Drug product attributes predict clinical efficacy in betibeglogene autotemcel gene therapy for β-thalassemia

Dustin Whitney,[1,3,4] Ilya Shestopalov,[1,4] Maeva Fincker,[1,2] Marc d'Anjou,[1] Kelly Kral,[1] Marisa Gayron,[1] Francis J. Pierciey,[1] and Richard A. Colvin[1]

[1]bluebird bio, Inc., Somerville, MA 02145, USA

Ex vivo autologous hematopoietic stem cell lentiviral-based gene therapy with betibeglogene autotemcel has been studied in patients with transfusion-dependent β-thalassemia in Phase III clinical trials (HGB-207 and HGB-212), with 90% of patients reaching transfusion independence (TI). Here, we explore manufacturing parameters, drug product quality attributes, and limited patient characteristics that had an impact on clinical efficacy in HGB-207 and HGB-212. Retrospective analysis revealed that the peripheral blood vector copy number (VCN) was related to TI, with a strong correlation between peripheral blood VCN at 6 months and gene therapy–derived therapeutic protein (HbA^{T87Q}) expression at 6 months (correlation coefficient, 0.8681; p < 0.0001; R^2 = 0.7536). A peripheral blood VCN threshold of ≥0.75 copies per diploid genome at 6 months post betibeglogene autotemcel infusion provided a stringent surrogate biomarker for TI and was used as the outcome variable for multivariate analysis using a random forest classifier. The top predictive feature of clinical efficacy was found to be the percentage of lentiviral vector–positive cells in the drug product. This retrospective analysis is critical to understanding the key product quality attributes that can predict clinical efficacy in lentiviral vector gene therapy within this clinical trial population.

INTRODUCTION

Transfusion-dependent β-thalassemia (TDT) is a severe, rare genetic disease characterized by impaired production of the β-globin chain of adult hemoglobin (Hb). This results in a deficiency of functional red blood cells (RBCs), chronic anemia, serious complications, and a shortened lifespan compared with the general population.[1,2] Patients undergo regular, lifelong blood transfusions and iron chelation; however, transfusions are burdensome, and many patients continue to experience TDT-related comorbidities due to iron overload despite improvements in care.[1,2]

Gene therapy provides a potentially curative treatment option for patients with TDT.[3] *Ex vivo* lentiviral gene therapy with betibeglogene autotemcel (beti-cel; also known as Zynteglo) integrates a modified *HBB* gene (β^{A-T87Q}-globin) into the hematopoietic stem cells

(HSCs) of patients with TDT, aiming to enable lifelong, stable production of functional adult Hb and transfusion independence (TI).[4,5] Gene therapy with beti-cel showed positive results in patients with TDT in Phase I/II[4] and Phase III[5,6] studies. In 2022, the US Food and Drug Administration approved Zynteglo as the first cell-based gene therapy to treat adult and pediatric patients with β-thalassemia who require regular blood transfusions.

A total of 41 patients were treated with beti-cel in the Phase III clinical trials HGB-207 and HGB-212. As of July 2022, 37 of 41 (90%) patients treated in HGB-207 and HGB-212 had achieved TI (defined as a weighted average Hb level ≥9 g/dL, starting 60 days after the last transfusion in patients who had not received RBC transfusions ≥12 months), including 21 of 23 (91%) patients with non-β^0/β^0 mutations in HGB-207,[5] and 16 of 18 (89%) patients with β^0/β^0, IVS-I-110 homozygous, or IVS-I-110/β^0 mutations in the HGB-212 trial.

