

ARTICLE OPEN



Early dynamics of chronic myeloid leukemia on nilotinib predicts deep molecular response

Yuji Okamoto^{1,2,23,24}, Mitsuhiro Hirano^{2,24}, Kai Morino^{1,3}, Masashi K. Kajita^{1,4,5}✉, Shinji Nakaoka^{1,6,7}, Mayuko Tsuda², Kei-ji Sugimoto⁸, Shigehisa Tamaki⁹, Junichi Hisatake¹⁰, Hisayuki Yokoyama¹¹, Tadahiko Igarashi¹², Atsushi Shinagawa¹³, Takeaki Sugawara¹⁴, Satoru Hara¹⁵, Kazuhisa Fujikawa¹⁶, Seiichi Shimizu¹⁷, Toshiaki Yujiri¹⁸, Hisashi Wakita¹⁹, Kaichi Nishiwaki²⁰, Arinobu Tojo^{2,21} and Kazuyuki Aihara^{1,22}✉

Chronic myeloid leukemia (CML) is a myeloproliferative disorder caused by the *BCR-ABL1* tyrosine kinase. Although *ABL1*-specific tyrosine kinase inhibitors (TKIs) including nilotinib have dramatically improved the prognosis of patients with CML, the TKI efficacy depends on the individual patient. In this work, we found that the patients with different nilotinib responses can be classified by using the estimated parameters of our simple dynamical model with two common laboratory findings. Furthermore, our proposed method identified patients who failed to achieve a treatment goal with high fidelity according to the data collected only at three initial time points during nilotinib therapy. Since our model relies on the general properties of TKI response, our framework would be applicable to CML patients who receive frontline nilotinib or other TKIs.

npj Systems Biology and Applications (2022)8:39; <https://doi.org/10.1038/s41540-022-00248-3>

INTRODUCTION

Chronic myeloid leukemia (CML) originates from a hematopoietic stem cell affected by a reciprocal translocation between chromosomes 9 and 22, which results in the formation of the *BCR-ABL1* fusion gene in the short derivative of chromosome 22 (Ph chromosome)¹. In this rearrangement, latent *ABL1* tyrosine kinase is constitutively activated by its oligomerization mediated by the N-terminal region of BCR and causes an unregulated expansion of myeloid cells, especially mature granulocytes². CML is generally diagnosed during the stable chronic phase (CP). Since *BCR-ABL1*-positive cells (called CML cells in the following) are quite dependent on *BCR-ABL1* for their growth and survival, *ABL1*-specific tyrosine kinase inhibitors (TKIs) have clinically demonstrated substantial reduction of CML cells and restoration of normal hematopoiesis in the vast majority of CP-CML patients^{3–7}.

At present, frontline TKIs for newly diagnosed CP-CML include imatinib³, dasatinib⁴, nilotinib^{5–7}, and bosutinib⁸, among which the latter three are the second generation TKIs with more potent activity than the former one. The actual therapeutic response to TKIs is estimated by quantitative polymerase chain reaction analysis of *BCR-ABL1* transcripts in peripheral blood and is represented by the ratio of *BCR-ABL1* mRNA to *ABL1* mRNA on the international scale (IS)^{9–11}. The consensus optimal response to TKIs in CP-CML is major molecular response (MMR, IS $\leq 0.1\%$) within 1 year of therapy¹², whereas a

significant number of CP-CML patients achieve superior or deep molecular response (DMR; MR4.5, IS $\leq 0.0032\%$)^{13–15}, especially on frontline nilotinib or dasatinib^{4–7}. Furthermore, a recent series of clinical trials on TKI discontinuation revealed that nearly a half of CP-CML patients maintained treatment-free remission (TFR) following >2 years of DMR on TKIs¹⁶. Thus, to ultimately reach TFR, the practical target of CML treatment is to achieve DMR as early as possible after frontline TKI therapy, and the switch to another TKI may be a therapeutic option when DMR is not likely to be obtained. Thus, an algorithm to enable the early prediction of DMR achievement on TKI therapy would be quite valuable.

In this paper, we propose a mathematical framework that can predict the efficacy of nilotinib in each CML patient only from short time-series data. We constructed a simple classification method that can estimate whether a CML patient reaches DMR or not within 24 months after nilotinib administration. The key to success is to focus on the dynamical behavior of CML cells.

RESULTS

Current scoring systems and guidelines are not suitable for DMR classification

The dataset analyzed in this study was obtained in a phase II clinical trial named the N-road study¹⁷ that aimed to examine the

¹Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan. ²Division of Molecular Therapy, Advanced Clinical Research Center, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan. ³Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Fukuoka 816-8580, Japan. ⁴Department of Applied Chemistry and Biotechnology, Faculty of Engineering, University of Fukui, Fukui 910-8507, Japan. ⁵Life Science Innovation Center, University of Fukui, Fukui 910-8507, Japan. ⁶Faculty of Advanced Life Science, Hokkaido University, Hokkaido 060-0810, Japan. ⁷PRESTO, Japan Science and Technology Agency, Tokyo 102-0076, Japan. ⁸Division of Hematology, Juntendo University Urayasu Hospital, Chiba 279-0021, Japan. ⁹Department of Hematology, Japanese Red Cross Ise Hospital, Mie 516-8512, Japan. ¹⁰Department of Hematology, Japanese Red Cross Omori Hospital, Tokyo 143-8527, Japan. ¹¹Department of Hematology, National Hospital Organization, Sendai Medical Center, Miyagi 983-8520, Japan. ¹²Division of Hematology and Oncology, Gunma Cancer Center, Gunma 373-8550, Japan. ¹³Department of Internal Medicine, Hitachi General Hospital, Ibaraki 317-0077, Japan. ¹⁴Division of Hematology-Oncology, Chiba Cancer Center, Chiba 260-8717, Japan. ¹⁵Department of Hematology, Chiba Rosai Hospital, Chiba 290-0003, Japan. ¹⁶Department of Hematology, Chibaken Saiseikai Narashino Hospital, Chiba 275-8580, Japan. ¹⁷Department of Hematology, Tsuchiura Kyodo General Hospital, Ibaraki 300-0028, Japan. ¹⁸Third Department of Internal Medicine, Yamaguchi University, Yamaguchi 755-0046, Japan. ¹⁹Division of Hematology and Oncology, Japanese Red Cross Narita Hospital, Chiba 286-8523, Japan. ²⁰Division of Oncology and Hematology, Jikei University Kashiwa Hospital, Chiba 277-8567, Japan. ²¹Institute of Innovation Advancement, Tokyo Medical and Dental University, Tokyo 113-8510, Japan. ²²International Research Center for Neurointelligence (WPI-IRCN), The University of Tokyo, Tokyo 113-0033, Japan. ²³Present address: Department of Biomedical Data Intelligence, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan. ²⁴These authors contributed equally: Yuji Okamoto, Mitsuhiro Hirano. ✉email: masashi.kajita@gmail.com; kaihara@g.ecc.u-tokyo.ac.jp

safety and efficacy of nilotinib for newly-diagnosed CML patients. Complete blood counts (CBCs) and IS were measured in a single certified laboratory center every 3 months for 2 years. In this study, we analyzed the data from the eligible 32 patients. Hereafter, among the eligible patients, we defined 18 patients who have reached DMR by 24 months as DMR patients and remaining 14 patients as non-DMR patients (Supplementary Figs. 1–2). See Supplementary Materials for the detailed information.

