

Determinants of response to CAR T cell therapy in a multi-centric, multi-trial cohort of patients with non-Hodgkin lymphoma

Corresponding Author: Dr Davide Bedognetti

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this manuscript, Dr. Budka and colleagues from Kite study the impact of Juvenile T cells leading to prolonged survival without impacting safety. The study is run on samples from a multi-centric, multi-trial cohort of 580 patients with non-Hodgkin lymphoma treated with CAR T-cell therapy (axi-cel and brex-cel). The size of this cohort is impressive. Interestingly, they found that a high proportion of "juvenile" (CCR7+CD45RA+) T cells in the product was the most significant feature associated with outcome, including the reduction of disease progression risk by 56%. Interestingly, there was no association with $\geq$ G3 CRS/NE, supporting an improved therapeutic index if you can manufacture juvenile-rich products.

The manuscript is not conceptually novel compared to other studies, but evidentially, yes. This manuscript doesn't discover a new biology, but it delivers a large, well-controlled analysis (probably the largest) that materially strengthens the evidence that naïve-like T cells in the product are beneficial for long-term results. Effect size and independence are key features for this paper: a higher juvenile fraction is the single strongest pre-treatment predictor of PFS, even after multivariable adjustment (considering tumor burden, CRP, ferritin, LDH) and after adding CAR-T expansion to the model. The implications of this work are multiple, including manufacturing optimization with shorter time in culture and agents that can prevent differentiation.

I have comments that might increase the impact and novelty of this work.

Main comments and suggestions:

- I think it would be important to test an external validation set with non-axi-cel products to test generalizability (tisa-cel, liso-cel)
- The authors appropriately compared the juvenile fraction in product vs. apheresis, which partially addresses baseline immune status. However, I think it is also important to check the prognostic value of baseline lymphocyte counts, CD4:CD8 ratio, and additional phenotyping, such as product exhaustion markers and other T cell subsets (e.g., stem cell memory).
- Prior Nature Medicine publications from the Kite group and collaborators have demonstrated that multiple tumor-related biomarkers, including the Immunoscore (CD3/CD8 IHC), IS21, B-cell GES, SII, BPI, CD19 expression (mRNA/protein), carry prognostic value for CAR-T outcomes. To demonstrate the novelty and additive value of the juvenile T-cell fraction, when available (e.g., ZUMA1, 7), please compare juvenile T cells directly to these established predictors both individually and in combination. This benchmarking is essential to contextualize the juvenile fraction within the existing CAR-T prognostic landscape and to dramatically increase the impact of this work.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

Thank you for the opportunity to review this manuscript. In this study, Budka and Davis et al. analyzed data from 580 patients with non-Hodgkin lymphoma treated with axicabtagene ciloleucel (axi-cel) to identify modifiable product characteristics linked to long-term outcomes. A high proportion of naïve (CCR7+CD45RA+) T cells in the CAR T-cell product was strongly associated with improved progression-free survival (PFS). This association remained significant even after adjusting for

other clinical and biological factors, including tumor burden, inflammatory markers, and CAR T-cell kinetics. Importantly, a high proportion of the naïve T-cell phenotype did not correlate with increased toxicity. The study also found that more prior lines of therapy were associated with fewer naïve T cells, underscoring the potential value of earlier CAR T-cell therapy.

This large multi-cohort study confirms and builds on findings from smaller prior studies. The manuscript is well-written and supported by comprehensive, standardized data.

Comments:

- The term "juvenile" T cells is somewhat unconventional. Given the limited immunophenotyping and the absence of an anti-CD95 antibody necessary to identify stem cell-like T cells, I believe using the term "naïve" or "naïve-like" T cells would be more accurate and appropriate.
- The claim that the proportion of juvenile T cells in the product is the strongest predictor of PFS may be misleading if it is based solely on the P-value. For example, in the first multivariable analysis, the effect size of the proportion of juvenile T cells is similar to that of LDH, and in the second model, peak CAR T-cell expansion shows a larger effect size. To advance understanding in the field, it would be helpful to clarify the relative contributions of key factors influencing complete response and PFS—such as LDH, CAR T-cell kinetics, and the proportion of juvenile T cells in the product. Notably, Figure 4A suggests that a high proportion of juvenile T cells may not fully mitigate the adverse impact of high tumor burden. For the multivariable Cox regression models, it might be helpful to use the Wald statistic to rank predictor importance and perform variable drop-out analyses to compare the model's C-index. Please also provide Cox regression modeling using continuous variables as supplementary data to help estimate the effect size of each variable.
- Given the size of the dataset, it would be informative to explore the ratio of CAR T-cell expansion and/or proportion of juvenile T cells to tumor burden and/or LDH and its association with CR and PFS.
- The more pronounced differences in the proportion of juvenile T cells when comparing ongoing response rate rather than CR rates may suggest a role of CAR T-cell persistence in sustaining responses. Do you have data on the relationship between the proportion of juvenile T cells in the product and CAR T-cell persistence (e.g., time to last detection, AUC0-60, AUC0-90, etc.) or duration of B-cell aplasia?
- The progressively shorter PFS observed with lower juvenile T-cell proportions in the Mid and Low CAR T-cell expansion groups suggests that peak CAR T-cell expansion may be the primary driver of initial response, while a higher proportion of juvenile T cells—due to their greater self-renewal capacity—may contribute to improved persistence (AUC in the first 6 months) and, consequently, better PFS.
- Please provide a table summarizing the baseline characteristics across the different axi-cel studies. If available, a breakdown of prior lines of therapy into chemotherapy-containing versus targeted therapy regimens would be helpful. Additionally, it would be informative to assess the correlation between time from last treatment—particularly cytotoxic chemotherapy—and the proportion of juvenile T cells.
- The observation that the inverse correlation between the number of prior regimens and the proportion of juvenile T cells is stronger in the final product than in the apheresis is noteworthy and highlights the limitations of phenotyping alone in assessing T-cell state and function.

Minor comments:

- For Fig 1F, Extended Data Fig 2B, and Extended Data Fig 3B-C, please also provide ANOVA and/or Kruskal-Wallis test results.
- Were there any differences in the median number of prior lines of therapy between males and females, or between patients with ECOG 0 versus 1?
- Please clarify in the Results section that the AUC refers to the period from day 0 to day 28 following CAR T-cell infusion.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

This paper presents a multivariate analysis of 580 patients across 5 multi-center clinical trials who have received axi-cel, the CD19.28z CART product of KITE-Gilead, for a range of NHL subtypes. Multiple studies to identify correlates of response and response duration have been published, highlighting the importance of T-cell subsets in the apheresis product and the infusion product (going back to 2014) as well as disease burden, baseline inflammation (of the patient), prior lines of therapy and CART expansion and persistence. This analysis studies all of these parameters in a much larger group of patients with a median follow up of 41.9 months.

They find that the frequency of what they call "juvenile" cells, phenotypically CCR7+CD45RA+, a population that contains both naïve and T memory stem cells – cells with very different characteristics- to be the factor most positively associated with tumor response and duration in both univariate and multivariate analyses. Other factors such as baseline tumor burden, baseline inflammatory markers are also significant, but less so. The correlation held true across all five trials. These findings are not new or surprising, but are very clear and well presented in this manuscript. There was a correlation between the number of juvenile T-cells in the apheresis product and the infusion product, but their presence in the product was more significant.

Specific comments

1. You remark that the positive effect of a high frequency of juvenile T-cells is actionable, and refer to two clinical trials, but you do not describe how more juvenile T-cells could be obtained or preserved. You should discuss the many other studies that have attempted to increase the juvenile phenotype and other phenotypes.

2. It is important to know if naïve or TSCM's should be preserved. This distinction could have added innovation to this otherwise impressive, but descriptive work
3. Extended Data Figure 1 could have been better explained for non-statisticians
4. Was Supplementary Figure 1 describing the apheresis product or the infusion product
5. There is only a weak correlation between CART expansion and AUC with juvenile T-cells. Could this be influenced by the copy# per cell. Could you supply the (median and range), since the PCR does not necessarily reflect the number of circulating cells

(Remarks on code availability)

What code?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have addressed my comments. I think this revised manuscript provides a more comprehensive narrative and evidence.

Reviewer #2

(Remarks to the Author)

The authors have addressed the reviewers' comments in a thoughtful and thorough manner. Their revisions have improved the clarity and presentation of the manuscript and have helped place the findings in a more appropriate context. Overall, the manuscript has been significantly strengthened and, in its current form, is suitable for publication.

Reviewer #3

(Remarks to the Author)

The authors have responded extraordinarily to the reviewers' comments and have added much new data, incorporating patients from additional clinical trials and Cytof data.

The clinical dataset is exceptional, with 580 patients in the discovery cohort now validated by 203 patients in an independent validation cohort over multiple ZUMA studies with a long median follow-up. The statistical analyses appear rigorous, and the association between naïve-like product T cells and durable clinical benefit is convincing.

However, the manuscript remains primarily correlative and confirmative of earlier, smaller studies showing associations between naïve-like cells and PFS and between the frequency of circulating naïve cells and lines of prior chemotherapy. The mechanistic interpretation lacks supporting biological data. The authors favor the hypothesis that PFS is driven by TSCM cells, but since this population cannot be distinguished from naïve cells once activated, this remains speculative. The association between naïve-like T cells and PFS is strong, but remains an association so phrases implying a causal relationship should be modified. The manuscript would be strengthened by a deeper discussion of alternative biological explanations.

Specific comments

The Cytof data could be further exploited to identify pathways that distinguish Cluster 1. Does CXCR3 remain after adjusting for CCR7, are these cells metabolically distinct? Are the naïve-like cells polarized towards TH1/2/17 etc

The CyTOF analysis of the apheresis samples should have allowed you to distinguish naïve T cells from T memory stem cells (TSCM), for example using CD95, CD122, CXCR3, or other markers. If so, was the association with progression-free survival driven predominantly by naïve T cells, TSCM, or both?

Line 118: Your prediction that CD45RA+ CCR7+ naïve-like cells more closely resemble TSCMs is unlikely to be correct. TSCMs represent less than 2% of all T-cells in healthy individuals and would be expected to maintain a similar frequency during the 14 days expansion, unless they expanded preferentially.

Can you comment from the literature whether naïve-like cells are superior in functional assays, for example, in serial killing ?

The writing has become a little repetitive due to the additional data

(Remarks on code availability)

I do not have the expertise to review the code

Open Access This Peer Review File 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.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File 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 <https://creativecommons.org/licenses/by/4.0/>

Point-by-point response to referees

We sincerely thank the reviewers for their thoughtful, rigorous, and constructive evaluation of our work. We are grateful for the positive assessment of the study and for the insightful comments and suggestions provided throughout the review process. We agree that many of the questions raised were important and helped us substantially strengthen both the clinical and biological aspects of the manuscript. In response to the reviewers' comments, we have substantially expanded the scope of the study and performed numerous additional analyses and experimental investigations.

Collectively, these additions substantially improve the rigor, generalizability, and mechanistic interpretation of the findings.

Major revisions included in the revised manuscript:

- Validation of the primary findings in an independent meta-cohort of 203 additional patients, including patients treated with brexu-cel or axi-cel combination studies, expanding the overall study population from 580 to 783 patients.
- New analyses demonstrating that the association between product naive-like T cells and outcome remains significant when evaluated alongside previously reported tumor microenvironment biomarkers, including B-cell gene-expression signatures, CD19 expression, IS21, SII, and CD3/CD8 immune infiltration metrics.
- Addition of a new high-dimensional CyTOF analysis of 116 CAR T-cell products, providing deeper characterization of product-associated T-cell populations and identifying favorable naive-like CD8 T-cell populations enriched for markers of cellular fitness and memory.
- Expanded analyses of baseline immune status, including baseline lymphocyte counts and product and apheresis CD4/CD8 ratios.
- Additional multivariable modeling, including continuous-variable Cox regression, variable-importance analyses / dominance analyses.
- New analyses examining tumor-burden-normalized CAR T-cell and naive-like T-cell metrics.
- Additional evaluation of prior therapies, timing of previous cytotoxic chemotherapy, demographic factors, and CAR T-cell persistence.
- Replacement of the term "juvenile T cells" with "naive-like T cells" throughout the manuscript

A total of 60 new figure panels (12 new figures: 2 main, 3 extended, and 7 supplementary) and 3 new (supplementary) tables have been added to the revised manuscript. For the reviewers' convenience, all new and revised figures and tables are also included throughout this response document.

Point-by-point response

Reviewer #1 (Remarks to the Author)

In this manuscript, Dr. Budka and colleagues from Kite study the impact of Juvenile T cells leading to prolonged survival without impacting safety. The study is run on samples from a multi-centric, multi-trial cohort of 580 patients with non-Hodgkin lymphoma treated with CAR T-cell therapy (axi-cel and brex-cel). The size of this cohort is impressive. Interestingly, they found that a high proportion of "juvenile"

(CCR7+CD45RA+) T cells in the product was the most significant feature associated with outcome, including the reduction of disease progression risk by 56%. Interestingly, there was no association with $\geq$ G3 CRS/NE, supporting an improved therapeutic index if you can manufacture juvenile-rich products.

The manuscript is not conceptually novel compared to other studies, but evidentially, yes. This manuscript doesn't discover a new biology, but it delivers a large, well-controlled analysis (probably the largest) that materially strengthens the evidence that naive-like T cells in the product are beneficial for long-term results. Effect size and independence are key features for this paper: a higher juvenile fraction is the single strongest pre-treatment predictor of PFS, even after multivariable adjustment (considering tumor burden, CRP, ferritin, LDH) and after adding CAR T expansion to the model. The implications of this work are multiple, including manufacturing optimization with shorter time in culture and agents that can prevent differentiation. I have comments that might increase the impact and novelty of this work.

We thank the reviewer for the assessment of our work. We appreciate the recognition of the size and breadth of the cohort, as well as the acknowledgment that this study provides robust evidence supporting the clinical relevance of product naive-like T cells and has multiple implications. Indeed, in this *Analysis* manuscript we aimed at providing arresting conclusions with broad impact. We are grateful for the constructive suggestions provided below, which helped us further, and significantly, strengthen the manuscript through additional analyses, validation studies, and expanded characterization.

Main comments and suggestions:

1.1 - I think it would be important to test an external validation set with non-axi-cel products to test generalizability (tisa-cel, liso-cel)

We appreciate the insightful comment and agree that there is utility in evaluating the association with naive-like T cells and clinical outcomes particularly in a non-axi-cel product. Indeed, our analysis did not include a validation cohort.

We evaluated whether the proportion of naive-like T cells within the infused product was associated with clinical outcome in an independent meta-cohort of 2023 patients including subjects treated with brexu-cel and axi-cel combinations (axi-cel and rituximab or atezolizumab). A new figure (**Fig. 5**), new extended data figures (**Extended Data Figures 7 and 8**) and a Supplementary Figure (**Supplementary Fig. 5**) have been added. We were able to validate the reported findings using the same cut-offs generated in the original cohort. Methods, results, and discussion were updated accordingly.

We did not have access to tisa-cel or liso-cel samples, nor to publicly available phenotyping datasets with matched markers and clinical outcomes for these products. We contacted several colleagues from different centers but were not able to find a sizable collection of samples that could be analyzed. Instead, most have access to pretreatment PBMCs, which highlights the unicity of our product dataset.

While this limitation prevents us from generalizing the identified cut-off as a biomarker, as the optimal threshold may vary across CAR-T products and manufacturing processes, our findings conceptually align with previous reports that have identified similar biological signals in product or apheresis samples from smaller cohorts of patients treated with 4-1BB CAR-T therapies. For

instance, a naive-like T-cell population in apheresis has been associated with favorable clinical outcomes following 4-1BB CAR-T treatment (Fraietta et al., *Nat Med*, 2018; Chen et al., *Nat Biomed Eng*, 2026; Cuffel et al., *Blood Adv*, 2022). However, our study demonstrates that the association between naive-like T cells and clinical efficacy is stronger when this population is assessed in the final CAR T product rather than in the apheresis material, highlighting a potentially actionable feature during manufacturing. Consistent with this observation, a recent study reported that higher frequencies of CD8 naive-like CAR T cells at day 7 after axi-cel infusion were associated with complete response (Schneider et al., *Nat Commun*, 2026; 19 patients with CAR T expansion data analyzed). Notably, in this cohort, differences in CD8 naive-like T-cell frequencies were not detected in the corresponding apheresis samples, likely owing to limited statistical power. Nevertheless, these findings are consistent with the in vivo expansion of CD8 naive-like T cells already present within the infused product, a population that our analysis identified as strongly associated with clinical response. Additional CyTOF analyses included in the revised manuscript in response to Reviewer 1 (comment 1.2) further support this interpretation.

