IN THE SPOTLIGHT

Rare Sequences Make Sense of CAR T-cell Therapy Outcomes

Sara Ghorashian[1,2] and Jack Bartram[1,2]

Summary: In this issue, Pulsipher and colleagues used next-generation sequencing to detect leukemia-specific sequences following tisagenlecleucel therapy of acute lymphoblastic leukemia. A challenge for the field currently is to identify which patients will have therapy failure and to do so early enough to allow planning for further treatment, for example, stem cell transplantation. Detection of disease below the standard detection level for this technique (less than one per million cells) at day 28 was associated with poorer outcomes and potentially therefore could be used to identify those that might benefit from adjunctive therapies.

See related article by Pulsipher et al., p. 66 (9).

CD19-targeted chimeric antigen receptor (CAR) T-cell therapy provides a therapy option with curative potential for those with advanced, relapsed/refractory acute lymphoblastic leukemia (ALL). The pivotal ELIANA study documented a 12-month event-free survival (EFS) of 50% for recipients of tisagenlecleucel infusion (1). This and subsequent reports noted that responses were durable, providing proof of concept that CAR T cells could be delivered as a stand-alone therapy. Further, outcomes were not impacted by cytogenetic risk, prior stem cell transplantation (SCT), and disease burden (1).

A deeper dive into the data is needed to understand the true incidence of therapy failure following tisagenlecleucel infusion. In the ELIANA study, 15 of 61 responding patients were censored for EFS because they went on to receive further therapy while in remission. This was SCT in eight patients and other cancer therapy in seven patients, presumably either for the loss of CAR T-cell persistence or emergence of minimal/measurable residual disease (MRD)–level disease. While registry data have generally confirmed that outcomes from tisagenlecleucel delivery in the real world remain as successful as in the ELIANA study (2), it is clear that when taking into account those that require further therapy in remission due to emergence of MRD or early B-cell recovery, the 12-month EFS may be significantly lower.

A key challenge for treating clinicians is the ability to identify those who will have therapy failure early on after CAR T-cell infusion. This allows for timely planning of adjunctive therapies such as SCT, a successful strategy following therapy of pediatric ALL with short-persisting CD19 CAR T cells (3). An alternative strategy would be to deliver adjunctive SCT to all responders, but this comes at the cost of toxicity, especially when undertaken as a second procedure, including the impact on fertility and late effects, as well as a significant health resource allocation for such costly sequential therapies. However, validated predictive biomarkers of subsequent therapy failure in responding patients have been lacking to date.

Tests for CAR T-cell persistence are useful in assessing the risk of relapse, as this is known to be higher where CAR T-cell persistence is short, and certainly where the duration is less than 3 months (4). Standardized, validated assays detecting CAR T cells are lacking outside the context of clinical studies, and most centers therefore use absence of significant populations of B cells in the peripheral blood and bone marrow (B-cell aplasia) to infer persistence of CAR T cells targeting CD19. However, there is a need for a consensus definition of B-cell aplasia, and different thresholds have been used to date. Evolution of CD19-negative leukemia is a mechanism for relapse in the presence of CAR T-cell persistence and means that ongoing B-cell aplasia alone is insufficient to predict the likelihood of relapse.

Therefore, most centers also follow the assessment of disease MRD. To date, there have been two accepted methods: either multiparametric flow cytometry (MFC-MRD) with a quantitative range usually of 1 in 10,000 cells (10^{-4}) or, as has been established widely in Europe, allele-specific real-time quantitative PCR of leukemia-specific immune-receptor gene rearrangements [PCR-MRD, either immunoglobulin (IG) or T-cell receptor (TCR) sequences]. As depicted in Fig. 1, this is normally quantitative down to the level of 1 in 100,000 cells (10^{-5}) but only allows quantitation of 1 or 2 dominant leukemia clones with a patient-specific assay.