The aim of this study was to explore the impact of manufacturing parameters, drug product quality attributes, and limited patient characteristics on clinical efficacy in the HGB-207 and HGB-212 Phase III clinical trials using a retrospective multivariate analysis (MVA). The MVA was performed with an earlier dataset (August 2021) for 37 subjects, in which 31 reached TI, 4 did not reach TI, and 2 were not evaluable for TI status (Table S1). The patient characteristics chosen were limited to readily available attributes at the time of the analysis: subject age, disease genotype, clinical site, and splenectomy status. A comprehensive analysis of patient characteristics and their relationship to efficacy of beti-cel is a separate study and outside the scope of this paper.
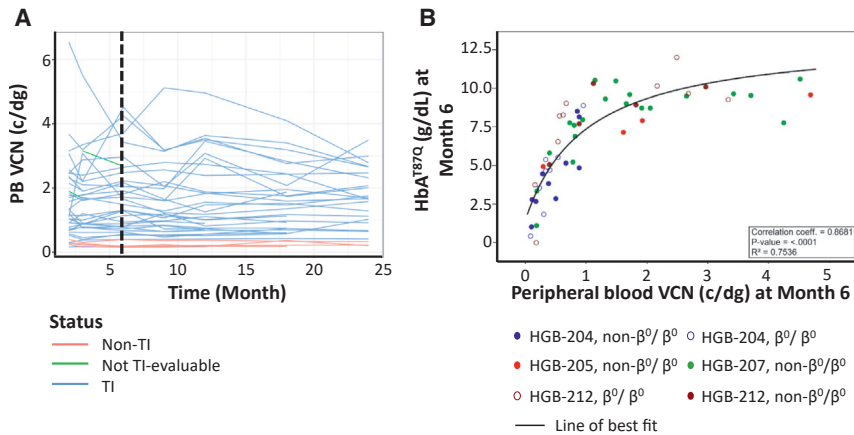
Figure 1. PB VCN at month 6 is an early indicator of TI
(A) PB VCN and TI are closely related. (B) Correlation between PB VCN at month 6 and HbA^{T87Q} concentration at month 6.

tients treated in the HGB-207 and HGB-212 Phase III studies of beti-cel were included in the analysis (Table S1). The remaining patients were either lacking some of the 27 variables or their clinical follow-up was <6 months at the time of this analysis.

The top predictive features were found to be drug product quality attributes and not manufacturing parameters or limited patient characteristics. Specifically, all of the predictive features were correlated measures of transduction efficiency (Figure 2A). These included the percentage of transduced lentiviral vector–positive (LVV+) cells in the drug product (%LVV+ cells), percentage of LVV+ burst-forming unit-erythroid (BFU-E) colonies, mean VCN among BFU-E colonies, VCN from cytokine cell culture, and VCN from pooled colony-forming cells. Figure 2A shows that the %LVV+ cells are the most predictive attribute of PB VCN at month 6 and subsequently TI, with higher %LVV+ cells being positively associated with PB VCN and, thus, ultimately a surrogate for TI. All TI-evaluable patients, regardless of genotype, with a PB VCN ≥0.75 c/dg at 6 months achieved TI (Figure 2B). All of the patients who did not achieve TI had PB VCN levels <0.75 c/dg at 6 months.

Many features did not predict clinical efficacy as defined by the above-established PB VCN biomarker, despite their importance in determining the success of allogeneic HSC transplantation (allo-HSCT). The minimum cell dose for the HGB-207 and HGB-212 studies was 5×10^6 cells/kg; however, the total cell dose (defined in Table S2) was not found to be predictive of achieving PB VCN ≥0.75 c/dg at 6 months. The dose of CD34$^+$ subpopulations, such as CD90$^+$CD133$^+$ HSCs and CD90$^−$CD133$^+$ cells that include multipotent progenitors (MPPs), was not found to be predictive.

RESULTS

As of July 2022, all 41 patients were evaluable for TI, and 37 (90%) achieved TI. A post hoc MVA was used to identify manufacturing parameters, drug product quality attributes, and limited patient characteristics that determined whether a patient would reach TI. However, given that only 4 of 41 patients within the cohort did not reach TI, TI itself could not be used as a biomarker of clinical efficacy in our model.