We believe our data establishes landmark parameters that can serve as a reference for future comparisons in appropriately assembled cohorts of patients receiving different CAR T cell therapies.

Below are details regarding the cohort analyzed and the findings. Related figures included in the manuscript are also provided here for reviewers' convenience.

Brexu-cel is an autologous CD19-directed CAR T-cell product approved for the treatment of relapsed/refractory mantle cell lymphoma (MCL) after ≥ 1 prior line of therapy, based on the results of the ZUMA-2 study. One hundred fifty-four evaluable patients from ZUMA-2 Cohorts 1 and 3 were included in this correlative analysis; Cohort 2 was not included as it evaluated a reduced CAR T-cell dose. The brexu-cel manufacturing process includes a T-cell selection and lymphocyte enrichment step (Wang, et. al., *NEJM*, 2020), a process which is necessary in certain B cell malignancies where high levels of circulating lymphoblasts are common, such as in MCL. We also included studies where axi-cel was administered with combination therapies atezolizumab (ZUMA-6, N = 23) or rituximab (ZUMA-14, N = 26). A total of 203 patients were included in the validation meta-cohort (**Fig. 5A**), with a median follow-up of 27 months. Additional details on ZUMA-2, ZUMA-6, and ZUMA-14 are provided in the methods section of the manuscript.

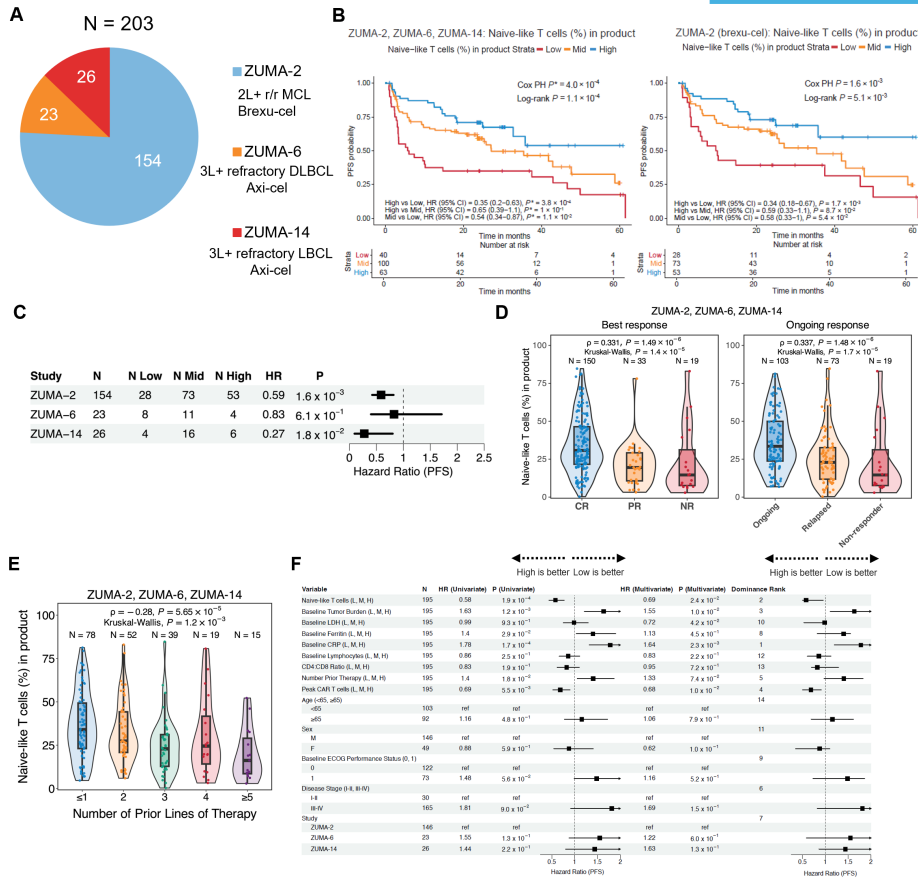
Kaplan-Meier PFS analysis was performed with the proportion of naive-like T cells in the product (combined validation meta-cohort and brexu-cel specific) converted to a categorical variable using the same numeric cutoffs defined in the original ZUMA-1, -5, -7, -12 meta-cohort., Cox proportional hazard regression analysis demonstrated that patients with a *High* proportion of naive-like T cells in the product had reduced risk of disease progression by 65% and 66% compared to patients with *Low* naive-like T cell proportions in the validation meta-cohort (ZUMA-6, ZUMA-4, and ZUMA-2 combined) the brexu-cel specific (ZUMA-2 cohort) respectively (**Fig. 5B**). This is remarkable as the reduction of risk of relapse was numerically greater than what observed in the original cohort (54% reduction of risk of disease progression). The proportion of naive-like T cells was significantly associated with PFS using log-rank P value for categorical comparisons (*Low*, *Mid*, *High*) and Cox proportional hazards (PH) P value with study as a covariate for ordinal assessment (*Low* = 1, *Mid* = 2, *High* = 3) in both validation cohorts. We evaluated the impact of the proportion of naive-like T cells in the product on PFS within individual studies and saw consistency across trials in the validation meta-cohort (**Fig. 5C**), including the brexu-cel (ZUMA-2) specific cohort ($P = 1.6 \times 10^{-3}$). Further exploration into the

impact of naive-like T cells on complete response and durability of response showed a progressive increase in CR and ongoing response rate according to the proportion of naive-like T cells in the validation meta-cohort (**Fig. 5D**). Consistent with the observations from analysis of the discovery meta-cohort, there was a negative correlation observed between number of lines of prior therapy and the proportion of naive-like T cells in the product (Spearman $\rho = -0.28$, $P = 5.7 \times 10^{-5}$; Kruskal-Wallis, $P = 1.2 \times 10^{-3}$; **Fig. 5E**) within the validation meta-cohort. To identify whether naive-like T cells are independently associated with improved response rates, we performed a multivariate PFS analysis in which continuous clinico-pathologic variables were categorized in the same way as the naive-like T cell proportion (*Low* = $\leq 25^{\text{th}}$ percentile, *Mid* = $25^{\text{th}} - 75^{\text{th}}$ percentile, *High* = $> 75^{\text{th}}$ percentile; **Fig. 5F**). The association between the naive-like T cell proportion and PFS remained significant in the multivariate analysis ($P = 2.4 \times 10^{-2}$), irrespective of prior lines of therapy and CAR T peak expansion. Multivariate PFS analysis was also performed with clinico-pathologic variables as continuous and the significance with PFS was maintained ($P = 7.1 \times 10^{-3}$; **Supplementary Fig. 5**). Further analysis of the impact of naive-like T cells within different inflammatory settings demonstrated the positive influence of naive-like T cells in patients with more pronounced tumor burden (sum of the products of the diameters of target lesions (SPD)) and systemic inflammation at baseline (*Mid* and/or *High* CRP, Ferritin, tumor burden, or LDH; **Extended Data Fig. 7**) in the validation meta-cohort. Similar to what is seen in the axi-cel meta-cohort analysis, the proportion of naive-like T cells in the product showed no correlation with the worst grade of CRS (**Extended Data Fig. 8E**) or NEs (**Extended Data Fig. 8F**). Importantly, when stratifying the proportion of naive-like T cells into *Low*, *Mid*, and *High*, there were similar rates of high-grade (grade ≥ 3) CRS (**Extended Data Fig. 8G**) and NEs (**Extended Data Fig. 8H**) across groups.

Overall, in an independent validation meta-cohort, including brexu-cel treated patients, the proportion of naive-like T cells in the product was associated with improved efficacy, without increased rates of severe CRS or NE.

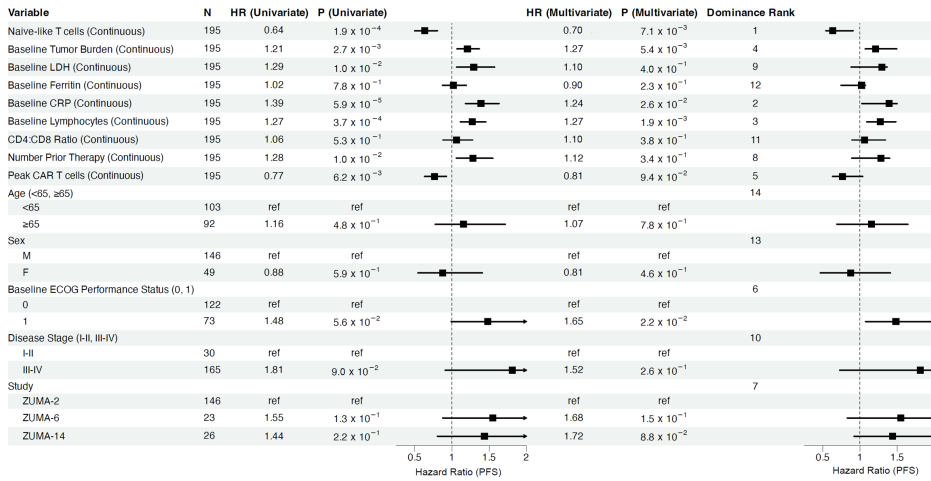
Fig. 5

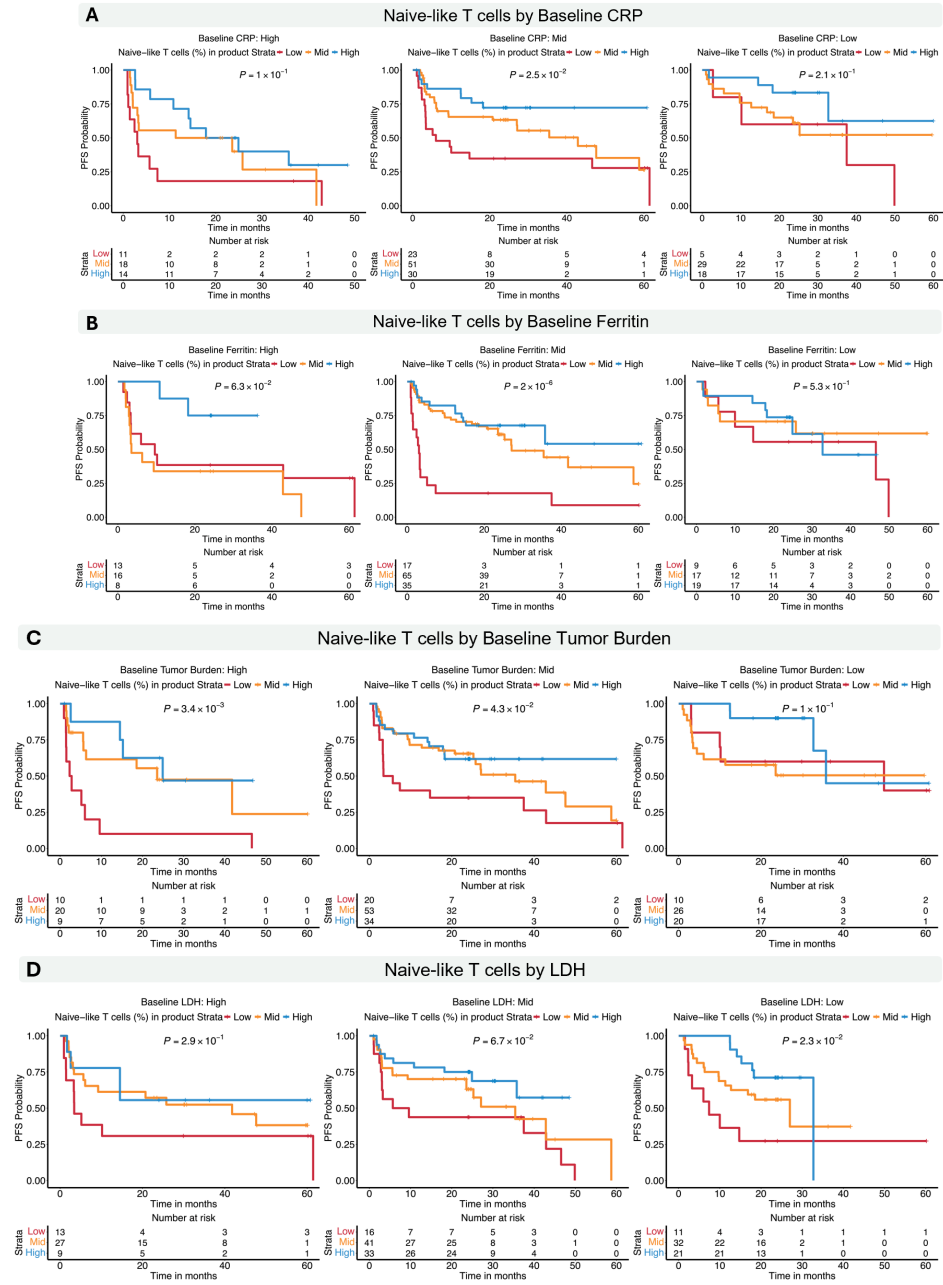
New Figure

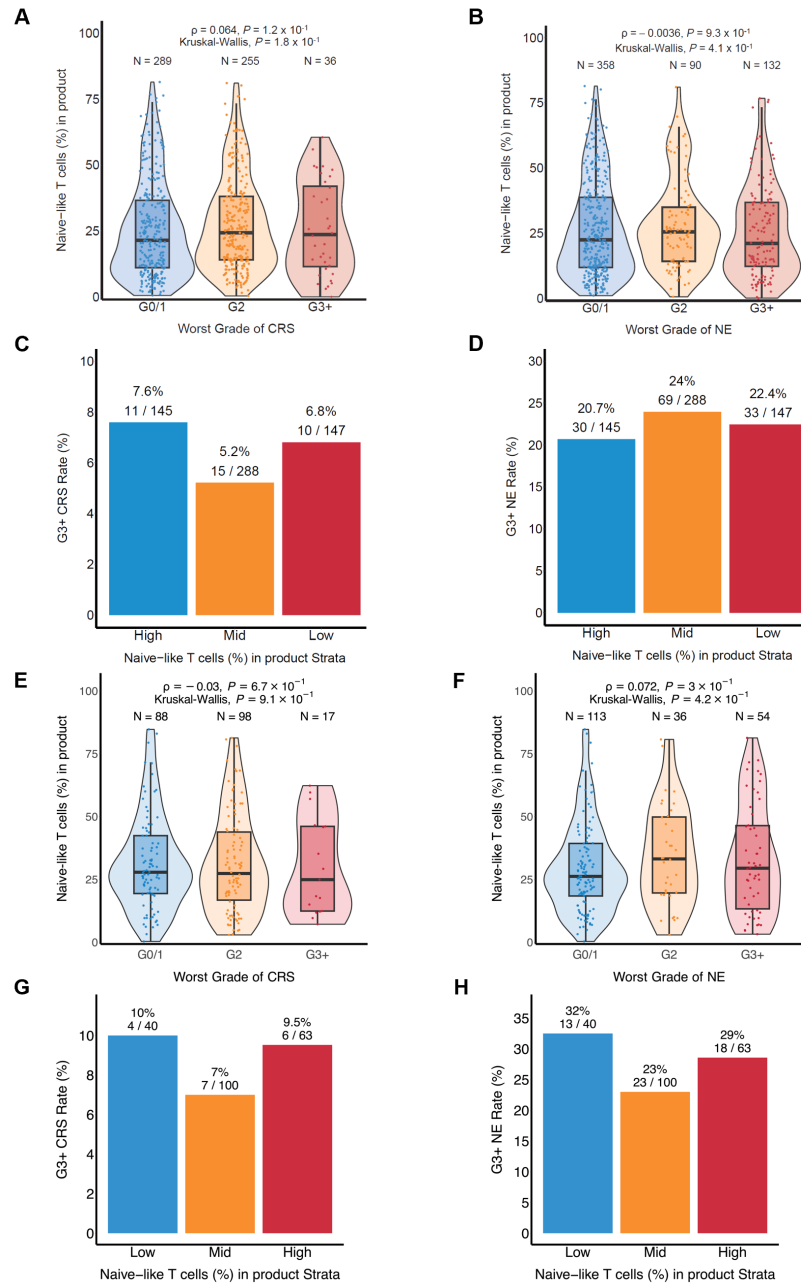


Supplementary Fig. 5

New Figure







1.2 - The authors appropriately compared the juvenile fraction in product vs. apheresis, which partially addresses baseline immune status. However, I think it is also important to check the prognostic value of baseline lymphocyte counts, CD4:CD8 ratio, and additional phenotyping, such as product exhaustion markers and other T cell subsets (e.g., stem cell memory).