Recently, K. Sasaki et al. reported optimal IS values that predict the highest probability of reaching a sustained DMR at specific time points between 3 and 12 months¹⁸. However, to the best of our knowledge, there are neither scoring systems nor algorithms for early prediction of DMR achievement in CML patients under TKI therapy. Therefore, we examined whether the current scoring systems and guidelines in CML clinical practice could be adapted to predict DMR. The European LeukemiaNet (ELN) guideline in 2013¹² recommends the EUTOS prognostic score for prediction of complete cytogenetic response (CCyR) and progression-free survival (PFS) on frontline imatinib. In the N-road study dataset¹⁷, the EUTOS score did not precisely discriminate DMR from non-DMR patients (Fig. 1a). We tested another prognostic score, the EUTOS long-term survival (ELTS) score, which is newly adopted in ELN guideline 2020¹⁵. The ELTS score was developed for prediction of long-term survival considering leukemia-related death. As well as the EUTOS score, the ELTS score also did not distinguish between DMR and non-DMR patients in the N-road study (Fig. 1b). In addition, the definition of response to frontline TKIs employed in the ELN^{12,15} required at least 12-month observation to discriminate the two populations in the same cohort (Fig. 1c–e). Then, we sought for key features to construct a better method of early DMR prediction.

Key features for DMR prediction

Potential key features should be found in the dynamics of CML cells in CML patients on TKIs because mathematical studies on CML^{19–25} explained the characteristic dynamics of CML cells in CML patients on TKIs and successfully predicted relapse. Particularly, M. Horn et al. suggested that relapse was predictable based on residual leukemic stem cells and reduction slopes of IS²³. However, for the purpose of DMR prediction from short-term measurements, most models are too complex to obtain suitable parameter values for each patient. Numerous approaches have been proposed to overcome these data-shortage problems such as high-dimensional data in general²⁶ and prostate cancer in specific^{27,28}. In this research, we overcome this problem by focusing on two important key features on the dynamics of CML cells in CML patients and by proposing an adequately-simple mathematical model, which can reproduce the two key features from a small dataset.

Two key features for early DMR prediction in CML patients are the decreasing rate and the convergence value of IS time-series data (Fig. 2). Both early and late dynamics of IS can be clearly represented by these two features (Fig. 2a, b). Additionally, the distributions of these two features in DMR and non-DMR patients seem to be distinguishable (Fig. 2c, d). Thus, we can discriminate DMR patients from non-DMR patients, if we accurately estimate these two features from short time series of IS values. Hereafter, we discuss a mathematical way to estimate these two features.

Our proposed method showed a good performance for distinguishing between DMR and non-DMR patients

First, we propose a mathematical model, which describes the time evolution of the white blood cell (WBC) count in CML patients (Eq. 1 in Methods). Our model includes two variables: the normal WBC (namely *BCR-ABL1*-negative) count and CML cell (namely *BCR-ABL1*-positive) count. Furthermore, each variable has two parameters: the recovery rate and the convergence value for the

normal WBC count, and the reduction rate and the convergence value for the CML cell count. Since both the cell counts are mathematically calculated by the measurement of IS and the total WBC count in our model (Eqs. 2–3 in Methods), we can estimate the variables and parameters (Fig. 2e, f). The first feature, the decreasing rate of IS, corresponds to the reduction rate of the CML cells. The second feature, the convergence value of IS, corresponds to the convergence value of the CML cells. Our mathematical model sufficiently approximates the time series of IS for various CML patients (Fig. 3a and Supplementary Fig. 2). These findings indicate the validity of our proposed model.

Second, we discuss how we deal with estimated parameter values to classify the TKI response (Fig. 3b–d). DMR and non-DMR patients were clearly separated on a parameter plane of the CML cell reduction rate and convergence value (Fig. 3c). To obtain appropriate threshold values of the two patients groups, we determined horizontal and vertical lines on the parameter plane by using all the eligible patient data points (see “DMR prediction criteria” in Methods). With the threshold values, we positively discriminated between DMR and non-DMR patients in the analyzed N-road dataset with 100 % accuracy (Fig. 3d). On the other hand, another set of parameters did not clearly separate patients (Fig. 3b). We note that, in Fig. 3, all the time points of all patients data were used to estimate our model parameters. However, the parameters can be estimated based only on three time points for each because of our model’s simplicity. To classify patients into DMR or non-DMR based on the estimated parameters, the threshold values that separate DMR and non-DMR patients in the parameter plane still need to be determined.

Third, we discuss an approach to determine the threshold values of the CML cell reduction rate and convergence value in a CML patient by utilizing information from other patients. We employed the leave-one-out cross-validation. The dataset with 32 patients is divided into two groups, one is the training set composed of 31 patients, and the other is the validation dataset composed of one patient. By applying the aforementioned framework to data from the training dataset, we obtained estimated parameter values of each patient and optimized the threshold values under a certain criterion. For the optimization, we employed three different objective functions, i.e., maximizing accuracy, maximizing sensitivity, and maximizing specificity. An optimal threshold value is given by maximizing a corresponding objective function (see also “DMR prediction criteria” in Methods). Based on the optimized threshold values, we predicted a patient of the test dataset whether the patient will reach DMR or non-DMR from short-term measurements. For the test patient, we set the correctness score as 1 or 0 if the DMR prediction successes or fails, respectively. After the cross-validation, i.e., the 32 times of the prediction test, we obtain the classification performance as an average of the correctness scores among 32 patients (see also “Performance of DMR prediction” in Methods). In Fig. 4, we summarize the performance of DMR predictions. Our proposed approach, which uses data only at 3 initial time points (at 0, 3, and 6 months), reached the performance of the ELN definition at 12 months of therapy. Furthermore, our method enhanced specificity and F1 scores compared to the EUTOS scoring system and ELN guidelines. As for the ELTS scoring system, it estimated most patients as “Intermediate risk” or “High risk”. By assuming that these classifications correspond to non-DMR, we consider it difficult for the ELTS scoring system to detect non-DMR patients (see also Supplementary Fig. 6 for the confusion matrices of the ELTS score-based prediction). In Supplementary Figs. 7–8, we analyzed the EUTOS score-based and ELTS score-based prediction results based on the patient’s characteristics using the parameters for these scores (see also “EUTOS score-based and ELTS score-based DMR prediction” of Supplementary Materials). We note that the difference of our three methods relies on the way to obtain the threshold values, i.e., maximizing accuracy, maximizing