The biomarker peripheral blood vector copy number (PB VCN), which largely stabilizes after 6 months, was determined to be closely related to TI (Figure 1A). PB VCN at 6 months post-beti-cel infusion was, therefore, explored for its potential to serve as a surrogate clinical biomarker and early predictor of the likelihood of achieving therapeutic levels of HbA^{T87Q} (and consequently TI) by 12 months in a larger cohort of 63 TDT patients treated with beti-cel during overall clinical development, including Phase I/II trials HGB-204 and HGB-205 and Phase III trials HGB-207 and HGB-212. In these patients, a linear fit of PB VCN/(1 + PB VCN) revealed a strong correlation between PB VCN and HbA^{T87Q} expression in PB at 6 months post treatment (correlation coefficient, 0.8681; $p < 0.0001$; $R^2 = 0.7536$; Figure 1B). Separate analysis of data from non-β^0/β^0 patients showed only that patients with this genotype produced, on average, 2.5 g/dL endogenous Hb. Therefore, patients who achieved at least 6.5 g/dL of HbA^{T87Q} expression in PB at 6 months went on to reach sum total Hb levels of ≥9 g/dL (i.e., the Hb level that is necessary for achieving TI). Using the correlation in Figure 1B, a value of 6.5 g/dL HbA^{T87Q} expression in PB was equivalent to a PB VCN of 0.75 copies per diploid genome (c/dg). Patients with PB VCN <0.75 c/dg at 6 months had variable outcomes, with only some achieving TI. Thus, a PB VCN threshold of ≥0.75 c/dg at 6 months post-beti-cel infusion provided a stringent surrogate biomarker for TI in both non-β^0/β^0 and β^0/β^0 patients and was used as the outcome variable for the subsequent MVA using a random forest classifier.

Variables (n = 27) consisting of manufacturing parameters, drug product quality attributes, and limited clinical characteristics (Table S2) were included as features in the random forest classifier to assess which features best predicted clinical outcome, defined by PB VCN ≥0.75 c/dg at 6 months post infusion. A cohort of 37 pa-
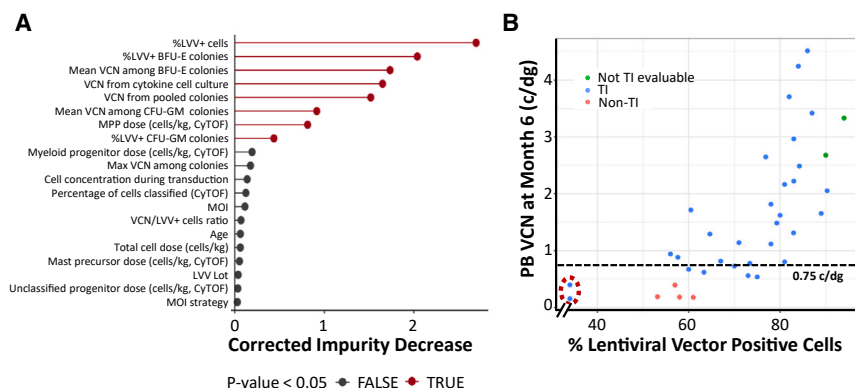
DISCUSSION

In the Phase I/II studies, HGB-204 and HGB-205, 12 of 14 patients aged 12–35 years with TDT and a non-β^0/β^0 genotype achieved TI after infusion of beti-cel.[4] However, only 3 of 8 patients with β^0/β^0 genotypes achieved TI in HGB-204. Based on these data, we hypothesized that higher transduction rates may be required for patients with β^0/β^0 genotypes to achieve TI, consistent with published *ex vivo* studies.[7] Therefore, process changes were implemented to improve transduction efficiency for clinical studies HGB-207 and HGB-212 and and future commercial manufacturing.[5] The aim of these process changes was to increase the percentage of CD34$^+$ cells in beti-cel that are transduced with LVV (%LVV+ cells). Consequently, the median VCN in beti-cel was higher in the HGB-207 and HGB-212 Phase III studies than in the Phase I/II studies, and

Figure 2. Top predictive features of clinical efficacy are drug product quality attributes

(A) Model performance and important variables from 50 resamples of the dataset. (B) %LVV+ cells are a key attribute that predicts PB VCN at month 6. The dotted black line represents a PB VCN of 0.75 c/dg. The red dotted circle highlights 2 patients who achieved TI with low %LVV+ cells in the drug product and a PB VCN at 6 months <0.75 c/dg. CFU-GM, colony-forming unit-granulocyte-macrophage; CyTOF, cytometry time of flight.

resulted in higher VCN in PB mononuclear cells, higher Hb levels, and a higher incidence of TI than in the previous studies.