We thank the reviewer for this comment. We performed all the suggested analysis.

We agree that representing additional features related to baseline immune status and further exploring phenotypic characteristics of the product are important for contextualizing the naive-like T cell (CCR7+CD45RA+, previously termed juvenile T cells) and providing further insight

into independent prognostic factors. To address this comment, we expanded the features in the univariate and multivariate PFS analyses to include baseline lymphocyte counts as well as the CD4 to CD8 ratio of the product in the axi-cel discovery meta-cohort (**Fig. 2F, Fig. 4D, Supplementary Fig. 2A, Supplemental Fig. 2B**) and the newly added validation meta-cohort (**Fig. 5F, Supplementary Fig. 5**). Importantly, in the 5-color panel, CD4 values were derived as $100 - (\% \text{ CD8}^+ \text{ of CD3 T cells})$, with the CD4/CD8 ratio therefore being mathematically equivalent to the percentage of CD8 for non-parametric based analyses.

The updated results highlight that naive-like T cells are maintained as one of the strongest PFS associated features when considered alongside other clinico-pathologic features. In the multivariate analysis, neither CD4 to CD8 ratio nor baseline lymphocyte count were significantly associated with PFS when categorized into *Low, Mid, High* strata in the axi-cel discovery meta-cohort (**Fig. 2F, Fig. 4D**) or validation meta-cohort (**Fig. 5F**). However, baseline lymphocyte count was negatively (yet marginally) associated with PFS in the axi-cel discovery meta-cohort when treated as a continuous feature ($P = 2.5 \times 10^{-2}$, **Supplementary Fig. 2A**). When CAR T expansion was included in the multivariate analysis, baseline lymphocyte count was negatively associated with PFS in the validation meta-cohort ($P = 1.9 \times 10^{-3}$, **Supplementary Fig. 5**), but not the axi-cel discovery meta-cohort ($P = 6.7 \times 10^{-2}$, **Supplementary Fig. 2B**). Flow-cytometry based CD4 to CD8 ratio was not associated with PFS in any of the univariate and multivariate analyses performed in either meta-cohort.

While the flow cytometry data utilized across the main and validation meta-cohorts was optimal for exploring consistently characterized phenotypes across multiple clinical trials, this data was generated with a 5-marker panel (CCR7, CD45RA, CD8, CD3, and Viability Dye), limiting the capacity to explore more advanced phenotyping in the meta-cohorts. To address the request for additional product characterization, we incorporated a cytometry by time-of-flight (CyTOF) analysis of 116 product samples across three LBCL clinical trials (ZUMA-1, ZUMA-6, and ZUMA-7). These samples represent a sub-cohort of the samples analyzed by 5-color flow cytometry.

This analysis employed 34 immune characterization markers (**Supplementary Table 4**) with a primary focus on markers that capture various aspects of T cell biology such as activation, proliferation, exhaustion, and memory status. Unsupervised clustering analysis for unbiased detection of populations within the axi-cel product was followed by univariate and multivariable (elastic net) Cox modeling to identify the product characteristics most strongly associated with PFS (**Fig. 6**). This analysis demonstrated that only one product cluster (cluster 1) significantly associated with improved survival, a CD8 naive-like T cell population expressing CCR7, CXCR3, CD45RA, CD95 (CD95 expression is not highlighted in the manuscript main results as a feature characterizing this clusters as it was broadly expressed across the majority of the cells, see **supplementary Figure 6A-B and in Extended Data Figure 9A**), NKG2D, and CD27, while three CD4 clusters (clusters 10, 18 and 26) expressing intermediate-high levels of activation markers such as ICOS, and CD25. were associated with disease progression (**Fig. 6C**). Cluster 26 demonstrated features of emerging exhaustion (OX-40 and PD-1 expression) while cluster 10 was distinguished by high CD27, absent CD45RA expression, intermediate CCR7, and elevated TIGIT and FOXP3, compatible with effector/activated T regulatory cells. Individual marker analysis based on percent positivity within CD3+CD45+ T cells identified consistent results as the cluster analysis, wherein markers of activation (CD25, ICOS, and OX-40) were linked with shorter PFS while markers of earlier T cell stages (CD27 and CD45RA) were associated with prolonged progression-free survival (**Fig. 6C**). Intriguingly, individual markers associated with disease progression were almost exclusively absent on the CD8 naive-like T cell population, cluster 1, while all markers associated with durability of response were

highly expressed in this population (**Fig. 6D and Fig. 6F**). Individual markers for all clusters are represented in **Supplementary Fig. 6**.

Explorative multivariable elastic net modeling incorporating all clusters and evaluated markers, identified naive-like T cell cluster 1, alongside naive-like T cell features (CD45RA) and cytotoxic marker NKG2D as independent predictors of improved PFS, whereas activated T cell clusters 10 and 18, alongside exhaustion/senescence cluster 6 (CD57^{hi}, TIGIT^{hi}) were associated with reduced PFS (**Fig. 6E**). Cluster abundance stratification demonstrated distinct PFS separation for the most beneficial (cluster 1) and adverse (cluster 18) clusters (**Fig. 6GH**). Given the presence of CCR7 and CD45RA in cluster 1, we performed a comparison of cluster 1 (%) from CyTOF with naive-like T cell (%) from the 5-color meta-cohort flow analysis and confirmed a strong correlation between these features (**Fig. 6I**). These findings support the positive relationship between naive-like T cells and response and provide added context in regard to the characteristics of these cells in the CAR T cell product.

Manual gating of the CyTOF data across the T-cell phenotypes defined by CCR7 and CD45RA (captured in the 5-color panel in the main and validation meta-cohorts) confirmed the lower frequency of exhaustion markers (TIGIT, PD-1, LAG-3) on CCR7+CD45RA+ T cells as compared to other T cell subsets (**Extended Data Fig. 9A**). Notably, CD95 was also highly expressed on all T-cell subpopulations, which is likely driven by the exogenous activation utilized during the CAR T-cell manufacturing process which confounds whether this population represents a naive or stem-cell memory population. Importantly, there is a high level of correlation between the naive-like T cell population identified using the 5-color flow cytometry panel and the CyTOF panel (**Extended Data Fig. 9E**) in matching subjects, with a greater proportion of naive-like T cells in earlier lines of therapy (**Extended Data Fig. 9FG**).

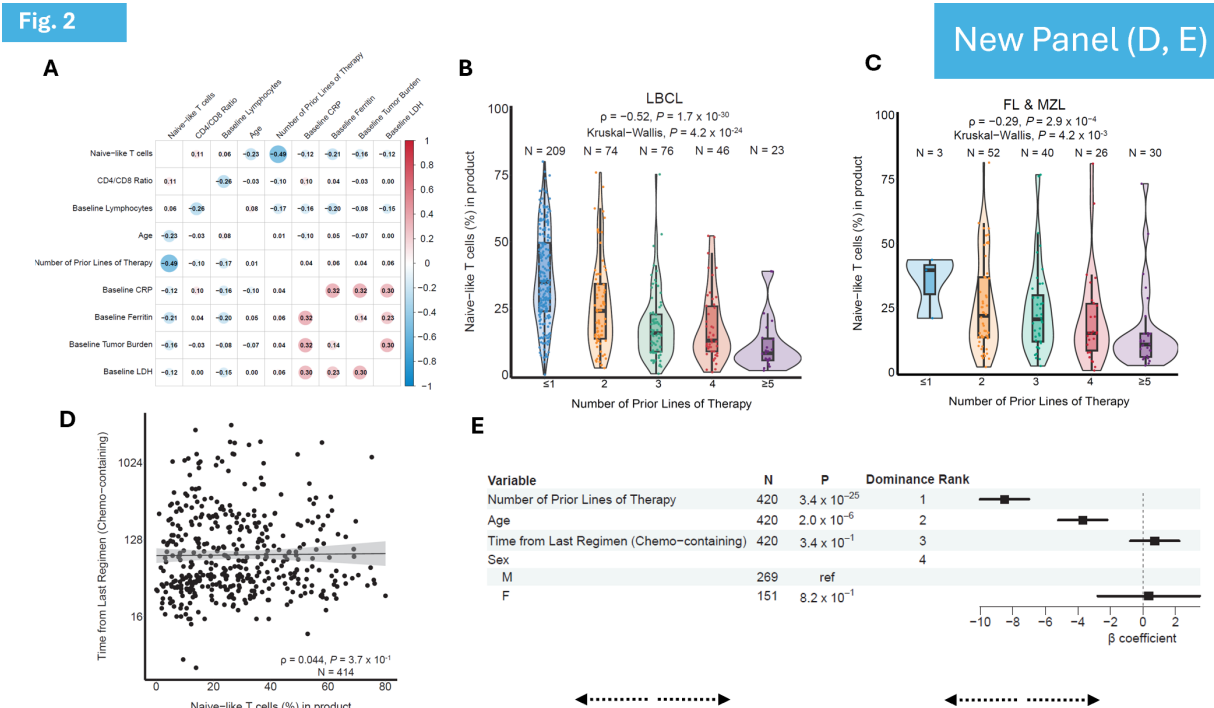
We further explored the CD4 to CD8 ratio within the CyTOF data as both markers were available within the CyTOF panel while only CD8 was captured within the Axi-cel meta-cohort 5-color panel, with CD4 values derived as $100 - (\% \text{ CD8}^+ \text{ of CD3 T cells})$. Although the 5-color panel did not contain both markers, we observed a strong correlation (Spearman $\rho = 0.89$, $P = 2.2 \times 10^{-16}$) between the CD4 to CD8 ratios generated using the CyTOF and the 5-color panel (**Extended Data Fig. 9D**). In fact, percentage of cells not expressing CD4 or CD8 was extremely low (Q1 = 3.7%, Median = 5%, Q3 = 6.925%).

While no correlations with PFS were observed with CD4/CD8 ratio in the main or validation cohorts analyzed by flow cytometry (**Fig. 1B 4D, 5F**), in the CyTOF analysis a high CD4/CD8 was associated with shorter PFS via univariable analysis (**Fig. 6C**).

Additionally, matched subject analysis between 5-color flow and CyTOF identified similar associations with response status with CD4 to CD8 ratios from both datasets (**Extended Data Fig. 9B and 9C**). While a significant association with response and survival were identified with the CD4 to CD8 ratio within the CyTOF analysis, multivariable elastic net modeling didn't identify CD4 or CD8 as key indicators of PFS, while activation markers and naive-like T cell clusters/markers were maintained in the model (**Fig. 6E**). This suggests that while CD4 and CD8 may be significant in CyTOF univariate analysis of product samples, their relationship with response may be redundant with, and outperformed by, the features maintained in the model. Through manual gating of the CyTOF data we've identified that CCR7+CD45RA+ cells (naive-like T cells) are predominantly CD8+, while CCR7-CD45RA- (effector memory) are predominantly CD4+ (**Extended Data Fig. 9A**). These findings suggest that the observed relationships between CD4 to CD8 ratio and differentiation states may contribute to the univariate association of CD4 to CD8 with response seen in this patient cohort. The absence of

an association of the CD4 to CD8 ratio with PFS in the main and validation meta-cohorts by univariate or multivariate modeling, along with the observed strong correlation between CD4/CD8 from the 5-color flow cytometry panel and the CyTOF analysis (**Extended Data Fig. 9D**), further supports that the CD4 to CD8 ratio is unlikely to be a major biomarker of response.

We additionally evaluated the CD4 to CD8 ratio in apheresis material (available for a subgroup of 235 patients), as the apheresis flow cytometry panel included both markers, and identified a strong correlation with the product ratio (**Extended Data Fig. 4G**). No correlations or trends were observed between the CD4 to CD8 ratio in the apheresis and PFS (**Extended Data Fig. 4H**), consistent with observations in the product (**Fig. 1B**). Consistently, a report in a small cohort of patients treated with the 4-1BB CAR T tisa-cel also did not find significant associations between CD4/CD8 ratio in apheresis and clinical response or PFS (Cuffel et al. Blood Advances 2022).



