reducing treatment-related side-effects. Indeed, assuming that the risk of acquiring a secondary mutation is proportional to the number of proliferating LSCs divisions, this risk increases by only 1.5% when half-dose is applied for three years (after an initial period of three years under standard dose) in comparison with the full-dose scenario (*Online Supplementary Text S12* and *Online Supplementary Figure S10*). We acknowledge that our estimates are based on the assumptions that there are no direct dose-dependent resistance mechanisms. This is supported by the observation that re-starting TKI treatment after relapse in cessation studies proved overall successful and did not suggest a higher tendency for TKI resistances.[33,34]

Our current model does not consider any immunological effects or other more detailed competition mechanism between LCs and their environment. Therefore, our predictions on the risk of molecular relapse after TKI cessation are solely based on the fraction of LCs at the time of TKI stop and the proliferation rate estimated from CML latency times. This implies that a lower residual LSC number ultimately results in a lower relapse risk. Because of this limitation, our model does not reflect potentially beneficial effects resulting from mildly increased abundance

of proliferating LCs which potentially stimulate a patient's immune response.

Although TKI treatment response is still substantially heterogeneous with respect to many clinical parameters, the well-defined CML phenotype and the accessibility of kinetic data on treatment response has made CML a primary target for mathematical model approaches in oncology.[4-6,22] It has also been recognized that the regulation of stem cell quiescence under continuing TKI therapy affects the kinetics of disease eradication.[5,22] Here, we consider a conservative scenario, which only assumes a direct cytotoxic TKI effect on proliferating LCs, but does not alter LSCs quiescence. If assuming an additional TKI-dependent reduction of the activation rate p_{XY}, as previously discussed,[5,17,35-37] any favorable dose de-escalation would lead to an even better long-term response (*Online Supplementary Figure S3*). Recently, using an analytical approach very similar to our formulation, Werner *et al.* developed a mathematical model that allows an estimation of LSC fractions to be made from longitudinal measurements of tumor load.[38] However, this model assumed no direct therapeutic effect on LSCs, which are, therefore, increasing even during treatment.

In summary, we show that the systematic assessment of available clinical data by means of mathematical models has direct clinical implications, but also reveals underlying disease and treatment mechanisms. Our results are substantiated by and support the interim findings of the ongoing DESTINY trial, thereby suggesting a change in current clinical practice and the consideration of TKI dose de-escalation strategies in maintenance therapy. As with any theoretical prediction, our results represent hypotheses that need to be validated in clinical trials. However, modeling approaches can substantially support the design of informative trials. We consider this simulation study as a proof-of-concept for the use of systems medicine to optimize treatment efficacy and to minimize health care costs.

Acknowledgments

Our analysis is based on previously published data from the German cohort of the IRIS trial and the CML IV study. We would like to acknowledge the PIs, study groups and reference labs involved in these studies for generating and providing the data. In particular, we thank Andreas Hochhaus, Rüdiger Hehlman and Martin Mueller.

Funding

The research of ACF was supported by the Excellence Initiative of the German Federal and State Governments (Dresden Junior Fellowship) and by CAPES/Pós-Doutorado no Exterior Grant number 88881.119037/2016-01. This work was further supported by the German Federal Ministry of Education and Research (www.bmbf.de/en/), Grant number 031A424 "HaematoOpt" to IR and Grant number 031A315 "MessAge" to IG. We acknowledge support from the Open Access Publication Funds of the SLUB/TU Dresden.