In this analysis, we found a positive association between %LVV+ cells in the drug product and PB VCN at 6 months in subjects from the Phase III studies HGB-207 and HGB-212 (Figure 2B). Moreover, with the exception of 2 nonevaluable patients at the time of this analysis (August 2021 dataset, Table S1), all of the patients with a PB VCN at 6 months of ≥0.75 c/dg achieved TI (Figure 2B), indicating that PB VCNs >0.75 c/dg are sufficient for clinical efficacy in this population. Increasing PB VCN up to 3 c/dg further increases HbA[T87Q] expression (Figure 1B), which is beneficial to patients beyond reaching the minimum requirement of TI. However, there is a strong negative correlation between endogenous Hb and HbA[T87Q] production, as demonstrated at months 6 and 12 (month 6: coefficient -0.79, $p < 0.0001$; month 12: coefficient -0.77, $p < 0.0001$; Table S3), but endogenous Hb appears to have no statistically significant impact on total Hb production either *in vitro* or in subjects in HGB-207 and HGB-212 (Figure S1; Table S3). Of note, Figure 2B shows that two patients with low %LVV+ cells in the drug product (both at 34%) and a PB VCN <0.75 c/dg at 6 months still reached TI. These patients achieved TI despite the low %LVV+ cells in the drug product because they had relatively high endogenous Hb levels. Therefore, although it is possible to achieve TI with PB VCN <0.75 c/dg at 6 months, exceeding this stringent biomarker threshold maximizes the odds of achieving TI. In addition, we confirmed our hypothesis that PB VCN levels are not influenced by the β-thalassemia genotype, supporting the use of this parameter as a surrogate clinical biomarker to represent drug product quality in our MVA model.

In 37 patients from the Phase III trials of beti-cel, HGB-207 and HGB-212, the attribute %LVV+ cells in the drug product were found to be the top predictor of achieving a PB VCN ≥0.75 c/dg, a stringent surrogate biomarker for TI. It is notable that a single drug product attribute measured before infusion can predict clinical efficacy despite the inherent physiological variability among the 37 subjects who met the HGB-207/HGB-212 clinical trial inclusion criteria,[5] broadly representing the real-world population that may benefit from treatment with beti-cel.

Additional analysis of Phase III data showed that there was a relationship between HbA[T87Q] expression at month 24 and the %LVV+ cells in the drug product (Figure 3B). Moreover, a threshold of %LVV+ cells was identified that could discriminate subjects who achieved TI from subjects who had variable outcomes (Figure 3A). This threshold was used to establish prospective acceptance criteria for future release and clinical use of beti-cel, thus providing increased assurance around the quality of the drug product and maximizing the chances of achieving TI with beti-cel treatment. A robust and defined manufacturing process was established that results in consistently efficient transduction such that CD34[+] cells from all of the subjects met the %LVV+ threshold necessary to achieve TI, regardless of donor–donor propensity to transduction with BB305 LVV. As a consequence of this manufacturing strategy, a substantial number of drug products end up with >80% %LVV+ cells, resulting in a PB VCN of >2 (Figure 2B).

Our study indicates that increasing %LVV+ cells affords a quantifiable benefit for achieving TI, but it also inevitably increases drug product VCN (Figure S2), and as a result, PB VCN therefore increases the hypothetical risk of insertional oncogenesis following beti-cel therapy.[8,9] The primary safety concern for LVV-transduced HSCs is the potential occurrence of LVV-mediated insertional oncogenesis, a form of genotoxicity. However, current LVVs used for gene therapy have inherent and engineered properties that limit this risk.[10] The BB305 LVV used in beti-cel trials is self-inactivating, with diminished potential for influencing the expression of nearby genes by the removal of viral enhancer and promoter sequences, and has a preference for integrations within introns.[11–14] Expression of the BB305 transgene is controlled by a tissue-specific promoter that limits expression to erythroid lineage cells.[4,5] Furthermore, a ratio of VCN to %LVV+ cells can be calculated before drug product infusion, and upper limits for drug product VCN and VCN/LVV+ cells can be used to control for disproportional transduction, where VCN increases without the beneficial increase to %LVV+ cells (Figure S2). With these design elements and controls in place, no cases of insertional oncogenesis or malignancy have been observed with the BB305 LVV to date (as of February 2023: 122 patients infused with BB305 LVV-derived drug products, 534 total patient-years of