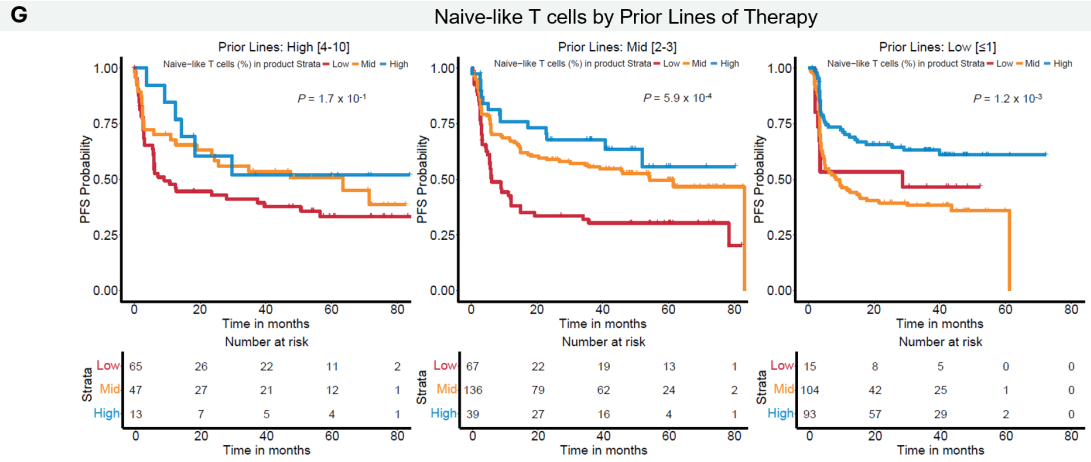
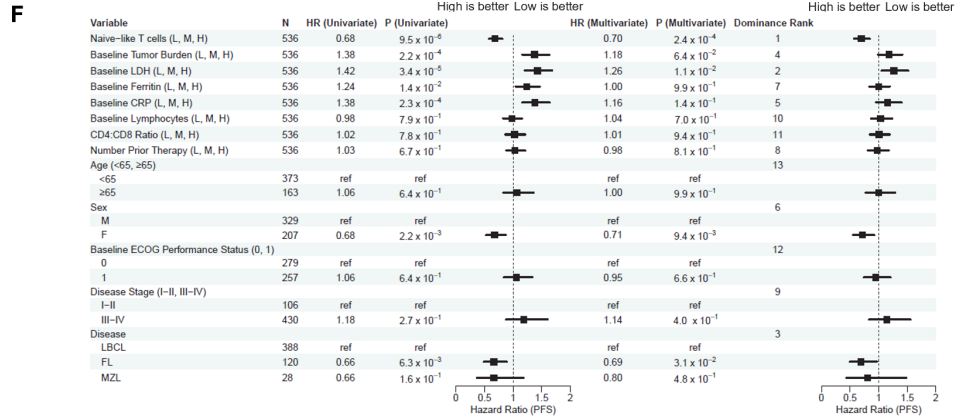
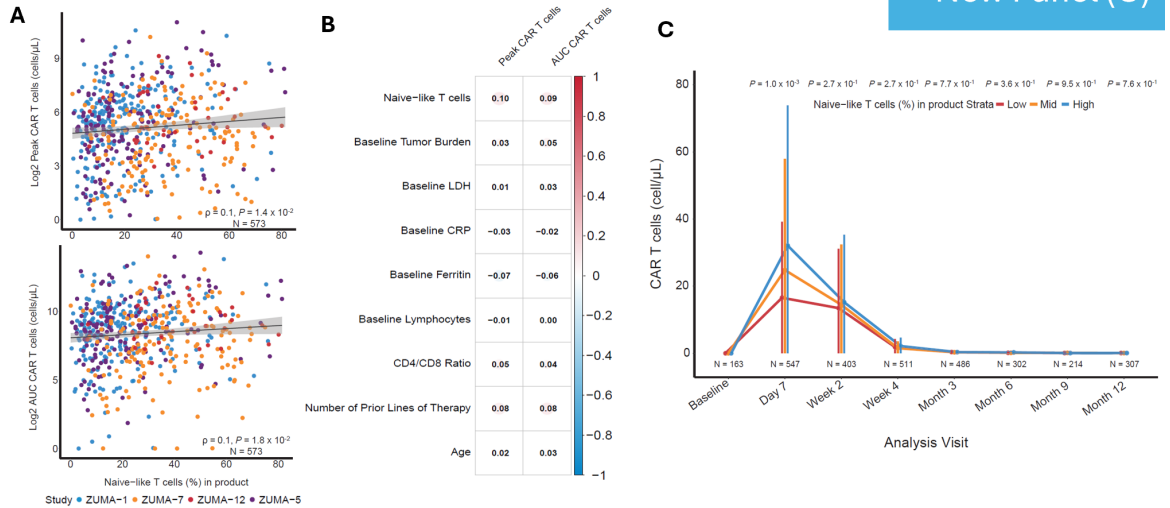
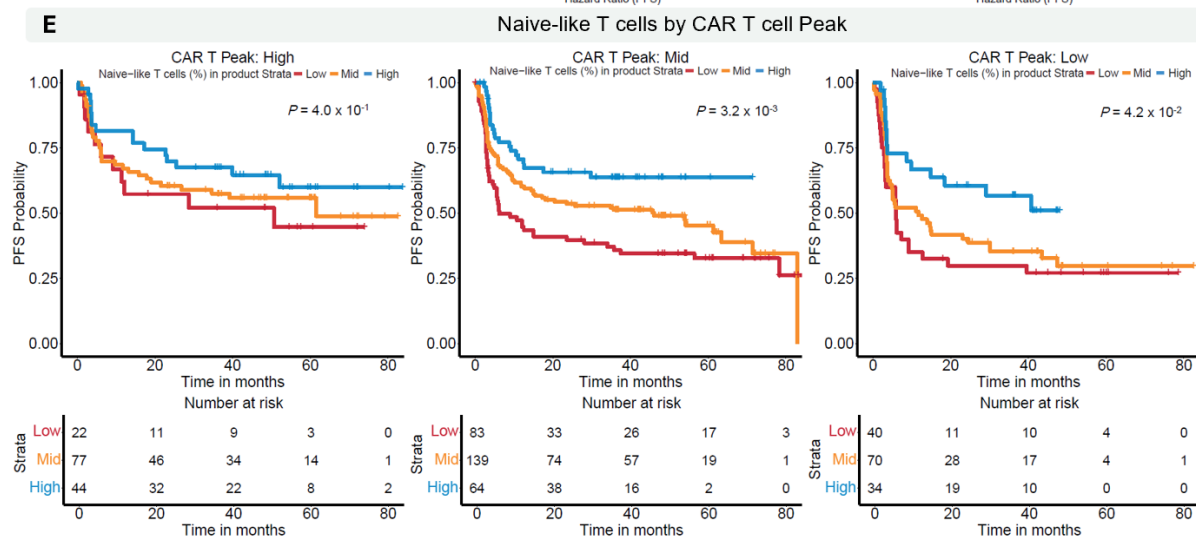
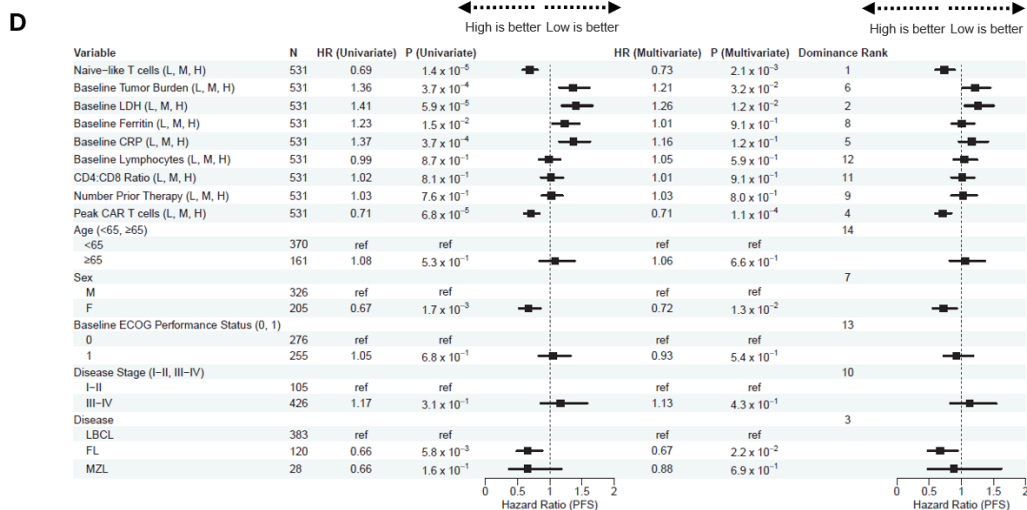
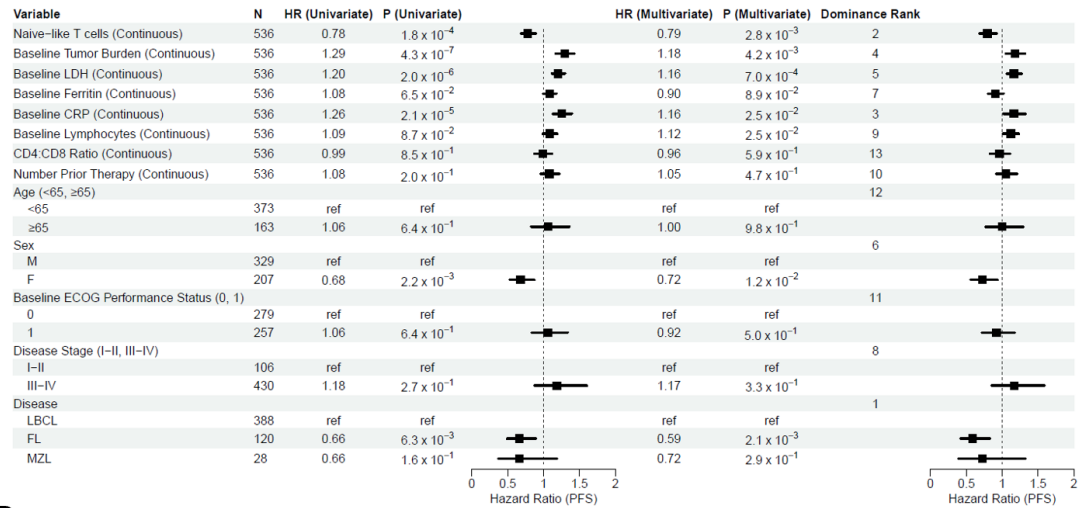


Fig. 4





A



B

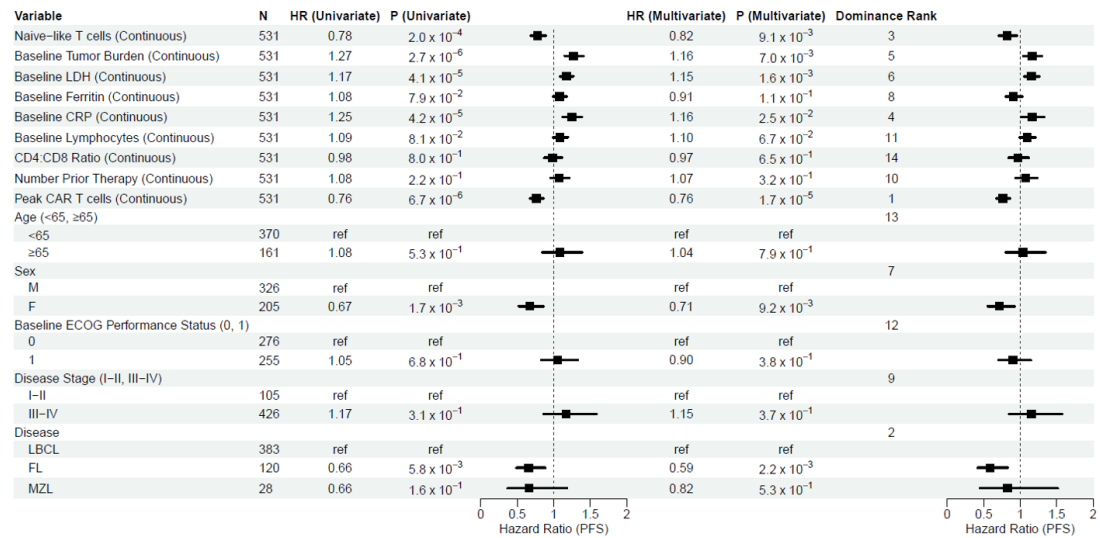
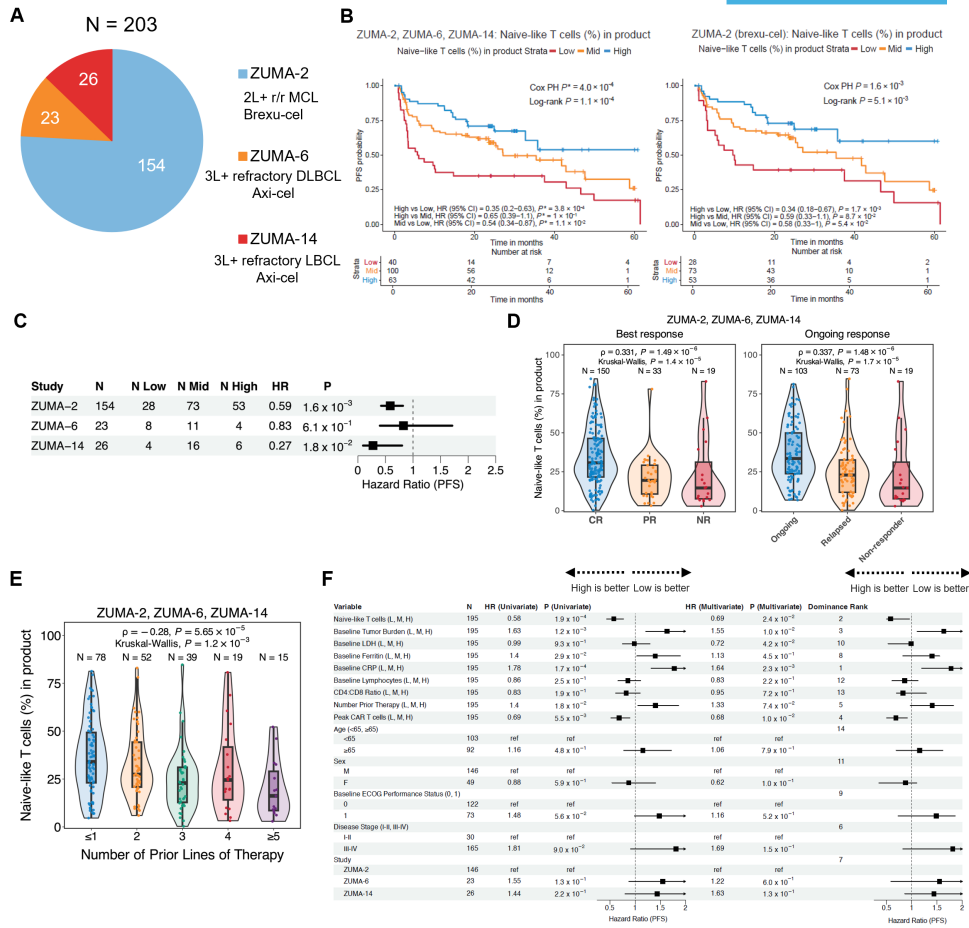