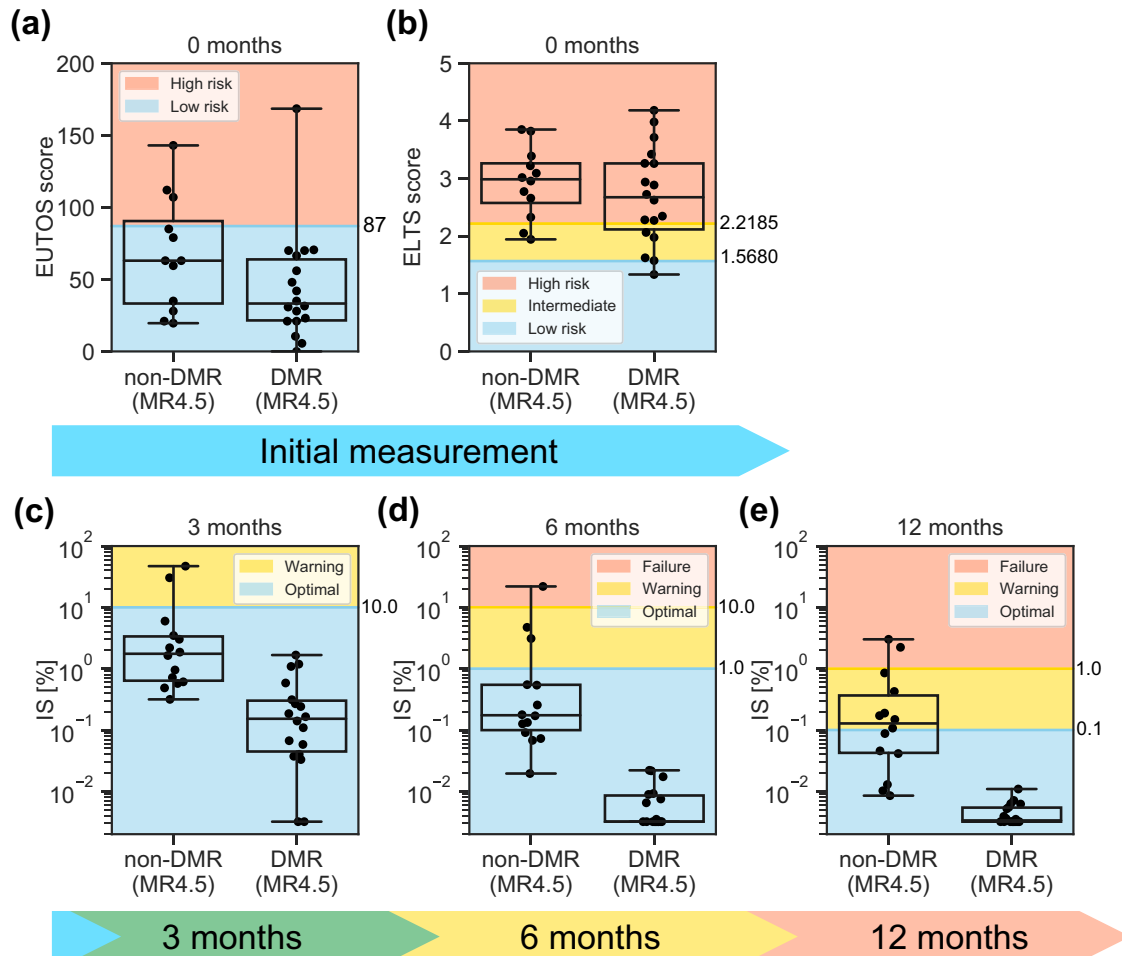


Fig. 1 Several criteria of the current scoring systems and guidelines are not suitable for DMR classification. Each dot indicates each patient value within the N-road dataset. The top and bottom bars of the boxes indicate the maximum and minimum values, respectively. The top, center, and bottom lines are the upper quartiles, medians, and lower quartiles, respectively. Patients are divided into two groups: a DMR patient set and a non-DMR patient set. **a** The EUTOS score did not predict non-DMR patients. The patient is considered “High risk” if the EUTOS score of a CML patient is >87 and “Low risk” otherwise. **b** The ELTS score did not also predict non-DMR patients. A CML patient is considered as “High risk” if the patient’s ELTS score is >2.2185 , “Low risk” if the score is 1.5680 or lower, and “Intermediate risk” otherwise. **c–e** According to the ELN guideline, we plotted patient IS and separated them into three groups: “Optimal”, “Warning”, and “Failure” for the data at **(c)** 3, **(d)** 6, and **(e)** 12 months after the initiation of TKI administration. In **c**, **d** many non-DMR patients were predicted as “Optimal”, indicating the failure of these predictions.

sensitivity, and maximizing specificity (see also “DMR prediction criteria” in Methods). The classification performances of our three optimization methods balanced the four scores: accuracy, sensitivity, specificity, and F1 scores. In addition, the performance scores of our three methods were equal to or higher than those of most current scoring systems and guidelines when we adopted MR4.5 for DMR definition (Fig. 4). The exceptions are the cases of specificity of the ELTS score and sensitivity of the ELN guidelines that achieved 100 % performance. These seemingly good scores come from the biased prediction of the score/guidelines (see Supplementary Fig. 6). The similar results were also obtained for MR4.0 and complete molecular response (CMR) (Supplementary Fig. 5) (see Methods for the definitions of MR4.0 and CMR). Thus, our algorithm has high potential to predict non-DMR patients on frontline nilotinib, and our method should improve the current CML treatment.

DISCUSSION

Finally, we discuss whether the presented mathematical model is applicable to other frontline TKIs including imatinib and dasatinib. The *BCR-ABL1* transcript dynamics under imatinib treatment^{21,23}

was similar to the IS dynamics under nilotinib treatment analyzed in our study. Although further investigations are needed, we expect that our model can be applied to imatinib treatment. TKIs achieve rapid complete hematological response, including normal WBC counts, in a vast majority of patients with CML. However, among ever approved TKIs, only dasatinib reveals a peculiar property of rapid mobilization of non-leukemic lymphocytes, especially NK cells and $CD8 + CD57 +$ large granular lymphocytes, in peripheral blood peaking 1–2 h after an oral intake^{29,30}. This unique phenomenon frequently results in transient elevation of WBC counts in blood testing of CML patients even with stable molecular response. Similar effects on blood cell dynamics are not observed with any other TKI. Accordingly, the present mathematical model is likely to be applicable to TKIs except for dasatinib.

In conclusion, it is suggested that our mathematical method has an ability to predict the possibility to achieve DMR in a retrospective cohort of CML patients treated with frontline nilotinib. This outcome is due to the suitable degree of simplicity of our mathematical model: it is simple enough to operate with short-term time series but also complex enough to distinguish between DMR and non-DMR patients. The ultimate goal of CML treatment is likely to be TKI-free maintenance of DMR, and the

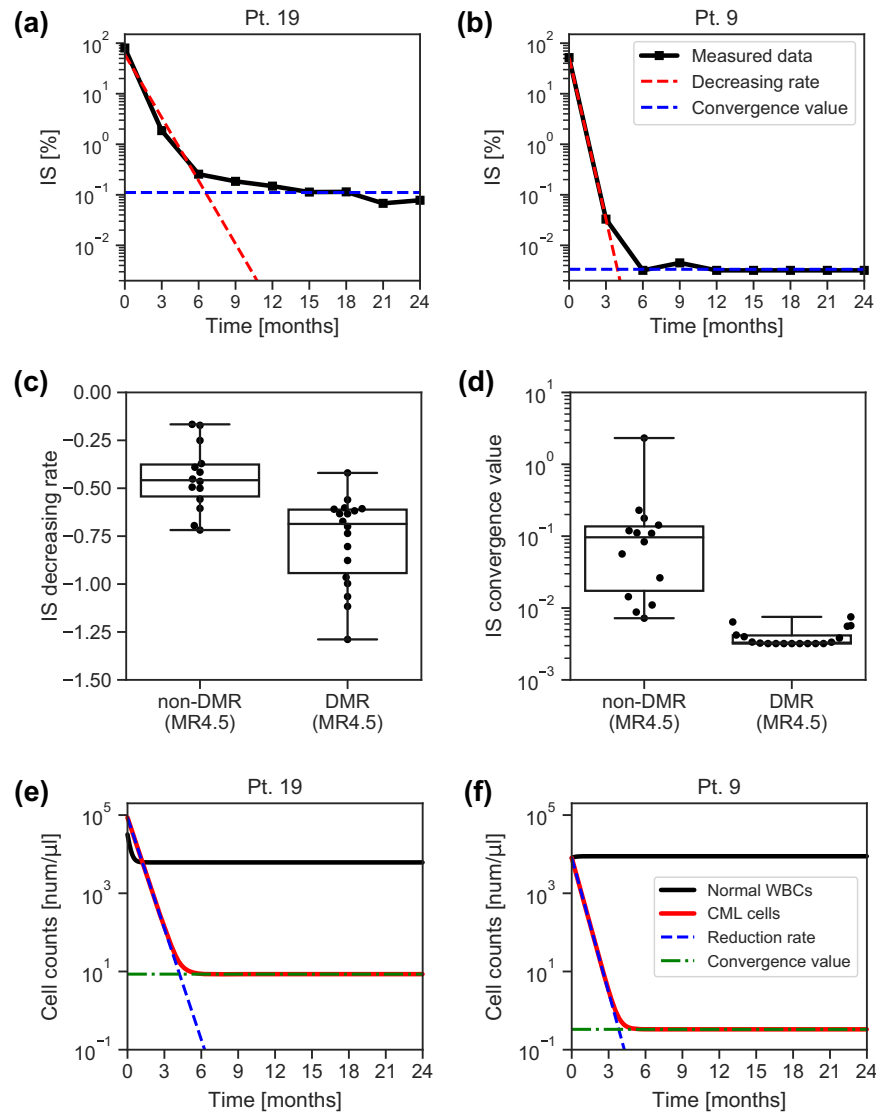


Fig. 2 Key features for DMR prediction: the decreasing rate and convergence value of IS, and the reduction rate and convergence value of CML cells. **a, b** The time series of IS are plotted in **(a)** for a non-DMR patient (patient # 19) and in **(b)** for a DMR patient (patient # 9). The decreasing rate and convergence value of IS are two key features for DMR prediction. A detailed information of the parameter estimation can be found in Methods. **c** The distributions of the IS decreasing rate for non-DMR and that for DMR patients were different. **d** The distributions of the IS convergence value for non-DMR and that for DMR patients were distinct. The definitions of the bars of the boxes, lines, and dots used in the boxplots (**c, d**) are the same as those in Fig. 1. **e, f** We show the estimated dynamics of normal WBCs and CML cells counts in **(e)** for a non-DMR patient (patient # 19) and in **(f)** for a DMR patient (patient # 9). These figures indicate that the reduction rate and convergence value of CML cells play a key role in the distinction between DMR and non-DMR patients.