A

B

- HGB-207, non-β^0/β^0 ○ HGB-212, β^0/β^0 ● HGB-212, non-β^0/β^0
— Line of best fit
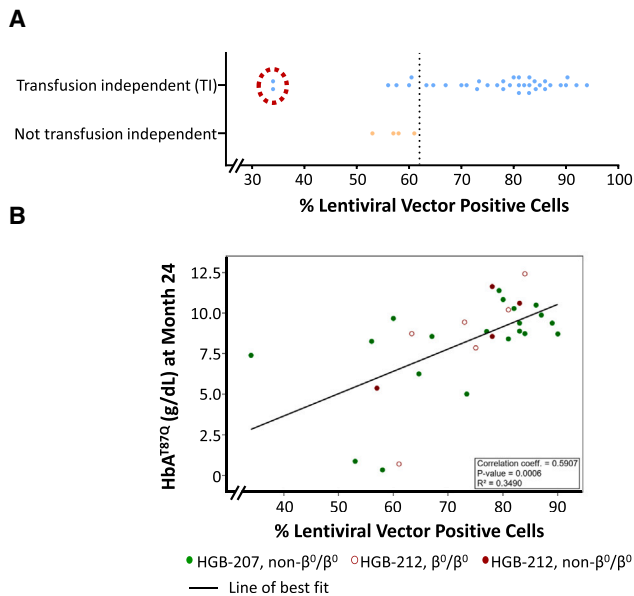
Figure 3. Relationship between % LVV+ cells in drug product and TI or HbA^{T87Q} concentration

(A) All subjects infused with drug products above the identified %LVV+ cell threshold achieved TI. (B) Correlation with HbA^{T87Q}, in g/dL, expression at month 24 versus %LVV+ cells in drug product. Phase III (HGB-207 and HGB-212) data are included. The red dotted circle highlights 2 patients who achieved TI with PB VCN at 6 months <0.75 c/dg (same 2 patients as in Figure 2).

follow-up). In cases in which oligoclonality was observed in a separate clinical study using a different vector design for the expression of β^{A-T87Q}-globin,[15] the abundance of the HMGA2-driven clone diminished over time[16] and has not led to oncogenesis. Thus, the risk of insertional oncogenesis related to higher %LVV+ cells that lead to higher drug product VCNs, higher PB VCNs, and increased odds of achieving TI remains only a hypothetical and nonquantifiable risk associated with beti-cel.

Allo-HSCT is a potentially curative treatment option for patients with TDT, but it is limited by donor availability and patient age and is associated with transplant-related complications.[1,3] The identification of new biomarkers and predictors for the clinical efficacy of autologous gene therapies is, therefore, important, because their availability could contribute to the development of therapeutic and prognostic strategies. Moreover, predictors of clinical efficacy help to determine the threshold for optimal clinical outcomes and generate improved clinical practices for the minimization of adverse effects. Lastly, they may support the establishment of product quality acceptance criteria for gene therapies, as demonstrated in Figure 3. In this context, our findings have a broader relevance and may be applicable to lentiviral gene therapy approaches in other disease areas.

Limitations of the study

This study focused on manufacturing parameters and product attributes and not on many possible clinical characteristics. For example,

it remains to be seen whether patient baseline iron levels and ineffective erythropoiesis are predictors of success in achieving TI with beti-cel therapy. This analysis is also limited to a post hoc analysis based on subjects who met the inclusion criteria for HGB-207 and HGB-212, and the findings may not be applicable in patients who do not meet the inclusion criteria. The identified threshold %LVV+ cells needed to reach TI is specific to the LVV construct and manufacturing process used for beti-cel, and as such the threshold should be reevaluated for other LVV constructs and clinical studies. Finally, the mechanism governing the plateau in HbA^{T87Q} levels observed with increasing VCN has not been fully investigated.