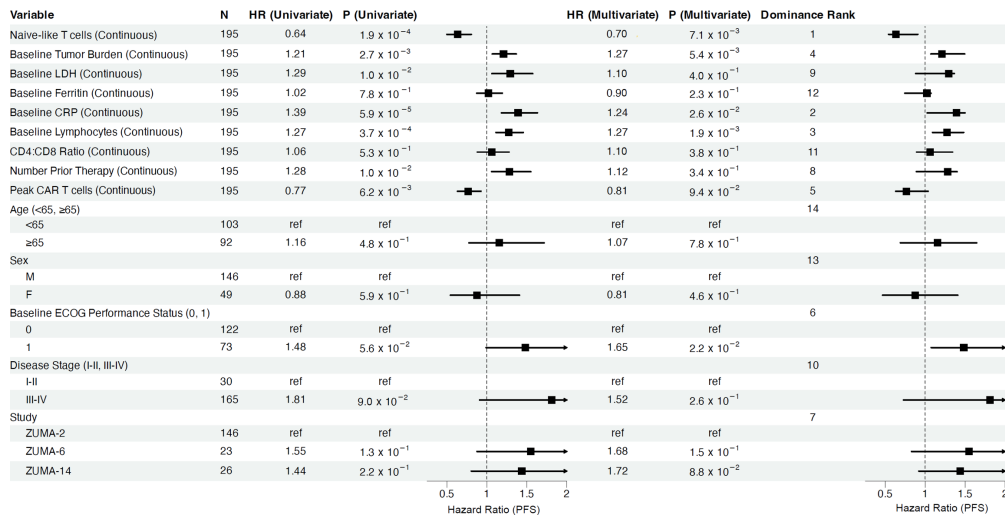
Fig. 5

New Figure



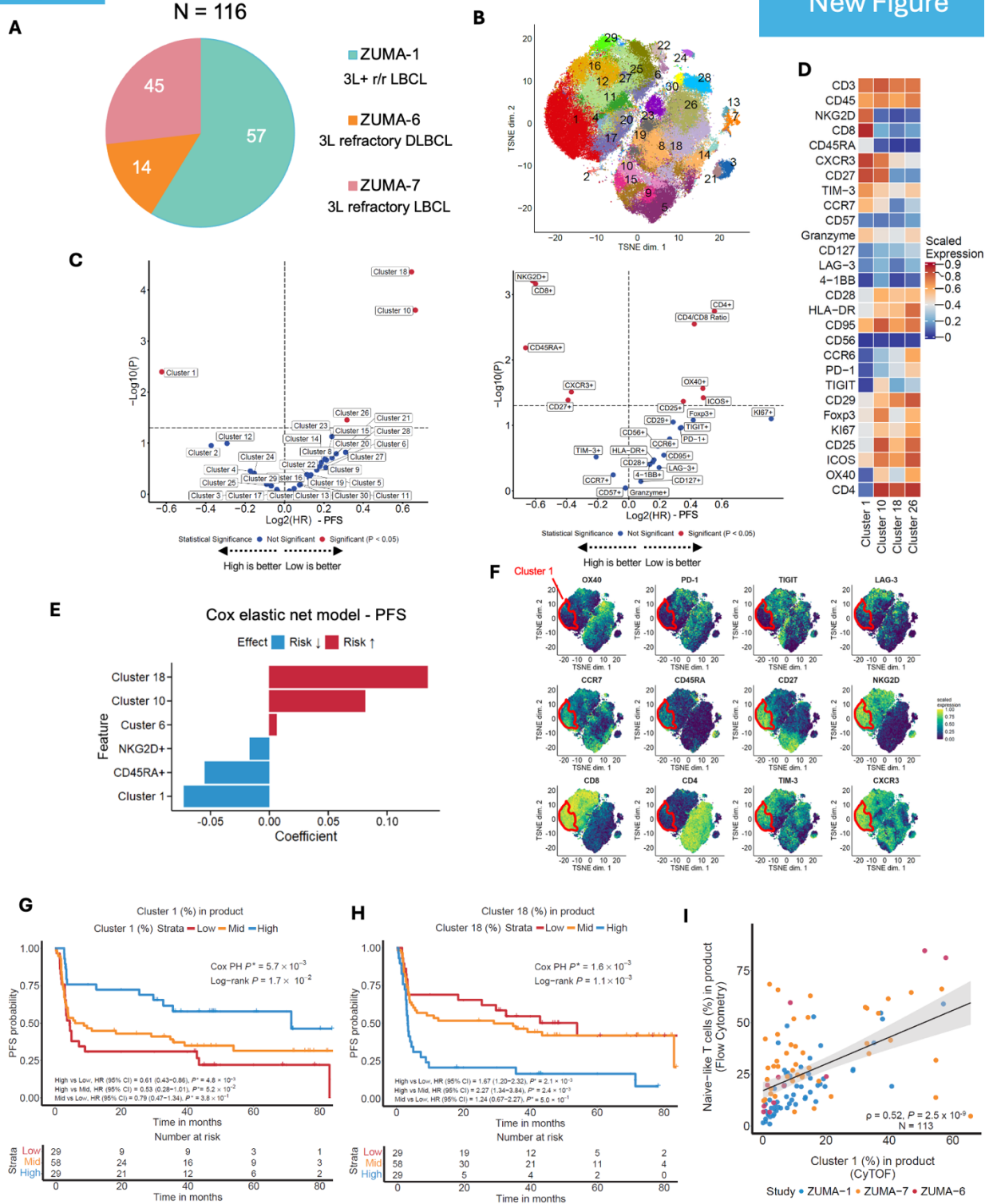
Supplementary Fig. 5

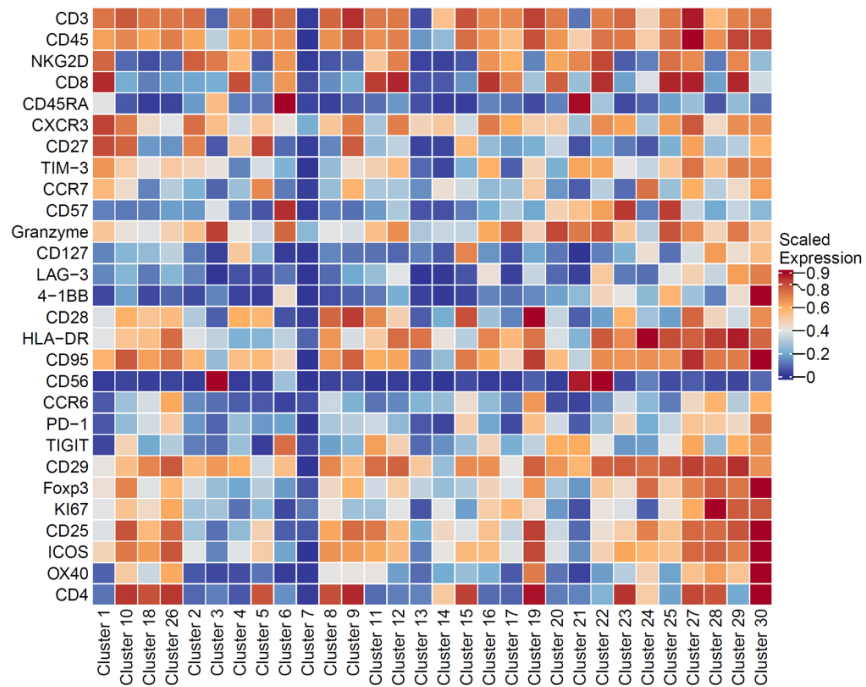
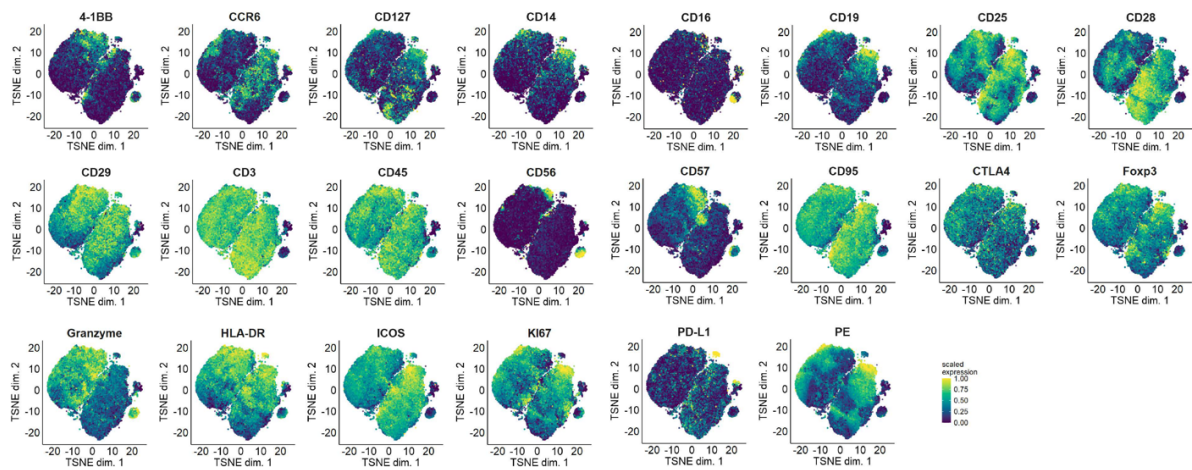
New Figure

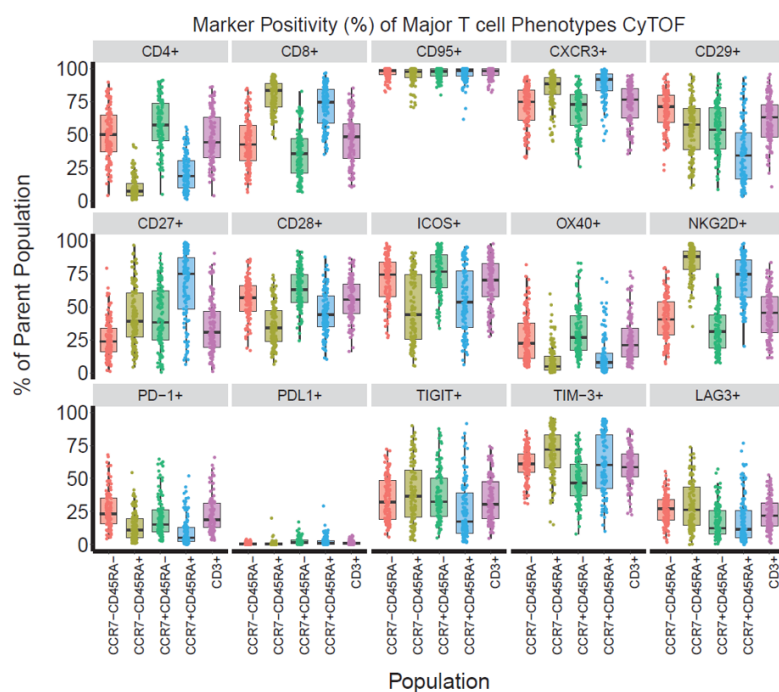
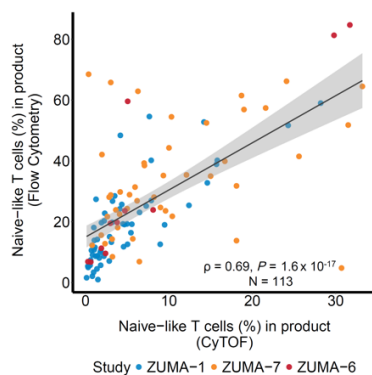
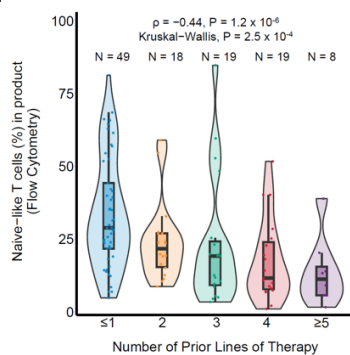
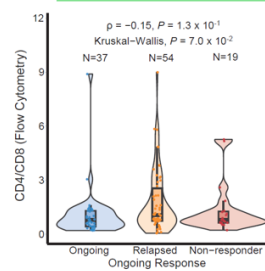
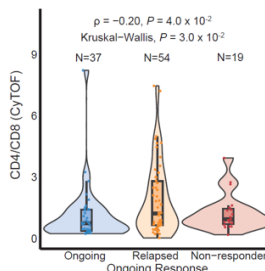
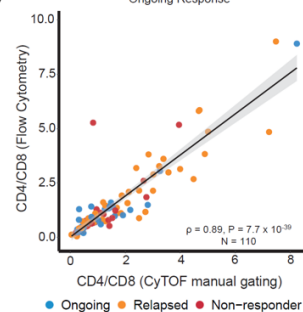
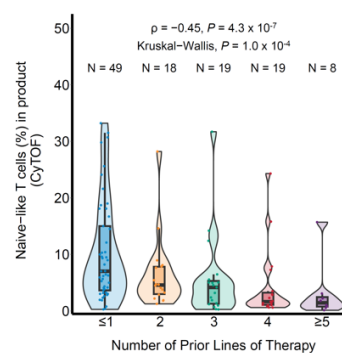


	Marker	Reactivity	Clone	Conjugate	Supplier
1	CD45	Human	HI30	89Y	Standard BioTools
2	HLA-DR	Cross	L243	146Nd	Biolegend
3	CD95	Human	DX2	164Dy	Standard BioTools
4	CD3	Human	UCHT1	154Sm	Standard BioTools
5	CD4	Human	SK3	174Yb	Standard BioTools
6	CD8	Human	SK1	168Er	Standard BioTools
7	CD25	Human	2A3	149Sm	Standard BioTools
8	CD127	Human	A019D5	143Nd	Standard BioTools
9	CD45RA	Human	HI100	169Tm	Standard BioTools
10	CCR7	Human	G043H7	159Tb	Standard BioTools
11	PD-1 (CD279)	Human	EH12.2H7	155Gd	Standard BioTools
12	CD19	Human	H1B19	142Nd	Standard BioTools
13	CD28	Human	CD28.2	160Gd	Standard BioTools
14	CD27	Human	L128	167Er	Standard BioTools
15	CD14	Human	M5E2	115In	Biolegend
16	CD16	Human	3G8	209Bi	Standard BioTools
17	CD56	Human	HCD56; NCAM16.2	176Yb	Standard BioTools
18	NKG2D (CD314)	Human	1D11; ON72	166Er	BD Biosciences
19	ICOS (CD278)	Human	C398.4A	148Nd	Standard BioTools
20	PD-L1 (CD274)	Human	29E.2A3	175Lu	Standard BioTools
21	Anti-CAR	Cross	Kip-1	145Nd	Kite
		Cross	1D4-C5	145Nd	Standard BioTools
22	CD29	Human	TS2/16	156Gd	Standard BioTools
23	CXCR3 (CD183)	Human	G025H7	163Dy	Standard BioTools
24	4-1BB (CD137)	Human	4B4-1	173Yb	Standard BioTools
25	OX40 (CD134)	Human	ACT35	150Nd	Standard BioTools
26	LAG-3 (CD223)	Human	11C3C65	165Ho	Standard BioTools
27	TIM-3 (CD366)	Human	F38-2E2	144Nd	Biolegend
28	CCR6 (CD196)	Human	G034H3	152Sm	Biolegend
29	CD57	Human	HCD57	172Yb	Standard BioTools
30	TIGIT	Human	MBSA43	153Eu	Standard BioTools
	Intracellular/Signaling				
31	Foxp3	Human	259D/C7; PCH101	162Dy	Standard BioTools
32	Ki67	Cross	B56	161Dy	Standard BioTools
33	CTLA4(CD152)	Human	14D3	170Er	Standard BioTools
34	Granzyme	Cross	GB11	171Yb	Standard BioTools