achievement of durable DMR is a prerequisite for this goal³¹. Furthermore, timely prediction of nilotinib failure to achieve DMR will enable the consideration of other therapeutic approaches³², which will allow sustained DMR leading to TFR. Therefore, the present study contributes to the progress of clinical treatment that is supported by quantitative measurement of disease dynamics and mathematical analysis.

METHODS

Definition of deep molecular response (DMR)

In this study, we employed DMR as the criterion for effective drug response in CML patients. There are multiple definitions of DMR such as MR4.0 ($IS \leq 0.01\%$) and MR4.5 ($IS \leq 0.0032\%$)^{9,12–15,32}. If a patient's IS becomes lower than one of the above threshold values at a certain time point, it is considered that the patient reaches the DMR corresponding to the employed threshold value. There is a

stricter criterion designated as CMR³². A patient is considered to reach CMR, if his/her expression of the *BCR-ABL1* gene is not measurable at two consecutive time points. In this paper, we designated patients as CMR who reached MR4.5 at both 21 and 24 months. Because lower IS implies a more favorable situation for CML patients, CMR is considered to be a better criterion. However, the measurement of low *BCR-ABL1* expression is technically challenging, i.e., non-negligible noise/error is often contained in such measurements. Therefore, we employed MR4.5 for this analysis (Figs. 1–4). In Supplementary Figs. 3–5, for MR4.0, MR4.5, and CMR, we compared the numerical results corresponding to Figs. 1, 3b–d, and 4.

Here, we note a more detailed definition of DMR patients. We designated patients whose IS decreased below 0.0032% at a certain time point as DMR patients. However, in a subset of these DMR patients, their IS fluctuated and/or again increased above 0.0032%. In this study, we treated these patients as DMR patients.

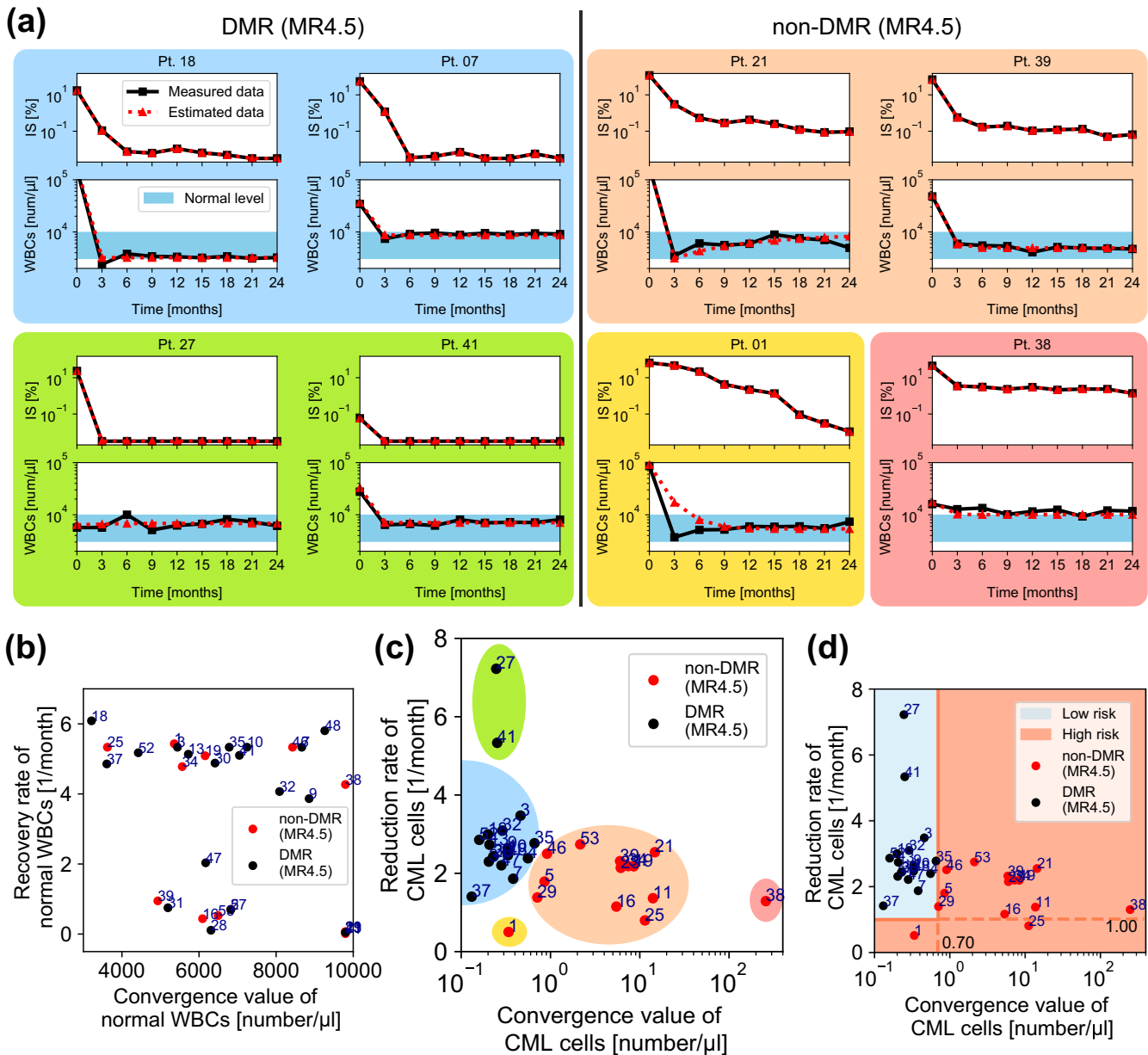


Fig. 3 DMR and non-DMR patients were separated on the plane of the reduction rate and convergence value of CML cells. **a** The estimated time series of CML dynamics (red dashed lines) sufficiently approximated the measured data for IS [%] and for WBC counts [number / μ l] (black solid lines). The patients shown in the colored boxes were chosen as examples from the patients within the same color regions shown in **(c)**. **b** The distribution of the convergence value and recovery rate [1/month] of normal WBCs did not show a correlation in terms of nilotinib response. The number indicated on each dot is patient #. **c** The distribution of the convergence value and reduction rate of CML cells was apparently classified into several subsets: a set including most DMR patients (the blue region), non-DMR patients (the orange region), and outlier patients (other color regions). **d** By choosing optimal threshold values, the distribution of DMR patients shown in **(c)** was perfectly separated from that of non-DMR patients. Notably, this result was obtained by using all patient data points.

In further investigation, this point may be reconsidered, if more detailed clinical information becomes available. One of the reasons for IS reoccurrence is the acquisition of drug resistance. However, we cannot distinguish patients who acquired drug resistance based on the available information, because observations over a longer time span are needed for that purpose.