MATERIALS AND METHODS
Patients
A total of 41 patients with TDT was treated with beti-cel in the HGB-207 (n = 23) and HGB-212 (n = 18) single-group, open-label, single-dose, Phase III clinical trials. Patients in the HGB-207 study had non-β^0/β^0 genotypes, whereas patients in the HGB-212 study had either β^0/β^0 genotypes or IVS-I-110 homozygous or IVS-I-110/β^0 mutations, which were considered in these studies to be equivalent to a β^0 mutation.

The beti-cel drug product was prepared as previously described,[5] and the manufacturing process is shown in Figure 4. In brief, HSCs were mobilized with granulocyte colony-stimulating factor plus plerixafor and then harvested by apheresis. The target minimum number of CD34$^+$ HSCs and progenitor cells to be collected for manufacturing the drug product was $\geq 12 \times 10^6$ CD34$^+$ cells/kg. The collection of sufficient cells for drug product manufacture and backup cells was achieved through 1 cycle of mobilization/apheresis for most patients (78% of patients underwent 1 cycle). CD34$^+$ cells were enriched, stimulated, and transduced ex vivo with BB305 LVV to create the beti-cel drug product. The BB305 LVV is a replication-defective, self-inactivating, third-generation HIV-1–based LVV carrying the human β-globin gene with a single modification at codon 87 ($\beta^{A87\ Thr:Gln}$ [β^{A-T87Q}]-globin).[5]

The presence of vector sequences in the genomic DNA (gDNA) of differentiated cells was detected using qPCR, with results expressed as VCN with units of vector copies per diploid genome. Drug product VCN was measured in gDNA from pooled colony-forming cells after in vitro culture in methylcellulose, and PB VCN was measured in gDNA from PB cells. The %LVV+ cells was measured by PCR using individual flow-sorted cells from the drug product after cell culture in cytokine media. Drug product VCN and %LVV+ cell assays are performed on every drug product lot, and the drug product must meet acceptance criteria for these parameters before the patient undergoes the conditioning regimen for beti-cel infusion. Details on cellular phenotypes included in this study are in Table S4.

The primary endpoint of both HGB-207 and HGB-212 was TI. TI was defined as a weighted average Hb level of at least 9 g/dL starting 60 days after the last transfusion in patients who had not received RBC transfusions for ≥ 12 months. Of the 41 patients who received
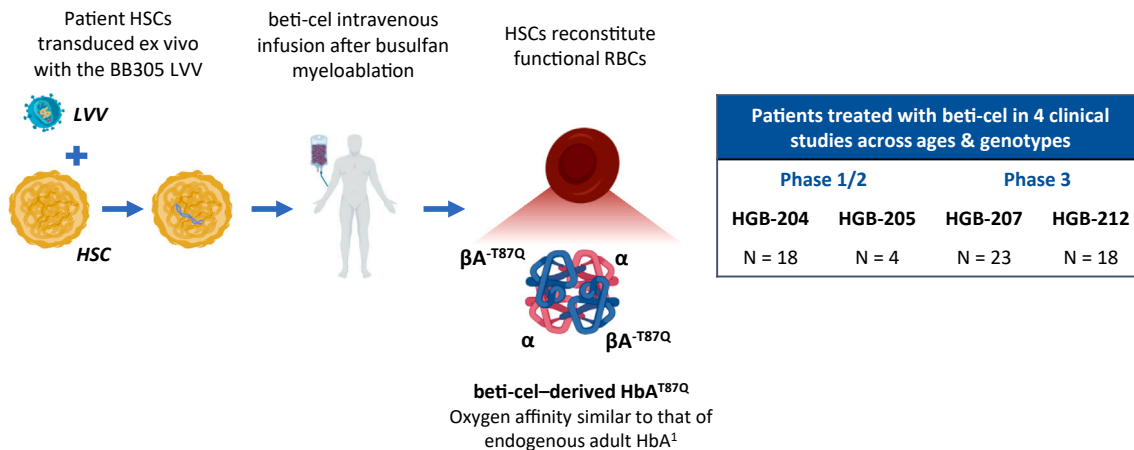
Figure 4. beti-cel gene therapy manufacturing process

beti-cel, 38 patients were evaluable for TI at the time of this analysis (August 2021).