Fig. 6



A**B**

A**E****F****B****C****D****G**

E. D. Fig. 4

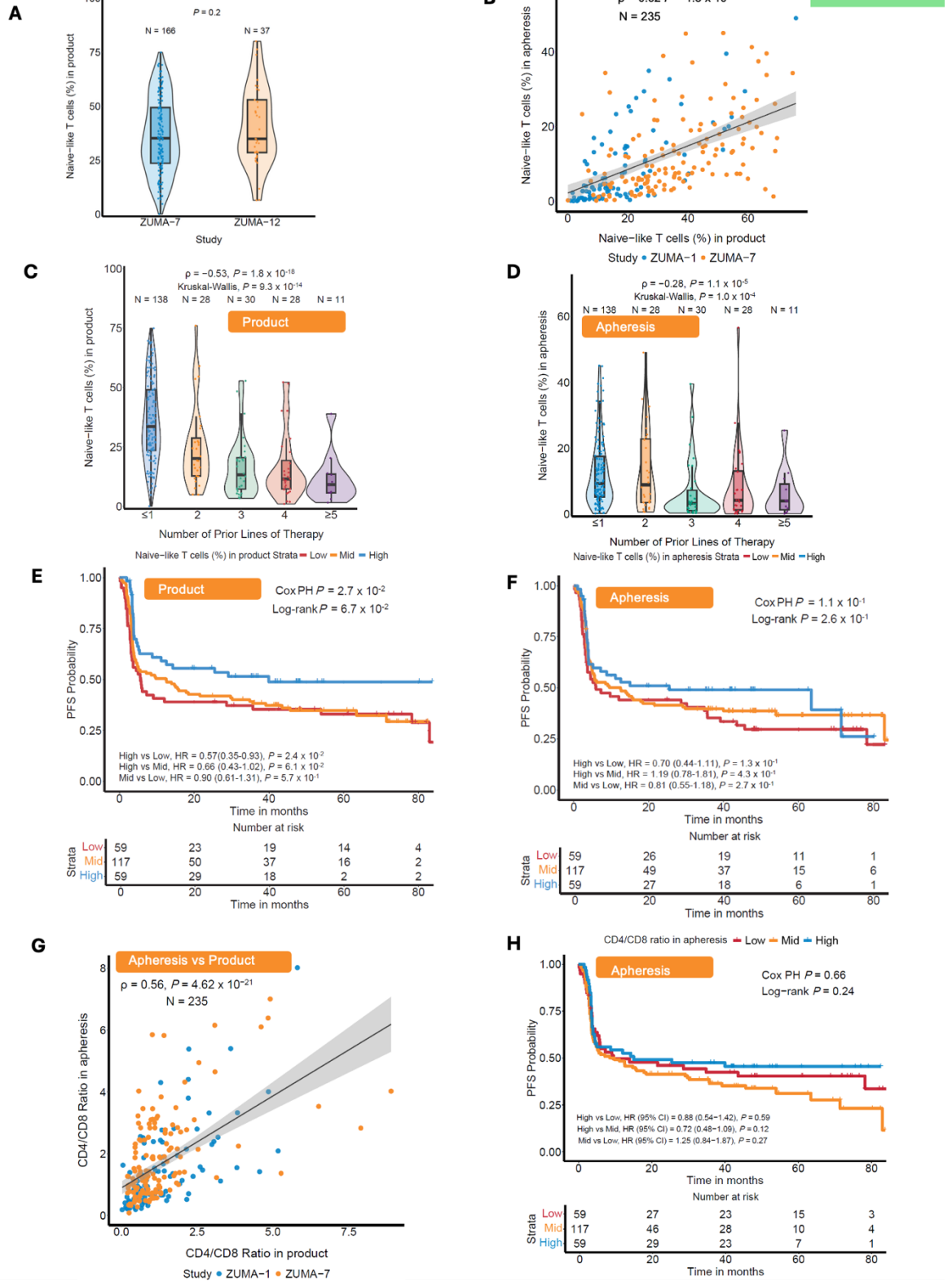
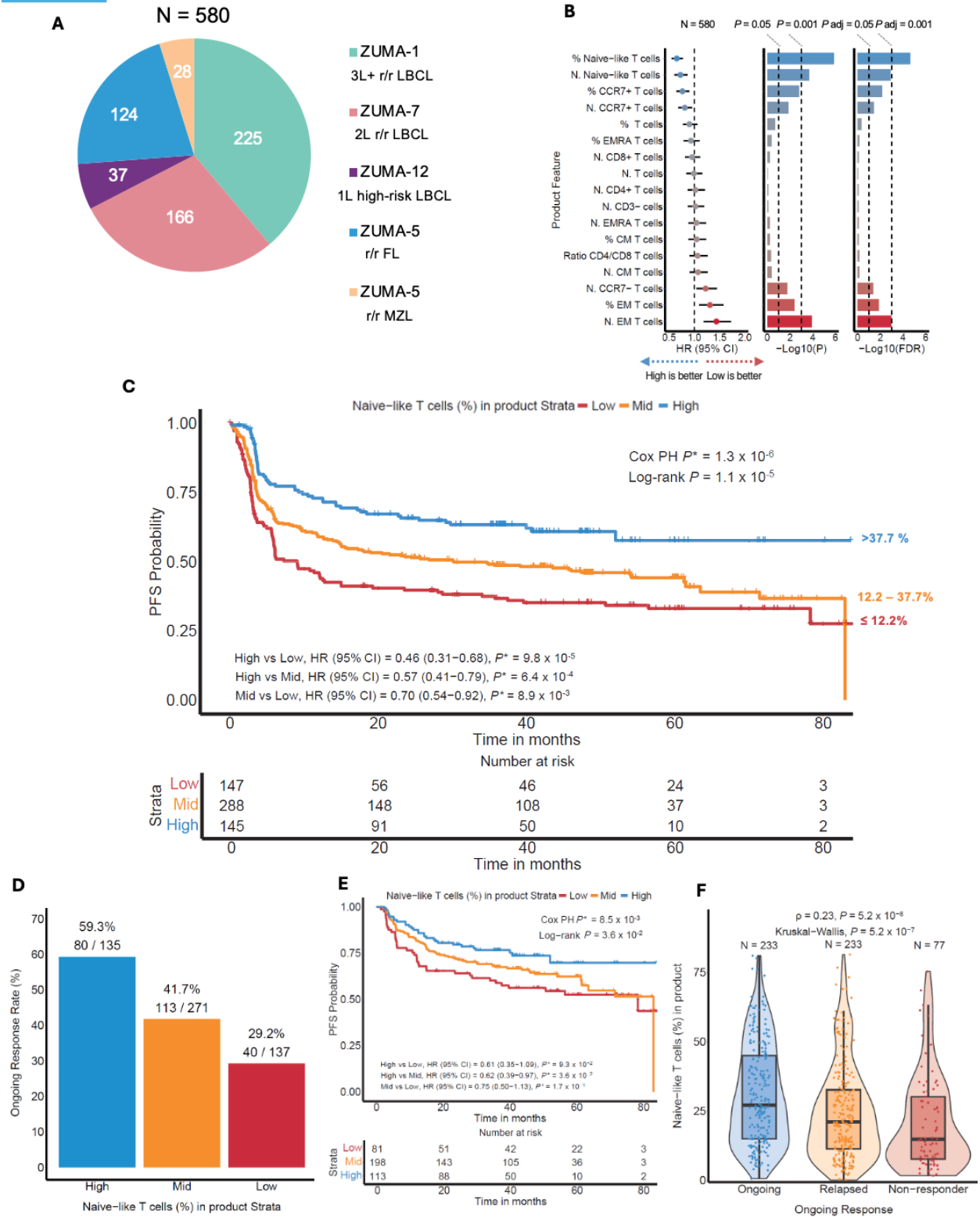
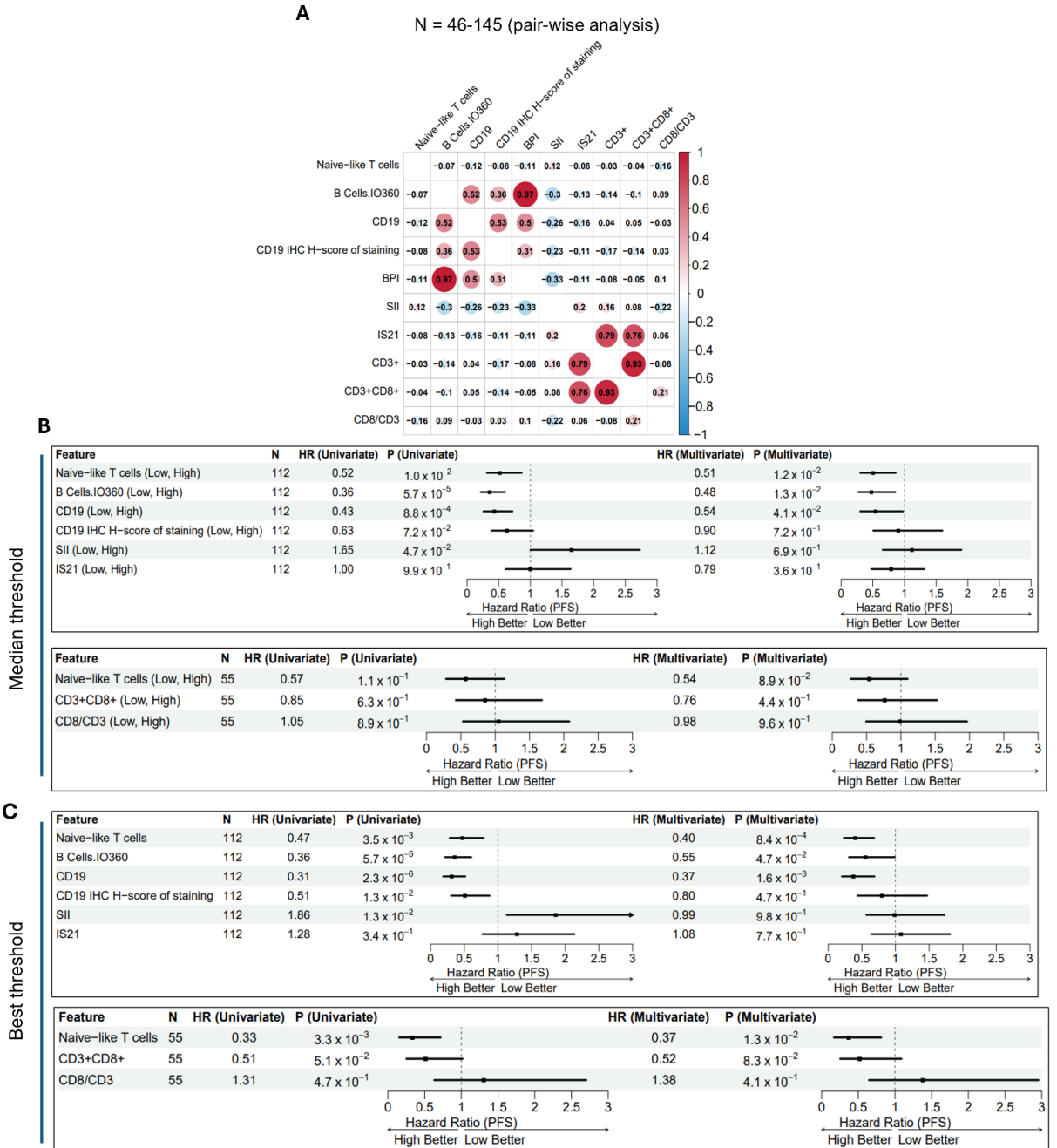


Fig. 1



1.3 - Prior Nature Medicine publications from the Kite group and collaborators have demonstrated that multiple tumor-related biomarkers, including the Immunoscore (CD3/CD8 IHC), IS21, B-cell GES, SII, BPI, CD19 expression (mRNA/protein), carry prognostic value for CAR T outcomes. To demonstrate the novelty and additive value of the juvenile T-cell fraction, when available (e.g., ZUMA1, 7), please compare juvenile T cells directly to these established predictors both individually and in combination. This benchmarking is essential to contextualize the juvenile fraction within the existing CAR T prognostic landscape and to dramatically increase the impact of this work.

In response to the reviewer's comment, we have performed correlative and multivariable analyses of proportion of product naive-like T cells with previously published prognostic co-variables pertinent to tumor microenvironment (TME). Considering that only 29 pre-treatment tumor biopsies were analyzed via nanostring in ZUMA-1 (Schollar, et. al. *Nature Medicine*, 2022), that the nanostring panels employed in ZUMA-1 and ZUMA-7 were different (pan-cancer vs IO360), with B cell, BPI and SII signatures not available in ZUMA-1 and further that product naive-like T cells does not show a significant association with outcome limited to the above 29 ZUMA-1 patients, for the analyses presented herein we leveraged the ZUMA-7 study, where both product phenotype and TME biomarkers were available in an adequate number of matching patients (Locke, et. al., *Nature Medicine*, 2024; Filosto, et. al., *Blood Cancer Discov.*, 2024). The results are presented in **Extended Data Fig. 6** of the revised manuscript. The correlogram in **Extended Data Fig. 6A** shows that naive-like T cells do not correlate with tumor B-cell GES, SII, BPI, CD19 expression (by mRNA or protein H-score), IS21 or the CD3 or CD8 cell densities (by IHC) or the CD8/CD3 density ratio. This is indicative of independent prognostic values of the product T-naive phenotype and the TME features. 112 patients in ZUMA-7 presented matching, non-missing naive-like T cells, nanostring-based GES and CD19 H-score values, while only 43 patients had all those co-variables with available CD3 and CD8 cell densities. Not considering the GES, 55 patients had naive-like T cells and CD3 and CD8 cell densities. Given the disparities in subject numbers by feature type, we performed two multivariate analyses (**Extended Data Fig. 6BC**): the first one included the percentage of naive-like T cells, the nanostring-based GES and CD19 H-score, and the second included naive-like T cells and the CD3 and CD8 cell density metrics. In the multivariate analyses, a co-linearity exclusion of variables was applied based on spearman $\rho \geq 0.7$. BPI was excluded from the multivariate analysis due to high correlation with B cell GES (spearman $\rho = 0.97$). In the second multivariate analysis, CD3 density was excluded due to high correlation with CD8 density (spearman $\rho = 0.93$). The percentage of naive-like T cells was associated with PFS regardless of TME features. The results were consistent whether dataset median (**Extended Data Fig. 6B**) or best-separation thresholds (**Extended Data Fig. 6C**) were utilized. B cell GES and CD19 expression were also significantly associated with PFS in the multivariate analyses. The above corroborated the independent prognostic value of the product naive-like T cell phenotype, regardless of the TME features, as well as the prognostic value of B cell GES and CD19 expression irrespective of product naive-like T cells, IS21 and SII.



Reviewer #2 (Remarks to the Author)

Thank you for the opportunity to review this manuscript. In this study, Budka and Davis et al. analyzed data from 580 patients with non-Hodgkin lymphoma treated with axicabtagene ciloleucel (axi-cel) to identify modifiable product characteristics linked to long-term outcomes. A high proportion of naive (CCR7⁺CD45RA⁺) T cells in the CAR T-cell product was strongly associated with improved progression-free survival (PFS). This association remained significant even after adjusting for other clinical and biological factors, including tumor burden, inflammatory markers, and CAR T-cell kinetics. Importantly, a high proportion of the naive T-cell phenotype did not correlate with increased toxicity. The study also found that more prior lines of therapy were associated with fewer naive T cells, underscoring the potential value of earlier CAR T-cell therapy. This large multi-cohort study confirms and builds on findings from smaller prior studies. The manuscript is well-written and supported by comprehensive, standardized data.

We thank the reviewer for the careful evaluation of our manuscript and for highlighting both the strengths and limitations of the study. We particularly appreciate the recognition that this work substantially strengthens the evidence linking naive-like T-cell content in the CAR T-cell product with long-term clinical outcomes. We agree that the principal contribution of this study lies in the strength of the evidence provided. Indeed, to our knowledge, this represents by far the largest and most comprehensive analysis of the relationship between product composition and clinical outcomes following CAR T-cell therapy. The revised manuscript has been further strengthened by the addition of high-dimensional CyTOF characterization of the infusion products, providing additional biological context for the observed associations as well as the validation in an additional meta-cohort of patients. The reviewer's comments prompted several additional analyses and revisions that have improved the rigor, clarity, and overall impact of the manuscript.

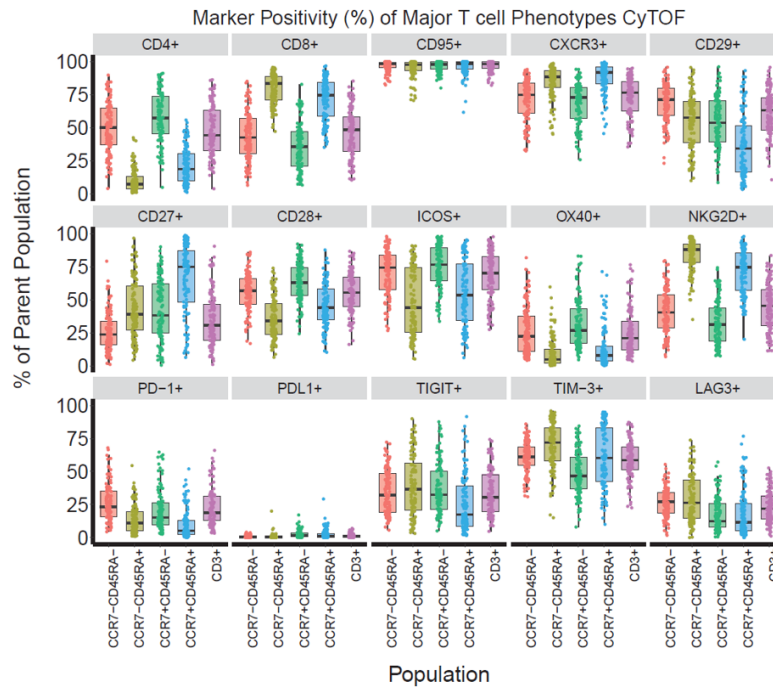
Comments:

2.1 - The term "juvenile" T cells is somewhat unconventional. Given the limited immunophenotyping and the absence of an anti-CD95 antibody necessary to identify stem cell-like T cells, I believe using the term "naive" or "naive-like" T cells would be more accurate and appropriate.

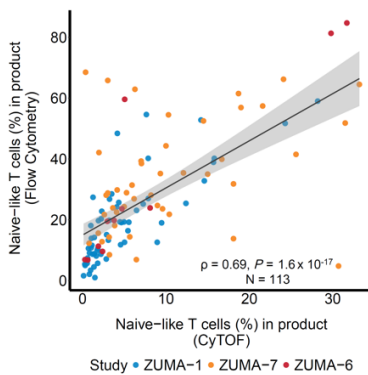
We thank the reviewer for these important comments. The markers used (CD45⁺ CCR7⁺) are shared by both naive T cells and memory stem cells (TSCM), which initially led us to use the term "**juvenile**". We also recognize that common markers used to differentiate T cell subsets might not apply to cells that undergo stimulation(s), such as the ones included in manufacturing of the product. At the same time, we recognize that this terminology (juvenile) is unconventional. In the revised manuscript, "**naive-like**" T cells is replacing "**juvenile**" T cells to avoid confusion.

To address Reviewer 1's comment on additional phenotyping markers and Reviewer 3's question regarding preservation of naive or TSCM subsets, we evaluated CD95 expression on 116 infused product samples from ZUMA-1, ZUMA-6, and ZUMA-7 using CyTOF data. CD95 was highly expressed across all T cell phenotypes (**Extended Data Fig. 9A**), reflecting ex vivo activation during CAR T cell manufacturing (**see also response 3.2**). This finding indicates that distinguishing TSCM-like from other subsets is not feasible in this context, supporting the use of "naive-like" rather than "juvenile" to describe this population.

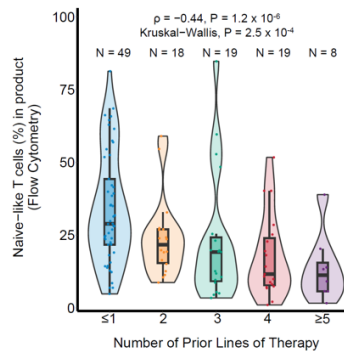
A



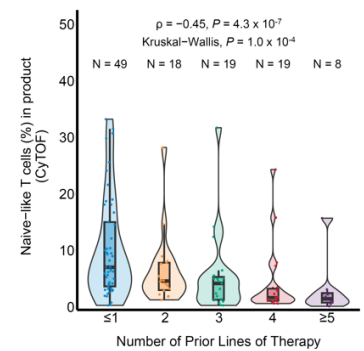
E



F

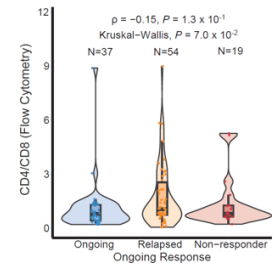


G

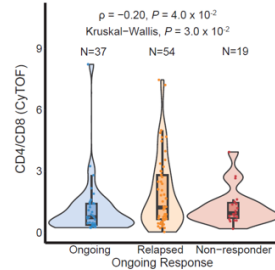


New Figure

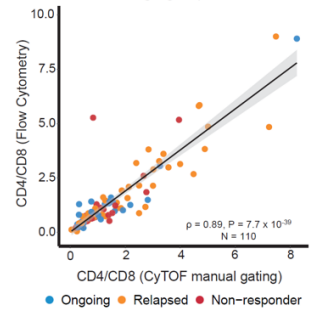
B



C



D



2.2 - The claim that the proportion of juvenile T cells in the product is the strongest predictor of PFS may be misleading if it is based solely on the P-value. For example, in the first multivariable analysis, the effect size of the proportion of juvenile T cells is similar to that of LDH, and in the second model, peak CAR T-cell expansion shows a larger effect size. To advance understanding in the field, it would be helpful to clarify the relative contributions of key factors influencing complete response and PFS—such as LDH, CAR T-cell kinetics, and the proportion of juvenile T cells in the product. Notably, Figure 4A suggests that a high proportion of juvenile T cells may not fully mitigate the adverse impact of high tumor burden. For the multivariable Cox regression models, it might be helpful to use the Wald statistic to rank predictor importance and perform variable drop-out analyses to compare the model's C-index. Please also provide Cox

regression modeling using continuous variables as supplementary data to help estimate the effect size of each variable.