Detailed information of the N-road trial

We used the dataset from the N-road trial¹⁷, which was a phase II, single-arm, open-label, multi-center clinical study for newly diagnosed CP-CML patients. This study was organized based on previous trials on nilotinib⁵⁻⁷. The CP-CML diagnosis had to be

made by chromosome banding or fluorescent in situ hybridization (FISH) or reverse transcription polymerase chain reaction. The N-road aimed to investigate the safety and efficacy of nilotinib in newly diagnosed CP-CML patients based on the early achievement of CMR. The primary endpoint was the CMR (defined as MR4.5 in the N-road study) achievement rate within 24 months. This trial included 53 patients and was conducted from August 2012 until July 2017. Nilotinib 300 mg twice daily was administered to patients for 24 months. If patients did not reach the criteria for optimal response defined by the N-road study referring to the ELN 2013 definition¹² at any time point, the administration of nilotinib up to 400 mg twice daily as second-line therapy was allowed.

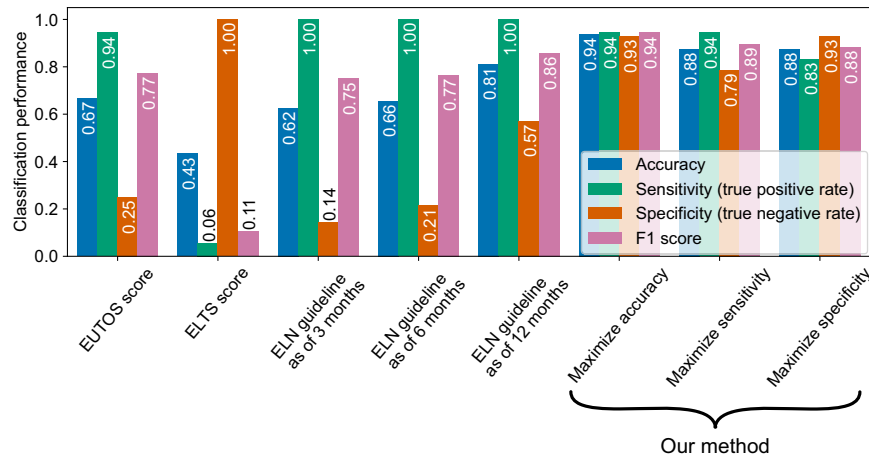


Fig. 4 Our prediction method achieved favorable performance in comparison with the criteria of the current scoring systems and guidelines. We summarize the performance for accuracy (blue bars), sensitivity (green bars), specificity (orange bars), and F1 scores (purple bars) of the EUTOS score, the ELTS score, the ELN guideline for the cases of 3, 6, 12 months, and our proposed method (presented with the three rightmost labels). In the current study, “positive” and “negative” indicate DMR and non-DMR patients, respectively. In the cases of the EUTOS scores, the ELTS scores, and the ELN guidelines, we decided to estimate a target patient as “positive”; if the patient was classified as “Low risk” or “Optimal”. Under this estimation, we obtained accuracy, sensitivity (the true positive rate), and specificity (the true negative rate) for each case. In our setting, the prediction of non-DMR patients is highly important, because the administrated TKI should be promptly changed for them. Thus, our method was superior to the EUTOS score and the ELN guideline because our method showed higher specificity than them. The ELTS score cannot distinguish between non-DMR and DMR patients. Notably, each of our three results was optimized in terms of accuracy, sensitivity, and specificity through the training period. See also “DMR prediction criteria” and “Performance of DMR prediction” in Methods.

Notably, our mathematical framework requires only IS and WBC counts. Other laboratory findings were also obtained but were not used in this paper. In addition, in this study, we analyzed the data from the eligible 32 patients; the remaining 21 patients were ineligible for analysis due to missing data of one or more CBC and/or IS measurements. Furthermore, in this paper, moving to the second-line therapy or not was not thematized.

This trial was approved by the institutional review boards of the participating institutions and was conducted in accordance with the ethical guidelines for medical and health research involving human subjects. This study also adhered to the ethical principles of the Declaration of Helsinki. All patients provided written informed consent to participate in the study. This N-road study is registered in University hospital Medical Information Network (UMIN) system, and its ID is UMIN000008565¹⁷.

Details of the EUTOS score, the ELTS score, and the criteria of the current guideline

Here, we introduce details of the EUTOS score¹⁹, the ELTS score¹⁵, and the ELN guidelines from 2013¹² and 2020¹⁵ to be related to our study. Both the EUTOS score¹⁹ and the ELTS score¹⁵ are the prognostic scores to estimate the survival risk at baseline (i.e., before the TKI administration). The EUTOS score¹⁹ was proposed as a score to predict CCyR and subsequently PFS. The EUTOS score is calculated by the spleen size below costal margin and the basophil ratio as follows:

$$\text{EUTOS score} := 4 \times \text{spleen size below costal margin (cm)} + 7 \times \text{basophil ratio (\%)}$$

If the EUTOS score of a patient is >87, the patient is designated as being in a “High risk” situation. The ELTS score was recently adopted in the ELN guideline 2020¹⁵ and proposed for long-term survival considering leukemia-related death. At baseline, the ELTS score is calculated by the age in completed years, the spleen size below costal margin, the percentage of blasts in peripheral blood,

and the platelet count as follows:

$$\begin{aligned} \text{ELTS score} := & 0.0025 \times \left(\frac{\text{age in completed years}}{10} \right)^3 + 0.0615 \\ & \times \text{spleen size below costal margin (cm)} \\ & + 0.1052 \times \text{blasts in peripheral blood (\%)} \\ & + 0.4104 \times \left(\frac{\text{platelet count}}{1000} \right)^{-0.5} \end{aligned}$$

A CML patient is considered as “High risk” if the patient’s ELTS score is >2.2185, “Low risk” if the score is <1.5680, and “Intermediate risk” otherwise.

In the ELN guideline 2013¹², the condition of a patient is classified into three stages according to the relationship among IS, chromosome banding analysis (CBA) from bone marrow cells, FISH of blood interphase cell nuclei, and the duration since the beginning of TKI administration: The three stages are “Failure”, “Warning”, and “Optimal” starting from the worst. At the 3-month time point, the threshold value between “Warning” and “Optimal” is set at 10% of IS. At the 6-month time point, the threshold value between “Failure” and “Warning” is set at 10%, and that between “Warning” and “Optimal” is at 1%. At the 12-month time point, these threshold values are set at 1% and 0.1%, respectively. The above information is illustrated in Fig. 1.

We should note that the ELN guideline 2020¹⁵ changed the criteria of the risk stage. However, the changes did not affect the results obtained from our dataset. Here, we explain the change related to our analysis and why it did not affect our result. One of the changes related to our analysis is the “Failure” stage criteria at the 3-month time point. The criteria of the guideline in 2013¹² do not explicitly refer to IS values, but those in 2020¹⁵ refer to IS values. According to the guideline in 2020, a patient is confirmed as the “Failure” stage if the IS value satisfies the following both conditions: the IS value is >10% at the 3 months; the same IS level is confirmed again before the 6 months. The N-road dataset has IS values every 3 months. Therefore, no patient satisfied the second condition. Thus, the changes in the ELN guideline 2020 did not affect our results.

We tested the performance of the EUTOS score, the ELTS score, and the ELN guideline for DMR prediction in our dataset as shown in Fig. 4. It should be noted that the EUTOS score, the ELTS score, and the guideline were not proposed for DMR prediction. In the cases of the EUTOS score and the guideline with 3 months and 6 months (labeled as “ELN guideline as of 3 months” and “ELN guideline as of 6 months”), the results indicated high “sensitivity” and low “specificity”; therefore, the prediction of non-DMR patients based on them is very difficult. On the contrary, for the ELN guideline as of 12 months, specificity was higher at 57%. Thus, for accurate prediction of non-DMR patients based on the guideline the 12 months duration is necessary. The ELTS score has higher “specificity” than the ELN guideline as of 12 months. However, the “sensitivity” and “accuracy” are significantly low. This result indicates that the ELTS score is not suitable for the prediction of DMR patients (see Supplementary Figs. 3 and 5 for the performance of the EUTOS score, the ELTS score, and the guideline in the cases of MR4.0, MR4.5, and CMR). The performance of the current guidelines, excluding the EUTOS score and the ELTS score, was tested in 32 patients. Since spleen size data, which are required to calculate the EUTOS and the ELTS scores, were not available for two patients, the EUTOS and ELTS scores were evaluated in the remaining 30 patients.