More recently, novel technologies such as highly parallel high-throughput sequencing of rearranged variable, diversity, and joining (VDJ) genes of the IG/TCR genes have been developed and are commonly termed next-generation sequencing (NGS)-MRD. This provides a single highly sensitive test applicable to almost all patients, quantifying sequences for all persisting or evolving leukemia clones with the ability to detect emerging clones that were initially at low frequency and would have been missed by biased sequencing assays
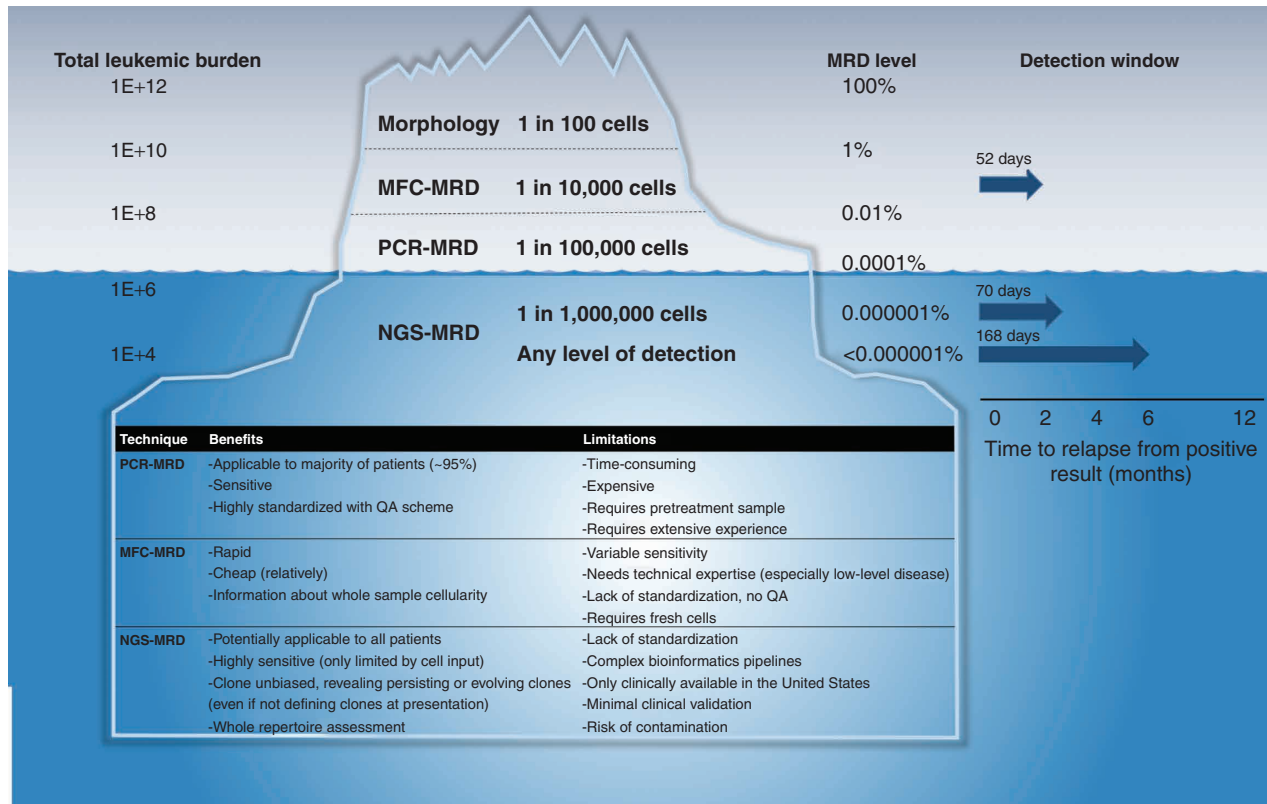
Figure 1. Comparison of relative sensitivities of MRD assessment methods and corresponding levels of disease. The left column refers to the estimated total leukemic cells (leukemic burden) in the body at a given MRD level. The median interval to relapse from detection of a positive MRD result, the "detection window" adapted from Pulsipher and colleagues (9), is shown in the far-right column for MFC-MRD, for NGS-MRD at a level of sensitivity cutoff at 10^{-6}, and for NGS-MRD at any detectable level. QA, quality assurance.