The P values provided represent the result of the Wald test for each coefficient in the Cox regression model. The Wald statistic is the estimated beta coefficient divided by the coefficient's standard error, this value is used to derive P values from a chi-squared distribution. This indicates that the ranking by P value is directly related to the ranking by the Wald statistic, which supports the claim that naive-like T cells are a top pre-treatment feature of PFS. We extracted the model coefficients and utilized ANOVA testing of the fitted model to get the log-likelihood based rankings and additionally performed a C-index based dominance analysis which captures the idea of adding/removing variables to identify the most important features in a multivariable model (https://cran.r-project.org/web/packages/dmir/vignettes/dmir_basics.html).

In brief, the dominance analysis generates 2^P sub-models, where P is the number of independent variables in the model, and records the fit statistic for each model (in this case the C-index). The dominance rankings provided are obtained by exploring all sub-model results and determining the relative impact on the model fit statistic of each independent feature relative to others features in the model, across all possible orders and model sizes. Naive-like T cells (categorical) was ranked as the top feature associated with PFS with or without the inclusion of CAR T expansion in the axi-cel meta-cohort (**Fig. 2F and Fig. 4D**) and was 2nd ranked feature, behind baseline CRP, in the validation meta-cohort analysis (**Fig. 5F**) via dominance analysis.

Additionally, we have included Cox regression modeling using continuous variables in the multivariate analysis, with (**Supplementary Fig. 2A**) and without CAR T expansion (**Supplementary Fig. 2B**) to better estimate effect sizes.

Overall, **Fig. 4A** (**Fig. 3C** in the revised manuscript) was intended to be descriptive, whereas the statistical significance and relevance of the association with clinical parameters were assessed through multivariable modeling. It should also be noted that the number of patients in the tumor burden *High* group was relatively small after stratification into the *Low*, *Mid*, and *High* categories, which may have limited the ability to detect significant differences in this group (as a note, in this subgroup analyses we found appropriate to indicate only the overall P values by log-rank test rather than the P values from a cox regression comparing each T naive-like strata). Moreover, although SPD is a measure of tumor burden, it has inherent limitations. Other parameters, such as LDH and ferritin, are also influenced by tumor burden, and stratification based on these markers revealed a more pronounced separation of naive-like T cell levels within the high-risk groups. Furthermore, a clear separation according to naive-like T-cell strata was observed in patients with High SPD in the validation cohort (**Extended Data Fig. 3C**, $P = 3.4 \times 10^{-3}$).

Fig. 2

New Panel (D, E)

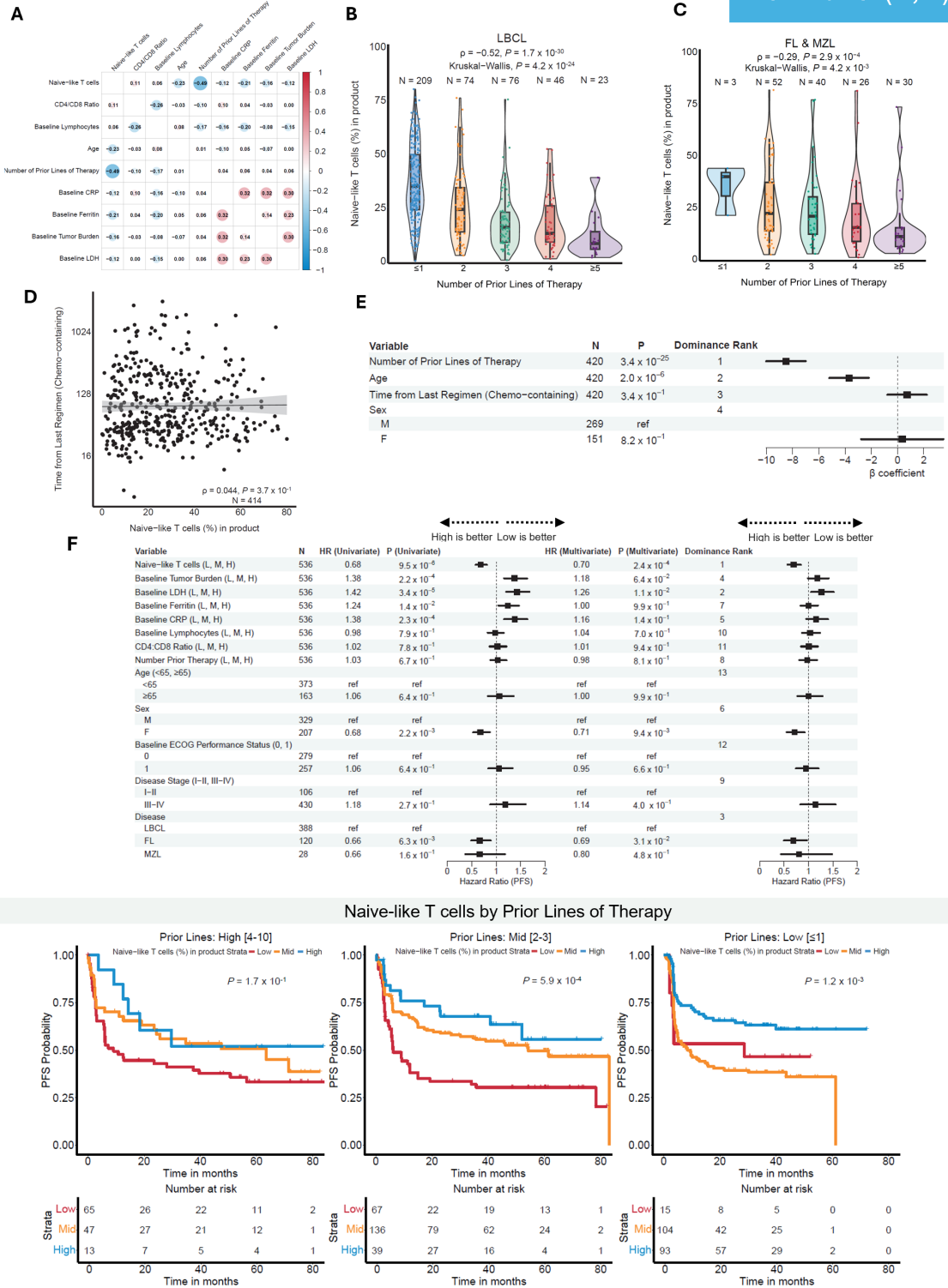
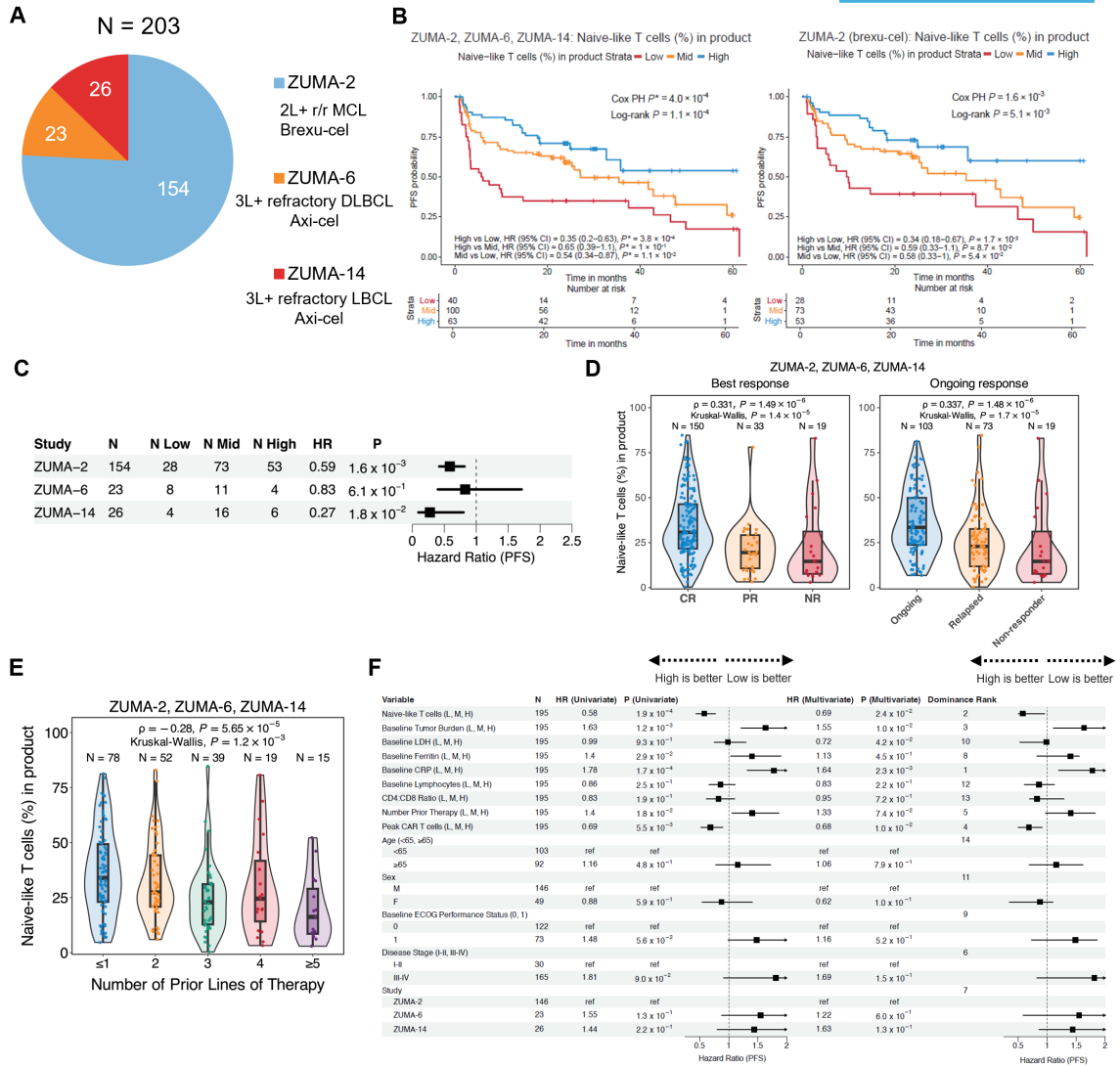


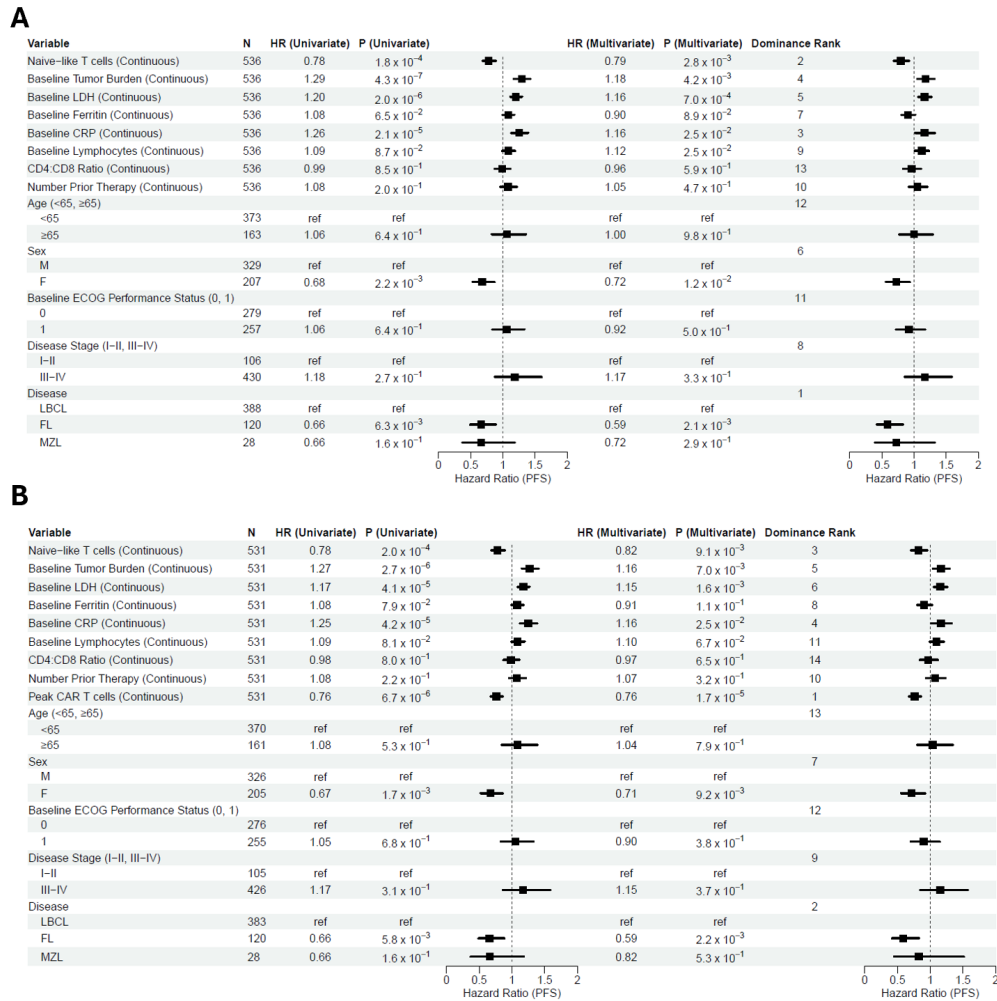
Fig. 4



Fig. 5

New Figure





2.3 - Given the size of the dataset, it would be informative to explore the ratio of CAR T-cell expansion and/or proportion of juvenile T cells to tumor burden and/or LDH and its association with CR and PFS.

Thank you for the comment, these ratios have been explored in previous publications and represent intriguing metrics that could provide insight into the relationship of effector to target ratio and clinical responses. The naive-like T cells to tumor burden (TB) ratio could furthermore represent a pre-treatment predictor of response. Univariate and multivariate Cox proportional hazard regression analysis of PFS (**Fig. 4D**) revealed opposing associations, with higher levels of naive-like T cells in product, or greater CAR T-cell peak expansion, correlated with improved PFS, as compared to baseline tumor burden. These findings further support the idea that simple ratios could demonstrate a strong association with PFS.

Kaplan-Meier PFS analysis was used to evaluate the naive-like T cells / TB ratio, with the ratio converted into categorical variables (*Low*, *Mid*, *High*). The log-rank P value for categorical comparisons and Cox proportional hazards (PH) P value for ordinal (*Low* = 1, *Mid* = 2, *High* = 3) is provided, with study included as a covariate. P values, HRs, and 95% CIs are reported for pairwise comparisons between strata, utilizing Cox proportional hazards modeling with study as cohort included as a covariate in the model (**Supplementary Fig. 4A**). A Cox proportional

hazard (PH) P value of 1.3×10^{-8} and log-rank P value of 1.0×10^{-8} was observed. Similarly, a higher naive-like T cells / TB ratio was identified in patients who had a CR as their best response (**Supplementary Fig. 4B**) and in subjects with ongoing response (**Supplementary Fig. 4C**), with significance determined using both spearman and Kruskal-Wallis testing.