Dynamical model of CML cells

Here, we introduce the details of our dynamical model. We modeled the dynamics within peripheral blood of CML patients. Administration of an effective TKI results in the reduction of CML cells (*BCR-ABL1*-positive WBC) and the increase of normal WBCs (*BCR-ABL1*-negative WBCs) in peripheral blood. In this study, we assumed that the number of CML cells and that of normal WBCs exponentially change with time. We also assumed that these numbers asymptotically converge to constant values. Then, the time series of these numbers can be described as a couple of linear differential equations as follows:

$$\begin{aligned} \frac{dx_i(t)}{dt} &= a_i x_i^* - a_i x_i(t), \\ \frac{dy_i(t)}{dt} &= b_i y_i^* - b_i y_i(t), \end{aligned} \quad (1)$$

where $x_i(t)$ and $y_i(t)$ represent the number of normal WBCs and that of CML cells for the i th patient at time t , respectively. Their convergence values are denoted by x_i^* and y_i^* . The coefficients a_i and b_i represent the recovery rate of normal WBCs and the reduction rate of CML cells for the i th patient, respectively. Examples of the dynamics generated by the model are shown in Fig. 2e and f.

Relationship between measured values and CML dynamics

We explain the relationship between our model and measurements from CML patients. We can convert two measurement variables, the WBC count and IS, into two state variables in our model, $x_i(t)$ and $y_i(t)$, via the following procedure. The WBC count clearly agrees with the sum of $x_i(t)$ and $y_i(t)$. Then, the WBC count for the i th patient at time t , denoted by $WBC_i(t)$, is described as follows:

$$WBC_i(t) = x_i(t) + y_i(t). \quad (2)$$

The definition of IS is the ratio of the *BCR-ABL1* gene (mRNA) expression to the *ABL1* gene (mRNA) expression: The *BCR-ABL1* gene is expressed in CML cells, whereas the *ABL1* gene is expressed in both CML cells and normal WBCs. Here, we assume that the total mRNA expression level is proportional to the number of cells. We denote the ratio of the total *ABL1* mRNA expression level in normal WBCs to that in CML cells by c . Thus, IS for the i th patient at time t , denoted by $IS_i(t)$ [%], is described as

follows:

$$IS_i(t) = \max\left(\frac{y_i(t)}{cx_i(t) + y_i(t)} \times k_{IS} \times 100, 0.0032\right), \quad (3)$$

where k_{IS} is the conversion factor of IS, and our dataset of the N-road adopts $k_{IS} = 1.2$. Because the detection limit of IS measurements in this dataset is 0.0032%, we set the lower bound as 0.0032.

These Eqs. (2) and (3) can transform the WBC count $WBC_i(t)$ and IS $IS_i(t)$ into $x_i(t)$ and $y_i(t)$, respectively. However, the constant ratio c is generally unknown. Thus, we estimated the ratio c and the parameters of the CML dynamical model a_i , b_i , x_i^* , and y_i^* from the measurements. See Methods for the detailed methods for the parameter estimation.

IS is an international standard value that is converted using a laboratory-specific conversion factor (CF), which is calculated by comparing the actual measured values of patients' samples between each laboratory and the international reference laboratory¹¹. Bio Medical Laboratories (BML), which performed IS testing in the N-road study, obtained $CF = 1.2$ from the Institute of Medical and Veterinary Science (IMVS) in Adelaide, Australia, which is the reference laboratory for international standards. We adopted $CF = 1.2$ for the value of k_{IS} in the Eq. (3).

Fitting IS time-series data with piecewise functions

In this part, we explain the method to estimate IS dynamics parameters shown in Fig. 2c, d. To obtain the decreasing rate and convergence value of IS, we defined a piecewise function with two domains as follows:

$$f_i(t) = \begin{cases} \log IS_i(0) + \frac{(\log IS_i^* - \log IS_i(0))}{\tau_i} \times t, & \text{if } 0 \leq t < \tau_i, \\ \log IS_i^*, & \text{if } t \geq \tau_i, \end{cases}$$

where $10^{IS_i^*}$ is the IS convergence value, $(\log IS_i^* - \log IS_i(0))/\tau_i$ is the IS decreasing rate of the i th patient, and τ_i [month] is a boundary between the two domains. Then we solve the following optimization problem:

$$\underset{IS_i^*, \tau_i}{\text{minimize}} \sum_{i \in P, t=0,3,\dots,24} (f_i(t) - \log \widehat{IS}_i(t))^2.$$

We used the Adam optimizer of the TensorFlow machine learning library to solve the problem.

Parameter estimation of the CML dynamical model

Here, we introduce the optimization method to estimate the parameters of Eqs. (1), (2), and (3). The measurement value of the WBC count and that of IS for the i th patient at the time t are denoted by $WBC_i(t)$ and $\widehat{IS}_i(t)$, respectively. Then, the optimization problem is described as follows:

$$\underset{x_i^*, y_i^*, a_i, b_i, c}{\text{minimize}} \sum_{i \in P, t=0,3,\dots,24} (\ln IS_i(t) - \ln \widehat{IS}_i(t))^2 + (\ln WBC_i(t) - \ln \widehat{WBC}_i(t))^2$$

$$\text{subject to } x_{\min}^* \leq x_i^* \leq x_{\max}^*, y_{\min}^* \leq y_i^* \leq y_{\max}^*, a_i, b_i, c > 0.$$

As for the fitting IS data with the piecewise function, we used the Adam optimizer to solve this optimization problem. We employed the method of Lagrange multipliers, and the constraint is

$$\sum_{i \in P} \{R(x_{\min}^* - x_i^*) + R(-x_{\max}^* + x_i^*) + R(y_{\min}^* - y_i^*) + R(-a_i) + R(-b_i)\} = 0,$$

where $R(x) := \max(x, 0)$ is the ramp function. In our setting, the Lagrange multiplier was set to 100. We set the lower bound $x_{\min}^* = 3100$ [count/ μ l] and the upper bound $x_{\max}^* = 9800$ [count/ μ l] in the constraint condition to maintain the convergence values of normal WBCs within the normal levels of healthy humans. Furthermore, we set the lower bound y_{\min}^* in the constraint

condition to keep $IS_i(t)$ beyond the detection limit of IS, i.e., $y_{i,\min}^*$ is obtained by solving $\frac{y_{i,\min}^*}{c \times \hat{IS}_i(t) + y_{i,\min}^*} \times k_{IS} \times 100 = 0.0032$. A set of patient indices is denoted by P_* . In Fig. 3, the parameters are estimated by utilizing all patient data, i.e., P_* contains all patients ($i = 1, \dots, 32$). In Fig. 4, we checked how our proposed method outperformed the existing guidelines in terms of DMR patient prediction using short time series. For this sake, we did leave-one-out cross-validation, i.e., P_* contained 31 patients used for the training dataset, and the remaining one patient was the prediction-target patient. We repeated this data splitting 32 times employing subsequently all 32 patients as the prediction target patient.