References

1. Pasic I, Lipton JH. Current approach to the treatment of chronic myeloid leukaemia. Leuk Res. 2017;55:65-78.
2. Rosti G, Castagnetti F, Gugliotta G, Baccarani M. Tyrosine kinase inhibitors in chronic myeloid leukaemia: which, when, for whom? Nat Rev Clin Oncol. 2017;14(3):141-154.
3. Apperley JF. Chronic myeloid leukaemia. Lancet. 2015;385(9976):1447-1459.
4. Michor F, Hughes TP, Iwasa Y, et al. Dynamics of chronic myeloid leukaemia. Nature. 2005;435(7046):1267-1270.
5. Roeder I, Horn M, Glauche I, Hochhaus A, Mueller MC, Loeffler M. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. Nat Med. 2006;12(10):1181-1184.
6. Stein AM, Bottino D, Modur V, et al. BCR-ABL transcript dynamics support the hypothesis that leukemic stem cells are reduced during imatinib treatment. Clin Cancer Res. 2011;17(21):6812-6821.
7. Tang M, Gonen M, Quintas-Cardama A, et al. Dynamics of chronic myeloid leukemia response to long-term targeted therapy reveal treatment effects on leukemic stem cells. Blood. 2011;118(6):1622-1631.
8. Cortes JE, Saglio G, Kantarjian HM, et al. Final 5-Year Study Results of DASISION: The Dasatinib Versus Imatinib Study in Treatment-Naive Chronic Myeloid Leukemia Patients Trial. J Clin Oncol. 2016;34(20):2333-2340.
9. Saussele S, Richter J, Hochhaus A, Mahon FX. The concept of treatment-free remission in chronic myeloid leukemia. Leukemia. 2016;30(8):1638-1647.
10. Mahon FX. Discontinuation of TKI therapy and 'functional' cure for CML. Best Pract Res Clin Haematol. 2016;29(3):308-313.
11. Caldemeyer L, Dugan M, Edwards J, Akard L. Long-Term Side Effects of Tyrosine Kinase Inhibitors in Chronic Myeloid Leukemia. Curr Hematol Malig Rep. 2016;11(2):71-79.
12. Abboud C, Berman E, Cohen A, et al. The price of drugs for chronic myeloid leukemia (CML) is a reflection of the unsustainable prices of cancer drugs: from the perspective of a large group of CML experts. Blood. 2013;121(22):4439-4442.
13. Hartmann JT, Haap M, Kopp HG, Lipp HP. Tyrosine kinase inhibitors - a review on pharmacology, metabolism and side effects. Curr Drug Metab. 2009;10(5):470-481.
14. Flynn KE, Atallah E. Quality of Life and Long-Term Therapy in Patients with Chronic Myeloid Leukemia. Curr Hematol Malig Rep. 2016;11(2):80-85.
15. Clark RE, Polydoros F, Apperley JF, et al. De-escalation of tyrosine kinase inhibitor dose in patients with chronic myeloid leukaemia with stable major molecular response (DESTINY): an interim analysis of a non-randomised, phase 2 trial. Lancet Haematol. 2017;4(7):e310-e316.
16. De- Escalation and Stopping Treatment of Imatinib, Nilotinib or sprYcel in Chronic Myeloid Leukaemia. Available from: https://ClinicalTrials.gov/show/NCT01804985.
17. Horn M, Glauche I, Muller MC, et al. Model-based decision rules reduce the risk of molecular relapse after cessation of tyrosine kinase inhibitor therapy in chronic myeloid leukemia. Blood. 2013;121(2):378-384.
18. STI571 Compared With Interferon Alfa Plus Cytarabine in Treating Patients With Newly Diagnosed Chronic Myelogenous Leukemia. Available from: https://ClinicalTrials.gov/show/NCT00006343.
19. Imatinib Mesylate With or Without Interferon Alfa or Cytarabine Compared With Interferon Alfa Followed by Donor Stem Cell Transplant in Treating Patients With Newly Diagnosed Chronic Phase Chronic Myelogenous Leukemia. Available from: https:// ClinicalTrials.gov/show/NCT00055874.
20. O'Brien SG, Guilhot F, Larson RA, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med. 2003;348(11):994-1004.
21. Hehlmann R, Lauseker M, Jung-Munkwitz S, et al. Tolerability-adapted imatinib 800 mg/d versus 400 mg/d versus 400 mg/d plus interferon-alpha in newly diagnosed chronic myeloid leukemia. J Clin Oncol. 2011;29(12):1634-1642.
22. Komarova NL, Wodarz D. Effect of cellular quiescence on the success of targeted CML therapy. PLoS One. 2007;2(10):e990.
23. Cortes JE, Kantarjian HM, Goldberg SL, et al. High-Dose Imatinib in Newly Diagnosed Chronic-Phase Chronic Myeloid Leukemia: High Rates of Rapid Cytogenetic and Molecular Responses. J Clin Oncol. 2009;27(28):4754-4759.
24. Kalmanti L, Saussele S, Lauseker M, et al. Safety and efficacy of imatinib in CML over a period of 10 years: data from the random-

ized CML-study IV. Leukemia. 2015;29 (5):1123-1132.
25. Deininger MW, O'Brien SG, Ford JM, Druker BJ. Practical management of patients with chronic myeloid leukemia receiving imatinib. J Clin Oncol. 2003;21 (8):1637-1647.
26. La Rosee P, Johnson K, Corbin AS, et al. In vitro efficacy of combined treatment depends on the underlying mechanism of resistance in imatinib-resistant Bcr-Abl-positive cell lines. Blood. 2004;103(1):208-215.
27. Radivoyevitch T, Hlatky L, Landaw J, Sachs RK. Quantitative modeling of chronic myeloid leukemia: insights from radiobiology. Blood. 2012;119(19):4363-4371.
28. Peng B, Lloyd P, Schran H. Clinical pharmacokinetics of imatinib. Clin Pharmacokinet. 2005;44(9):879-894.
29. Naqvi K, Jabbour E, Skinner J, et al. Early results of lower dose dasatinib (50 mg daily) as frontline therapy for newly diagnosed chronic-phase chronic myeloid leukemia. Cancer. 2018;35(15):e18551.
30. Russo D, Martinelli G, Malagola M, et al. Effects and outcome of a policy of intermittent imatinib treatment in elderly patients with chronic myeloid leukemia. Blood. 2013;121(26):5138-5144.
31. Hehlmann R, Lauseker M, Saussele S, et al. Assessment of imatinib as first-line treatment of chronic myeloid leukemia: 10-year survival results of the randomized CML study IV and impact of non-CML determinants. Leukemia. 2017;31(11):2398-2406.
32. Hochhaus A, Saglio G, Hughes TP, et al. Long-term benefits and risks of frontline nilotinib vs imatinib for chronic myeloid leukemia in chronic phase: 5-year update of the randomized ENESTnd trial. Leukemia. 2016;30(5):1044-1054.
33. Mahon FX, Rea D, Guilhot J, et al. Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre Stop Imatinib (STIM) trial. Lancet Oncol. 2010;11(11):1029-1035.
34. Ross DM, Branford S, Seymour JF, et al. Safety and efficacy of imatinib cessation for CML patients with stable undetectable minimal residual disease: results from the TWISTER study. Blood. 2013;122(4):515-522.
35. Graham SM, Jorgensen HG, Allan E, et al. Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to STI571 in vitro. Blood. 2002;99(1):319-325.
36. Glauche I, Horn M, Roeder I. Leukaemia stem cells: hit or miss? Br J Cancer. 2007;96(4):677-678.
37. Besse A, Lepoutre T, Bernard S. Long-term treatment effects in chronic myeloid leukemia. J. Math. Biol. 2017;75(3):733-758.
38. Werner B, Scott JG, Sottoriva A, Anderson AR, Traulsen A, Altrock PM. The Cancer Stem Cell Fraction in Hierarchically Organized Tumors Can Be Estimated Using Mathematical Modeling and Patient-Specific Treatment Trajectories. Cancer Res. 2016;76(7):1705-1713.