Model description

The retrospective Phase III clinical trial data were mined for variables that could serve as a surrogate biomarker for TI status, to provide an output feature that could be used in the MVA model to extract predictors of postinfusion outcomes. Graphical analysis revealed that PB VCN at 6 months was closely related to TI. A linear fit of PB VCN/(1 + PB VCN) was applied to calculate the correlation coefficient between HbA^{T87Q} levels and the surrogate biomarker PB VCN at 6 months. $p \leq 0.05$ was regarded as statistically significant.

The prediction model for TI was constructed using a random forest classifier, with PB VCN at 6 months as the output variable. A total of 27 variables consisting of limited clinical characteristics, manufacturing parameters, and drug product characteristics (Table S2) were available for 37 of 38 patients in the studied cohort at the time of analysis (August 2021; Table S1). These variables were input as the covariates of the prediction model (Figure 5). Other variables, such as baseline iron levels and markers of ineffective erythropoiesis, were not investigated in the model because the full numerical dataset was not available at the time of the analysis for all of the patients. The random forest is a popular tool for classification and regression, which shows a powerful ability to construct a predictive model for new biomarkers. The random forest is less prone to overfitting problems and can handle a large amount of noise.[17]

A random forest model was trained to categorize patients reaching PB VCN at 6 months ≥ 0.75 c/dg or not, using the ranger and tidymodels R packages[18,19] and the 27 variables described above as model input. Collinearity of the variables was assessed with Spearman rank correlation, and a representative variable for groups of highly correlated variables (rho >0.8) was automatically selected if necessary. The complete dataset (n = 37 patients; Table S1) was split randomly 80:20 into training and validation datasets. Upsampling strategy of the lowest abundance class was used during training to combat class imbalance but not model performance assessment or variable importance measurement. Hyperparameters were tuned using a 10-fold cross-validation strategy, and the parameters leading to the highest cross-validated receiver operating characteristic (ROC) area under the curve (AUC) were selected. The final model performance was assessed by calculating accuracy, F1 measure, ROC AUC, precision, recall, and negative predictive value on the validation set as well as on 50 resamplings of the full dataset (50 random versions of the validation
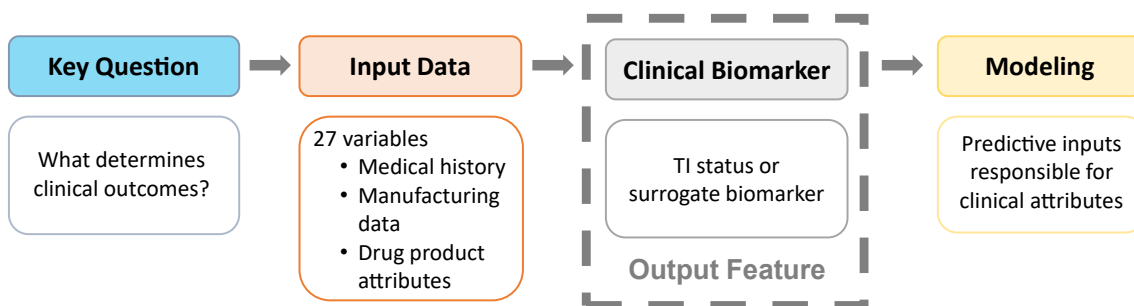

Figure 5. Model description

dataset). Similarly, the corrected impurity across validation and re-sampled sets was measured to identify variables that best predicted TI status. $p \leq 0.05$ was considered significant.

DATA AND CODE AVAILABILITY

Appropriately deidentified patient-level datasets and supporting documents may be shared following the attainment of applicable marketing approvals and consistent with criteria established by bluebird bio and/or industry best practices to maintain the privacy of study participants. For more information, please contact datasharing@bluebirdbio.com.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2023.101155.