Parameter estimation for the prediction-target patient

Here, we introduce an approach to estimate parameters of the prediction-target patient only from three data points. As we have discussed in Fig. 3b–d, the classification is performed on the plane of the CML cell convergence value y_i^* and its reduction rate b_i . Thus, we introduce a method to estimate the two parameters from three data points, which are specifically denoted by \hat{y}_i^* and \hat{b}_i . Notably, the gene expression ratio parameter c is already estimated in the previous training step. Hereafter, we assume that all ISs of the three data points are greater than the detection limit. Otherwise, clearly according to the definition of DMR, the patient must be a DMR patient. Under this assumption, from Eqs. (2) and (3), the estimated number of CML cells at time t , denoted by $\hat{y}_i(t)$, is obtained as follows:

$$\hat{y}_i(t) = \frac{c \times \hat{IS}_i(t)}{100 \times k_{IS} + (c - 1) \times \hat{IS}_i(t)} \widehat{WBC}_i(t).$$

Then, the estimated convergence value \hat{y}_i^* and the reduction rate \hat{b}_i are obtained as follows:

$$\hat{y}_i^* = \frac{\hat{y}_i(0)\hat{y}_i(2\Delta t) - \hat{y}_i(\Delta t)^2}{\hat{y}_i(0) - 2\hat{y}_i(\Delta t) + \hat{y}_i(2\Delta t)},$$

$$\hat{b}_i = \frac{1}{\Delta t} \ln \left(\frac{\hat{y}_i(0) - \hat{y}_i(\Delta t)}{\hat{y}_i(\Delta t) - \hat{y}_i(2\Delta t)} \right),$$

where Δt represents the interval of measurement points. In our dataset, the N-road, the data are measured every 3 months, i.e., $\Delta t = 3$ [month], meaning that we utilized the data $\{\widehat{WBC}_i(t), \hat{IS}_i(t)\}$ for $t = 0, 3$, and 6 [months]. Based on the estimated parameters \hat{y}_i^* and \hat{b}_i , our model predicts DMR or non-DMR for the target patient. Note that, according to our DMR definition, we regard a patient as DMR, if the patient's IS value satisfies the DMR definition ($IS \leq 0.0032\%$ for MR4.5) for $t = 0, 3$, or 6 [months].

DMR prediction criteria

We introduce an approach to determine the criteria for DMR prediction. As presented in Fig. 3d, DMR patients should show large enough b_i and small enough y_i^* . To estimate the parameters (\hat{y}_i^*, \hat{b}_i) , we define a parameter region $\{(\hat{y}_i^*, \hat{b}_i) | \hat{y}_i^* \leq Y \text{ and } \hat{b}_i \geq B\}$ for the criteria. If the estimated parameters (\hat{y}_i^*, \hat{b}_i) are included in this region, the patient is predicted to be a DMR patient. These threshold values Y and B are determined by using the training data, i.e., patients whose ID are included in a set P_* . We explored the optimal Y and B that maximize the accuracy, sensitivity, or specificity. We used the exploration intervals 0.1 and 0.05 for Y and B , respectively. These indices are the common concept in the classification context as follows. The accuracy is the ratio of patients who are successfully predicted. The sensitivity, the true positive rate, is the ratio of DMR patients who are successfully predicted. The specificity, the true negative rate, is the ratio of non-DMR patients who are successfully predicted. We note that

“positive” (“negative”) indicates “prediction as DMR” (“prediction as non-DMR”). In this sense, “true positive” (“false positive”) indicates that the prediction as DMR is “correct” (“incorrect”).

As described above, if a patient satisfies the DMR definition within its initial three time points, the patient is considered to be a DMR patient. In Fig. 4 and Supplementary Fig. 5, the numbers of such patients were 15 for MR4.0, 10 for MR4.5, and 2 for CMR. All these patients were estimated to be “Optimal” in the case of the ELN guidelines^{12,15}. For MR4.0 and MR4.5, only one patient of them was estimated to be “High risk” in the case of the EUTOS score, and the others were estimated to be “Low risk”. For CMR, all the patients were estimated to be “Low risk” in the case of the EUTOS score. In the case of the ELTS score, only one of the patients for MR4.0 and MR4.5 was estimated to be “Low risk” and the others were estimated to be “High risk” or “Intermediate risk”, and all of the patients for CMR were estimated to be “Intermediate risk”.

Here, for each accuracy, sensitivity, and specificity, we remark how to choose the optimal threshold (Y, B) . In the case of maximizing accuracy, we obtained the multiple possible sets of optimized (Y, B) that maximize accuracy. Then, from the possible sets, we selected the set (Y, B) closest to the average values of the possible sets. In the case of maximizing sensitivity, we obtained the multiple possible sets (Y, B) which achieve the largest sensitivity. However, under this condition of maximizing sensitivity, we can freely magnify the parameter region corresponding to DMR. To keep the area of the region reasonable, from the above possible sets, we again selected parameter sets that achieve the maximal true negative rate. Finally, we selected the largest Y and the smallest B among the sets that had passed the both selections as the optimal set. In the case of maximizing specificity, similar to the case of sensitivity, we obtained the multiple possible sets (Y, B) which achieve the largest specificity. Then, from the possible sets, we again selected parameter sets that achieve the maximal true positive rate. Finally, we selected the smallest Y and the largest B among them as the optimal set.

Performance of DMR prediction

Through the above framework, by maximizing the accuracy, sensitivity, or specificity, we obtained three prediction results as shown in Fig. 4 (for MR4.5) and Supplementary Fig. 5 (for MR4.0, MR4.5, and CMR). We used accuracy, sensitivity, specificity, and F1 score to evaluate the classification performance of the prediction results. F1 score is defined as $F1 \text{ score} := \frac{2TP}{2TP + FP + FN}$, where TP, FP, and FN indicate the number of true positive, false positive, and false negative cases, respectively.

We obtained the classification performances as follows. As we employed the leave-one-out cross-validation, the dataset with 32 patients is divided into two groups, one is the training set composed of 31 patients, and the other is the validation dataset composed of one patient. For the test patient, we set the correctness score as 1 or 0 if the DMR prediction successes or fails, respectively. After the cross-validation, i.e., the 32 times of prediction test, we obtain the classification performance as an average of the correctness scores among 32 patients.

Our results with specificity maximization provided good performance especially in terms of specificity. Thus, this method can successfully predict non-DMR patients from short-term measurements. In future studies analyzing a larger dataset, we plan to employ machine learning methods for the optimization of our framework. The prediction of characteristic features of CML patients may also represent a future research avenue. In Fig. 3c, one can find that some patients were outliers, i.e., they were separated from most DMR and most non-DMR patients. Further clinical analysis on these outliers may provide new characteristic features of CML patients. As an example of related works, M. Horn et al. showed that patient classification on parameter space

predicted relapse²³. Thus, prediction of relapse may be another possible application of our mathematical framework.

DATA AVAILABILITY

All data analyzed in this study are available in Supplementary Materials.

CODE AVAILABILITY

Code used in this study is available at the following link: <https://github.com/CMLDataScience/CML-DMR-Prediction>.

Received: 22 February 2022; Accepted: 16 September 2022;