such as PCR-MRD. Unambiguous assignment of IG/TCR sequences to leukemia represents a bioinformatic challenge when tracking clones with low frequency, as it requires taking into account frequencies of a given clonal IG/TCR index sequence within healthy B- or T-cell populations. NGS-MRD can identify children with very low risk of relapse after induction chemotherapy and has been shown to be more specific for relapse prediction than PCR-MRD immediately after SCT (5, 6). With this approach, sensitivities of detection of one in a million cells (10^{-6}) and lower can be routinely achieved with an adequate input cell number, which is the main limitation of this approach (Fig. 1).

Recent real-world data of tisagenlecleucel therapy have provided information on factors other than MRD detection and loss of CAR T-cell persistence, which are associated with therapy failure. These include preinfusion disease burden (12-month EFS for high vs. low disease burden was 34 vs. 69%; ref. 2), prior blinatumomab therapy (median EFS was 5.8 vs. 22.6 months; ref. 7), being MRD positive by *IGH* PCR at day 28 (HR of 3.81 for cumulative incidence of relapse, CIR), as well as loss of B-cell aplasia (HR of 3.21 for CIR; ref. 8).

In the article by Pulsipher and colleagues (9), analysis of MRD by NGS and multiparametric flow cytometry data were compared in a retrospective analysis of 143 patients receiving tisagenlecleucel on the ELIANA and ENSIGN trials whose combined safety data have recently been reported

(10). The median follow-up for the cohort, at 38 months, was relevant for disease outcomes. Data on the status of ongoing B-cell aplasia as well as MRD assessment are presented. These data document for the first time a significant difference in EFS when B-cell aplasia is lost by 3, 6, and 9 months after infusion compared with those with ongoing B-cell aplasia at these time points (see Pulsipher and colleagues, Supplementary Fig. S3), as well as an incremental improvement in EFS for those in B-cell aplasia at these time points (2-year EFS of 63%, 72%, 83%, and 88% at 3, 6, 9, and 12 months, respectively).

To date, there have been no reports systematically exploring use of molecular MRD assessment following therapy with tisagenlecleucel. There is utility in the data Pulsipher and colleagues provide comparing MFC-MRD with NGS-MRD beyond the already established increased sensitivity of the latter, but importantly, also in exploring the optimal threshold for positivity and lead time to relapse from a positive result in order to identify the best biomarker of relapse risk.

In this context, a useful biomarker is one that:
(i) Faithfully predicts relapse with high sensitivity and specificity, and a high positive predictive value
(ii) Is feasible to obtain with standard clinical sampling, has a low failure rate, uses a robust, reliable, accurate, and precise assay, and is ideally obtained by minimally invasive approaches

(iii) Separates populations with uniformly good outcomes who can be spared the toxicity of further therapy, as well as those with poor outcomes for whom further therapy is justified

(iv) Is reliably detected well before frank relapse, with a detection window long enough to allow planning for interventions including SCT

In this study, while MFC-MRD fared positively in relation to criteria 1 to 3, 50% of patients relapsed without a prior positive MFC-MRD result, and the median interval to relapse from detection of a positive MFC-MRD result was 52 days (Pulsipher and colleagues, Fig. 2B). Thus, relapse could occur before adequate time to initiate further therapy, namely SCT. Despite the greater sensitivity of NGS to detect MRD, the median interval from NGS-MRD detection at a standard cut-off (10^{-6}) to relapse was only 18 days longer than for MFC (70 days). Furthermore, the outcomes of patients with a day 28 result positive at a level of 10^{-6} or more were not significantly different than those with a negative result (Pulsipher and colleagues, Fig. 3A and B), suggesting that NGS-MRD at this detection threshold was a poor biomarker in relation to characteristic 3 above. However, greatly improved outcomes were noted for patients with a result that was negative at any level of detection (median EFS and overall survival not reached vs. 5.8 and 12 months, respectively). This also translated into a greater median interval before frank relapse of 168 days (range 47–330 days).