REFERENCES

1. Cappellini, M.D., Farmakis, D., Porter, J., and Taher, A. (2021). 2021 Guidelines for the Management of Transfusion Dependent Thalassemia ((TDT) (Thalassemia International Federation)).

2. Galanello, R., and Origa, R. (2010). Beta-thalassemia. Orphanet J. Rare Dis. 5, 11–15.

3. Dong, A.C., and Rivella, S. (2017). Gene Addition Strategies for β-Thalassemia and Sickle Cell Anemia. In Gene and Cell Therapies for Beta-Globinopathies Advances in Experimental Medicine and Biology, P. Malik and J. Tisdale, eds. (Springer New York), pp. 155–176.

4. Thompson, A.A., Walters, M.C., Kwiatkowski, J., Rasko, J.E.J., Ribeil, J.-A., Hongeng, S., Magrin, E., Schiller, G.J., Payen, E., Semeraro, M., et al. (2018). Gene Therapy in Patients with Transfusion-Dependent β-Thalassemia. N. Engl. J. Med. 378, 1479–1493.

5. Locatelli, F., Thompson, A.A., Kwiatkowski, J.L., Porter, J.B., Thrasher, A.J., Hongeng, S., Sauer, M.G., Thuret, I., Lal, A., Algeri, M., et al. (2022). Betibeglogene Autotemcel Gene Therapy for Non–β0/β0 Genotype β-Thalassemia. N. Engl. J. Med. 386, 415–427.

6. Kwiatkowski, J.L., Locatelli, F., Walters, M.C., Porter, J.B., Hongeng, S., Yannaki, E., Kulozik, A.E., Sauer, M.G., Thrasher, A.J., Thuret, I., et al. (2021). Improvement in Health-Related Quality of Life Following Treatment with Betibeglogene Autotemcel in Patients with Transfusion-Dependent β-Thalassemia Enrolled in Phase 3 Studies. Blood 138, 3085.

7. Breda, L., Casu, C., Gardenghi, S., Bianchi, N., Cartegni, L., Narla, M., Yazdanbakhsh, K., Musso, M., Manwani, D., Little, J., et al. (2012). Therapeutic hemoglobin levels after gene transfer in β-thalassemia mice and in hematopoietic cells of β-thalassemia and sickle cells disease patients. PLoS One 7, e32345.

8. Baum, C., Kustikova, O., Modlich, U., Li, Z., and Fehse, B. (2006). Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. Hum. Gene Ther. 17, 253–263.

9. Sadelain, M. (2004). Insertional oncogenesis in gene therapy: how much of a risk? Gene Ther. 11, 569–573.

10. Tucci, F., Galimberti, S., Naldini, L., Valsecchi, M.G., and Aiuti, A. (2022). A systematic review and meta-analysis of gene therapy with hematopoietic stem and progenitor cells for monogenic disorders. Nat. Commun. 13, 1315.

11. Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L., and Trono, D. (1998). Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. J. Virol. 72, 9873–9880.

12. Schröder, A.R.W., Shinn, P., Chen, H., Berry, C., Ecker, J.R., and Bushman, F. (2002). HIV-1 Integration in the Human Genome Favors Active Genes and Local Hotspots. Cell 110, 521–529.

13. Wu, X., Li, Y., Crise, B., and Burgess, S.M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. Science 300, 1749–1751.

14. Yu, S.F., von Rüden, T., Kantoff, P.W., Garber, C., Seiberg, M., Rüther, U., Anderson, W.F., Wagner, E.F., and Gilboa, E. (1986). Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. Proc. Natl. Acad. Sci. USA 83, 3194–3198.

15. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature 467, 318–322.

16. Payen, E., and Leboulch, P. (2012). Advances in stem cell transplantation and gene therapy in the β-hemoglobinopathies. Hematology. Am. Soc. Hematol. Educ. Program 2012, 276–283.

17. Breiman, L. (2001). Random Forests. Mach. Learn. 45, 5–32.

18. Wickham, H., Acerick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., Francois, R., Grolemun, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. 4, 1–6.

19. Wright, M.N., and Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J. Stat. Software 77, 1–17.