Published online: 13 October 2022

REFERENCES

- Shtivelman, E., Lifshitz, B., Gale, R. P. & Canaani, E. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. *Nature* **315**, 550–554 (1985).
- Melo, J. V. & Barnes, D. J. Chronic myeloid leukaemia as a model of disease evolution in human cancer. *Nat. Rev. Cancer* **7**, 441–453 (2007).
- Druker, B. J. et al. Five-year follow-up of patients receiving Imatinib for chronic myeloid leukemia. *N. Engl. J. Med.* **355**, 2408–2417 (2006).
- Talpaz, M. et al. Dasatinib in Imatinib-resistant philadelphia chromosome-positive leukemias. *N. Engl. J. Med.* **354**, 2531–2541 (2006).
- Saglio, G. et al. Nilotinib versus Imatinib for newly diagnosed chronic myeloid leukemia. *N. Engl. J. Med.* **362**, 2251–2259 (2010).
- Kantarjian, H. et al. Nilotinib in Imatinib-resistant CML and philadelphia chromosome-positive ALL. *N. Engl. J. Med.* **354**, 2542–2551 (2006).
- Kantarjian, H. M. et al. Nilotinib (formerly AMN107), a highly selective BCR-ABL tyrosine kinase inhibitor, is effective in patients with philadelphia chromosome-positive chronic myelogenous leukemia in chronic phase following imatinib resistance and intolerance. *Blood* **110**, 3540–3546 (2007).
- Cortes, J. E. et al. Bosutinib versus Imatinib for newly diagnosed chronic myeloid leukemia: Results from the randomized BFORE trial. *J. Clin. Oncol.* **36**, 231–237 (2018).
- Cross, N. C. P., White, H. E., Müller, M. C., Saglio, G. & Hochhaus, A. Standardized definitions of molecular response in chronic myeloid leukemia. *Leukemia* **26**, 2172–2175 (2012).
- Hughes, T. et al. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood* **108**, 28–37 (2006).
- Branford, S. et al. Desirable performance characteristics for BCR-ABL measurement on an international reporting scale to allow consistent interpretation of individual patient response and comparison of response rates between clinical trials. *Blood* **112**, 3330–3338 (2008).
- Baccarani, M. et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood* **122**, 872–884 (2013).
- Hehlmann, R. et al. Deep molecular response is reached by the majority of patients treated with Imatinib, predicts survival, and is achieved more quickly by optimized high-dose Imatinib: results from the randomized CML-Study IV. *J. Clin. Oncol.* **32**, 415–423 (2014).
- Mahon, F.-X. & Etienne, G. Deep molecular response in chronic myeloid leukemia: the new goal of therapy? *Clin. Cancer Res* **20**, 310–322 (2014).
- Hochhaus, A. et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia* **34**, 966–984 (2020).
- Ross, D. M. et al. Long-term treatment-free remission of chronic myeloid leukemia with falling levels of residual leukemic cells. *Leukemia* **32**, 2572–2579 (2018).
- Nishiwaki, K. et al. Optimal treatment strategy with nilotinib for patients with newly diagnosed chronic-phase chronic myeloid leukemia based on early achievement of deep molecular response (MR^{4.5}): the phase 2, multicenter N-Road study. *Cancer Med.* **9**, 3742–3751 (2020).
- Sasaki, K. et al. Prediction for sustained deep molecular response of BCR-ABL1 levels in patients with chronic myeloid leukemia in chronic phase. *Cancer* **124**, 1160–1168 (2018).
- Hasford, J. et al. Predicting complete cytogenetic response and subsequent progression-free survival in 2060 patients with CML on imatinib treatment: the EUTOS score. *Blood* **118**, 686–692 (2011).
- Clapp, G. & Levy, D. A review of mathematical models for leukemia and lymphoma. *Drug Discov. Today. Dis. Models* **16**, 1–6 (2015).

- Michor, F. et al. Dynamics of chronic myeloid leukaemia. *Nature* **435**, 1267–1270 (2005).
- Tang, M. et al. Dynamics of chronic myeloid leukemia response to long-term targeted therapy reveal treatment effects on leukemic stem cells. *Blood* **118**, 1622–1631 (2011).
- Horn, M. et al. Model-based decision rules reduce the risk of molecular relapse after cessation of tyrosine kinase inhibitor therapy in chronic myeloid leukemia. *Blood* **121**, 378–384 (2013).
- Roeder, I. et al. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nat. Med.* **12**, 1181–1184 (2006).
- Komarova, N. L. & Wodarz, D. Effect of cellular quiescence on the success of targeted CML therapy. *PLoS One* **2**, e990 (2007).
- Ma, H., Leng, S., Aihara, K., Lin, W. & Chen, L. Randomly distributed embedding making short-term high-dimensional data predictable. *Proc. Natl Acad. Sci. USA* **115**, E9994–E10002 (2018).
- Hirata, Y. et al. Intermittent androgen suppression: Estimating parameters for individual patients based on initial PSA data in response to androgen deprivation therapy. *PLoS One* **10**, e0130372 (2015).
- Morino, K. et al. Predicting disease progression from short biomarker series using expert advice algorithm. *Sci. Rep.* **5**, 8953 (2015).
- Mustjoki, S. et al. Rapid mobilization of cytotoxic lymphocytes induced by dasatinib therapy. *Leukemia* **27**, 914–924 (2013).
- Mustjoki, S. et al. Clonal expansion of T/NK-cells during tyrosine kinase inhibitor dasatinib therapy. *Leukemia* **23**, 1398–1405 (2009).
- Hochhaus, A. et al. Treatment-free remission following frontline nilotinib in patients with chronic myeloid leukemia in chronic phase: results from the ENESTFreedom study. *Leukemia* **31**, 1525–1531 (2017).
- Baccarani, M. et al. Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *J. Clin. Oncol.* **27**, 6041–6051 (2009).

ACKNOWLEDGEMENTS

This work was partially supported by JSPS KAKENHI Grant Number JP15H05707 (to K.M., M.K.K., S.N., and K.A.), JST PRESTO Grant Number JPMJPR16E9 (to S.N.), JST Moonshot R&D Grant Number JPMJMS2021 (to K.A.), AMED under Grant Number JP22dm0307009 (to K.A.) and Institute of AI and Beyond of The University of Tokyo (to K.A.). The N-road study had been funded by Novartis Pharma K.K.

AUTHOR CONTRIBUTIONS

M.H. and A.T. designed the project from the clinical viewpoint. Y.O., K.M., M.K.K., S.N., and K.A. designed the mathematical analyses of this project. Y.O., K.M., M.K.K., S.N., and K.A. created the theoretical method. Y.O., K.M., M.K.K., and S.N. analyzed the data. Y.O. and M.K.K. performed numerical analyses. M.H., K.N., and A.T. obtained the clinical data from the N-road and proposed the clinical implementations. Y.O., M.H., K.M., M.K.K., S.N., A.T., and K.A. wrote the manuscript. K.N., H.W. and A.T. designed the N-road study. K.S., S.T., J.H., H.Y., T.I., A.S., T.S., S.H., K.F., S.S., T.Y., H.W., A.T., and K.N. collected the clinical data in individual patients. M.T. performed quantitative analysis of Ph⁺ stem and progenitor cells in each patient. All authors checked the manuscript and agreed to submit its final version.

COMPETING INTERESTS

K.M. reports a grant from NEC outside the submitted work. S.N. reports a grant from Morinaga Milk Industry Co., Ltd. outside the submitted work. K.N. received a grant for this work from Novartis Pharma K.K. and reports grants from Zenyaku Kogyo Company, Ltd., Asahi Kasei Pharma, and Taiho Pharmaceutical Co., Ltd.; grants and personal fees from Chugai Pharmaceutical, Kyowa Hako Kirin Co., Ltd., Nippon Shinyaku Co., Ltd., Mochida Pharmaceutical Co., Ltd., Ono Pharmaceutical Co., Ltd., Takeda Pharmaceutical Co., Ltd., and Sumitomo Dainippon Pharma Co., Ltd.; and personal fees from Pfizer, Otsuka Pharmaceutical Co., Ltd., Janssen Pharmaceutical K.K., Eisai Co., Ltd., and Celgene K.K. outside of the submitted work. H.Y. reports grants from Celgene K.K. and Astellas Pharma Inc. outside the submitted work. A.T. reports personal fees from Sysmex, Otsuka Pharmaceutical Co., Ltd., Bristol-Myers Squibb, and Takeda Pharmaceutical Co., Ltd., Daiichi-Sankyo and grants and personal fees from Pfizer, and Chugai Pharmaceutical outside the submitted work. K.A. reports a grant and personal fees from KKE, a grant and personal fees from NEC, a grant from Sysmex, a personal fee from Novo Nordisk Japan, and a grant and personal fees from Toyota Central R&D Labs. outside the submitted work. Y.O. had worked at Sysmex Co., but the entire results of this paper are totally independent from his past job in Sysmex Co. The other authors have nothing to disclose.

ETHICS

This research has been approved by the Institutional Review Boards at The University of Tokyo, at The Institute of Medical Science, The University of Tokyo (IMSUT), at Kyushu University, at University of Fukui, and at Hokkaido University, and conducted in accordance with the ethical guidelines for medical and health research involving human subjects. This research also adhered to the ethical principles of the Declaration of Helsinki. All patients provided written informed consent for participation prior to initiation of any study procedures of the N-road study ¹⁷.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-022-00248-3>.

Correspondence and requests for materials should be addressed to Masashi K. Kajita or Kazuyuki Aihara.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022