The concern with making clinical decisions on NGS-MRD results well below the quantifiable range, as discussed above, is the likelihood of a false-positive result. The authors assess the value of repeated measurements and their consequence in terms of subsequent outcomes, demonstrating relapse in 19 of 26 patients with repeated positive NGS-MRD results from day 28 after infusion, whereas subsequently negative results were less concerning for relapse risk. Furthermore, the predictive value of positive NGS-MRD at any threshold of detection at months 3 and 6 after infusion was strongly associated with increased relapse risk, and there were no survivors without further therapy in this group.

While it is clear from these data that NGS-MRD has the potential to be a powerful predictor of relapse after infusion of tisagenlecleucel, multicenter prospective validation is needed to establish the wider applicability of this approach. Ideally, NGS-MRD should be compared with PCR-MRD, as this is the next most sensitive methodology most widely used in Europe, with rigorous standardization, interlaboratory collaboration, and quality assessment—efforts that have been in place for decades. Furthermore, the use of less invasive methods of MRD detection, such as peripheral blood NGS-MRD, should be explored in detail. Peripheral blood samples can be more easily obtained than bone marrow, thus having potential to increase the interval from detection to relapse. Further methodologic refinements such as establishing leukemic clones at the time of initial relapse as opposed to screening pre–CAR T-cell treatment may reduce the false-positive rate and quality control failure rate.

Authors' Disclosures

REFERENCES

1. Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, et al. Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. N Engl J Med 2018;378:439–48.
2. Schultz LM, Baggott C, Prabhu S, Pacenta H, Phillips CL, Rossoff J, et al. Disease burden impacts outcomes in pediatric and young adult B-cell acute lymphoblastic leukemia after commercial tisagenlecleucel: results from the Pediatric Real World CAR Consortium (PRWCC). Blood 2020;136:14–5.
3. Shah NN, Lee DW, Yates B, Yuan CM, Shalabi H, Martin S, et al. Long-term follow-up of CD19-CAR T-cell therapy in children and young adults with B-ALL. J Clin Oncol 2021;39:1650–9.
4. Finney OC, Brakke HM, Rawlings-Rhea S, Hicks R, Doolittle D, Lopez M, et al. CD19 CAR T cell product and disease attributes predict leukemia remission durability. J Clin Invest 2019;129:2123–32.
5. Kotrova M, van der Velden VHJ, van Dongen JJM, Formankova R, Sedlacek P, Brüggemann M, et al. Next-generation sequencing indicates false-positive MRD results and better predicts prognosis after SCT in patients with childhood ALL. Bone Marrow Transplant 2017;52:962–8.
6. Wood B, Wu D, Crossley B, Dai Y, Williamson D, Gawad C, et al. Measurable residual disease detection by high throughput sequencing improves risk stratification for pediatric B-ALL. Blood 2018;131:1350–9.
7. Myers RM, Taraseviciute A, Steinberg SM, Lamble AJ, Sheppard J, Yates B, et al. Blinatumomab nonresponse and high-disease burden are associated with inferior outcomes after CD19-CAR for B-ALL. J Clin Oncol 2021 Nov 12 [Epub ahead of print].
8. Dourthe ME, Rabian F, Yakouben K, Chevillon F, Cabannes-Hamy A, Méchinaud F, et al. Determinants of CD19-positive vs CD19-negative relapse after tisagenlecleucel for B-cell acute lymphoblastic leukemia. Leukemia 2021 May 17 [Epub ahead of print].
9. Pulsipher MA, Han X, Maude SL, Laetsch TW, Qayed M, Rives S, et al. Next-generation sequencing of minimal residual disease for predicting relapse after tisagenlecleucel in children and young adults with acute lymphoblastic leukemia. Blood Cancer Discov 2022;3:66–81.
10. Levine JE, Grupp SA, Pulsipher MA, Dietz AC, Rives S, Myers GD, et al. Pooled safety analysis of tisagenlecleucel in children and young adults with B cell acute lymphoblastic leukemia. J Immunother Cancer 2021;9:e002287.