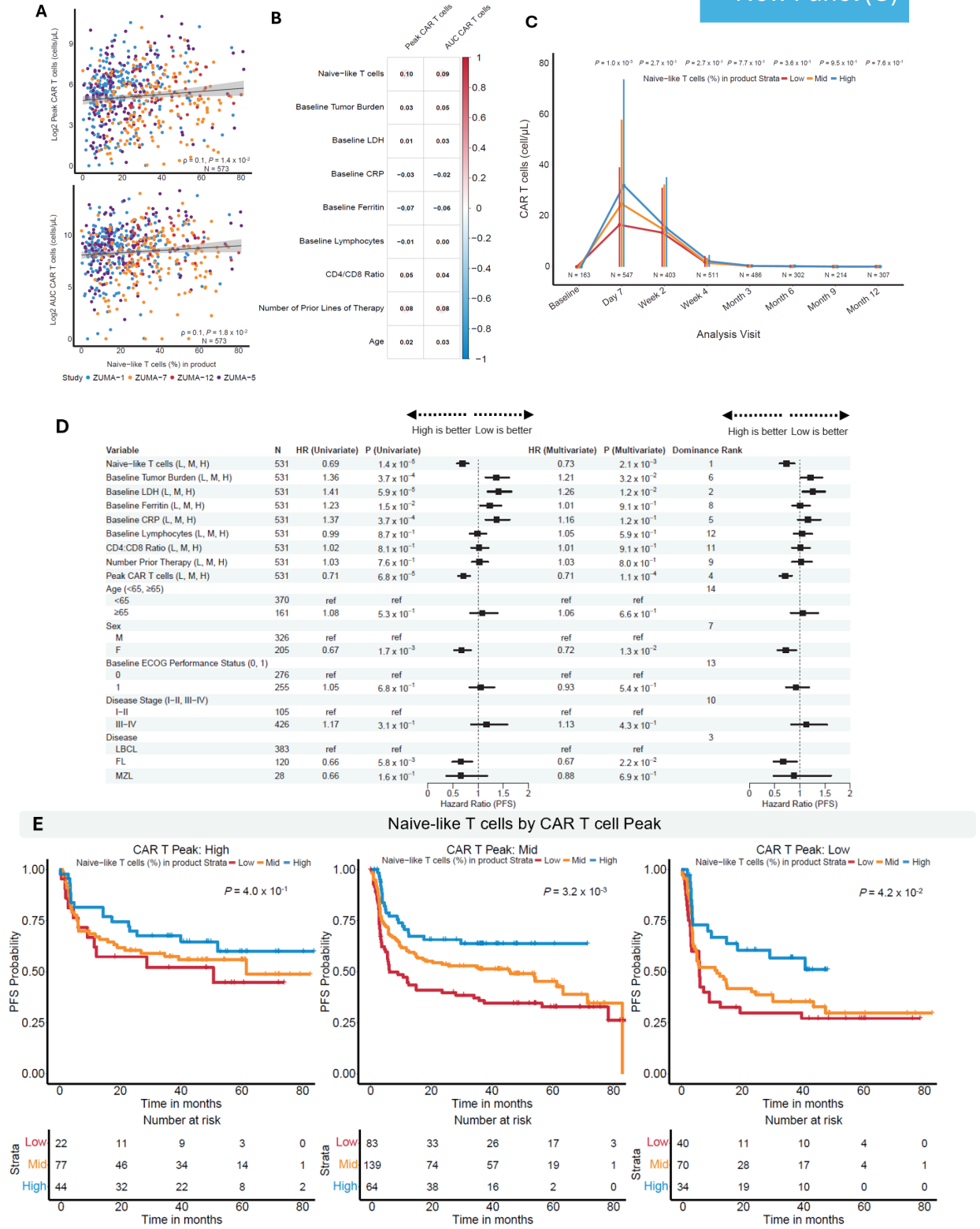
The same analyses as described above for naive-like T cells / TB were performed using the CAR T peak / TB ratio. A Cox PH P value of 1.2×10^{-9} and log-rank P value of 2.6×10^{-10} for PFS was observed for this metric (**Supplementary Fig. 4D**). Similarly, a higher CAR T peak/TB ratio was present in patients who had a CR as their best response (**Supplementary Fig. 4E**) and in patients that were ongoing responders (**Supplementary Fig. 4F**), with significance identified with both spearman and Kruskal-Wallis testing.

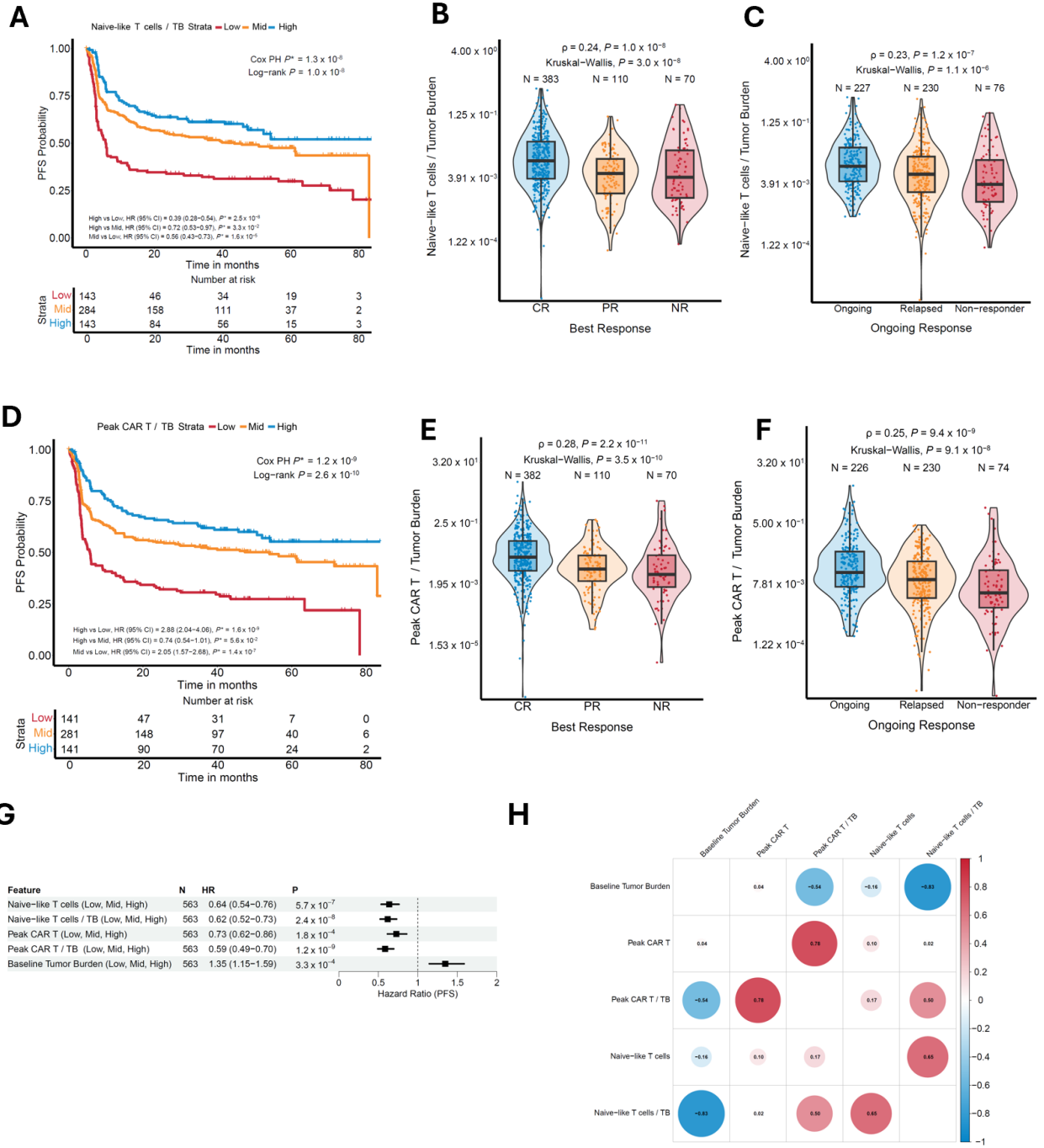
Univariate PFS Cox proportional hazards model, with study included as a covariate, for baseline tumor burden, the proportion of naive-like T cells in the product (original and normalized to tumor burden), and the peak CAR T cell expansion (original and normalized to tumor burden) was performed. We found that all features associated with PFS and naive-like T cells had similar hazard-ratios with or without tumor burden normalization while peak CAR T cell expansion demonstrated improvement in hazard-ratio with tumor burden normalization (**Supplementary Fig. 4G**). A pairwise spearman correlation analysis of these features shows that baseline tumor burden is inversely correlated with both ratio-based metrics (naive-like T cells per tumor burden and CAR T peak per tumor burden), and that the original representation of naive-like T cells and CAR T peak expansion correlate strongly with their respective ratios (**Supplementary Fig. 4H**).

Overall, the association between CAR T cell expansion and efficacy strengthened when CAR T peak levels were normalized to TB (i.e., by estimating effector-to-target cell ratios). However, this normalization had a marginal impact on the effect of the naive-like T cell estimates on PFS, consistent with the observation that the benefit of naive-like T cells was maintained across TB strata (**Supplementary Figure 4A–H**). These findings also suggested that the influence of naive-like T cells on PFS is not fully captured by peripheral CAR T cell expansion, and might reflect intrinsic properties of these cells, such as intratumoral proliferation, that might overcome an initially unfavorable effector to target cell ratio.

Fig. 4

New Panel (C)





2.4 - The more pronounced differences in the proportion of juvenile T cells when comparing ongoing response rate rather than CR rates may suggest a role of CAR T-cell persistence in sustaining responses. Do you have data on the relationship between the proportion of juvenile T cells in the product and CAR T-cell persistence (e.g., time to last detection, AUC0-60, AUC0-90, etc.) or duration of B-cell aplasia?

The progressively shorter PFS observed with lower juvenile T-cell proportions in the Mid and Low CAR T-cell expansion groups suggests that peak CAR T-cell expansion may be the primary driver of initial response, while a higher proportion of juvenile T cells—due to their greater self-renewal capacity—may contribute to improved persistence (AUC in the first 6 months) and, consequently, better PFS.

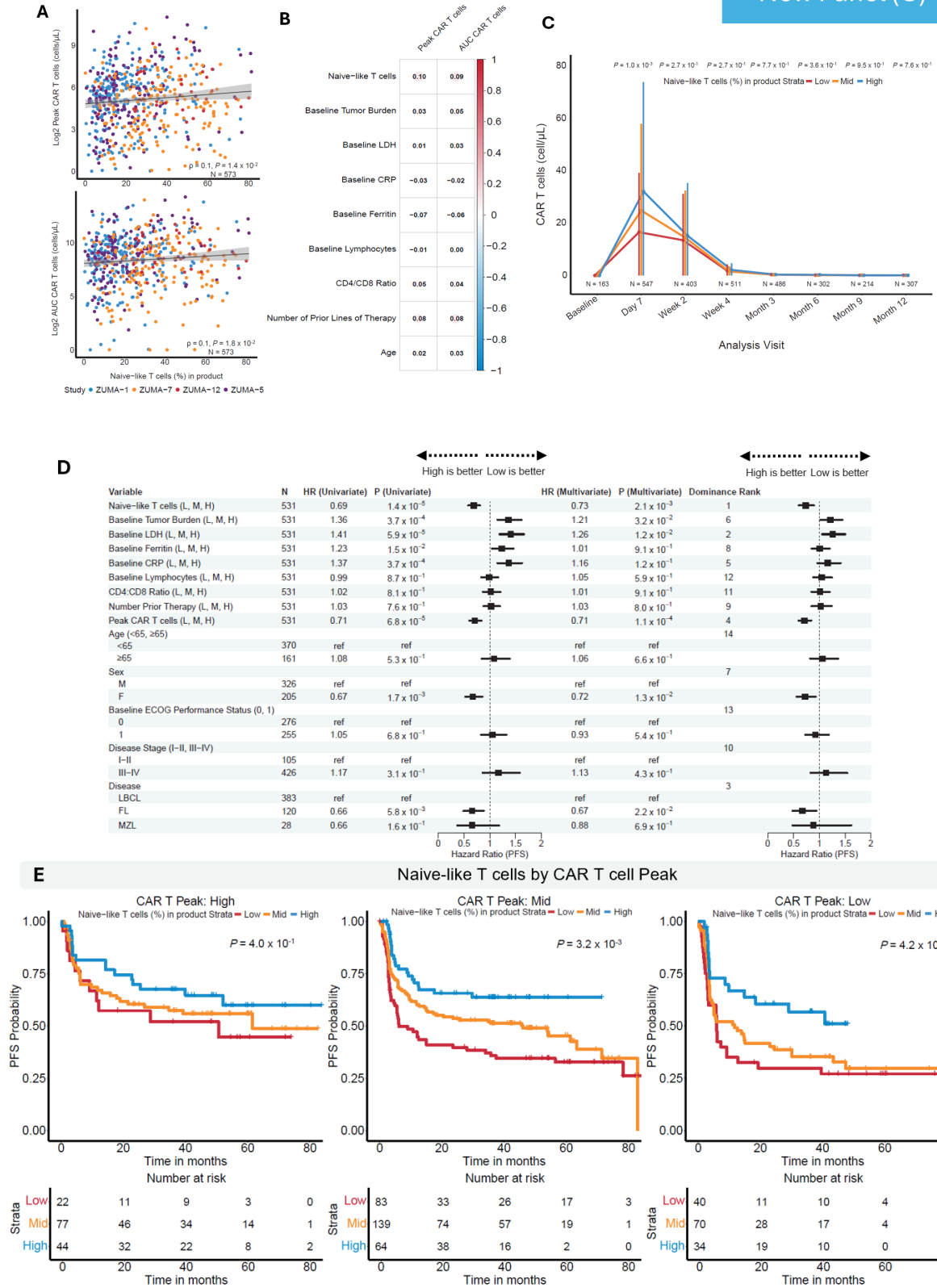
We analyzed CAR T cell concentrations (cells/ μ L of blood) from baseline through Month 12 stratified by the proportion of naive-like T cells in the infused product (*Low* = $\leq 25^{\text{th}}$ percentile, *Mid* = 25^{th} – 75^{th} percentile, *High* = $>75^{\text{th}}$ percentile) to evaluate potential differences in CAR T cell kinetics and persistence (**Fig. 4C**). For each time point, median $\pm$ median absolute deviation of CAR T cell concentrations were reported and Kruskal–Wallis testing was performed to identify statistical differences across naive-like T cell groups. Patients with a higher proportion of naive-like T cells exhibited significantly greater CAR T cell expansion on Day 7 and demonstrated a trend of higher expansion at Week 2, but no clear differences were observed beyond this point (Week 4, Months 3, 6, 9, or 12).

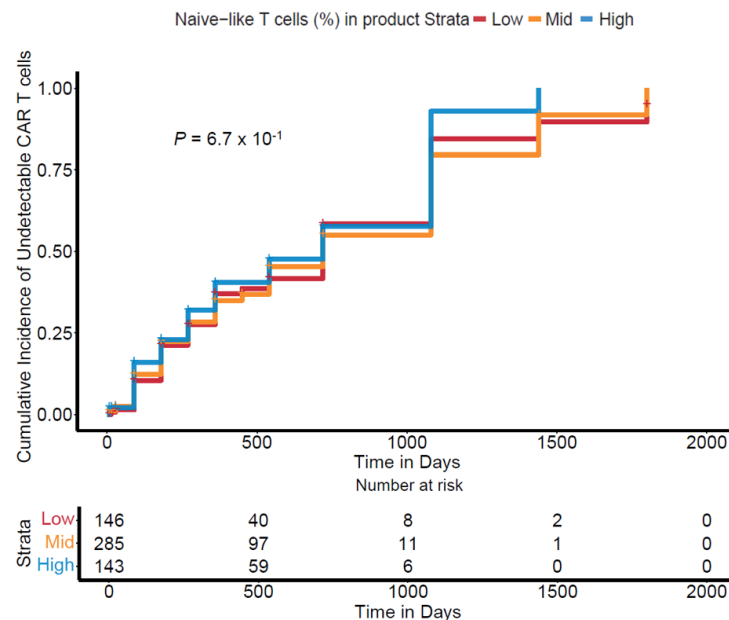
To further assess the relationship between naive-like T cells and CAR T cell persistence, we performed a cumulative incidence analysis of CAR T cell undetectability by naive-like T cell groups (**Supplementary Fig. 3**). CAR T cell levels were classified as undetectable once a value of 0 cells/ μ L of blood was captured, with no additional detectable occurrences, while subjects were censored if CAR T cell concentrations were detectable through the end of all available collections. Utilizing this approach, no significant differences were observed for the time to CAR T cell undetectability by naive-like T cell strata (log-rank $P = 0.67$).

These findings suggest that CAR T cell expansion is influenced by the proportion of naive-like T cells in the product during the first two weeks post-infusion, but beyond this point there is no clear demarcation of CAR T cell levels to support peripheral blood long-term persistence differences driven by naive-like T cell proportions in the product. However, we cannot rule out the possibility that naive-like T cells exhibit greater intra-tumoral proliferation or persistence, parameters that are not captured by assessing CAR T cells in peripheral blood. This point is added to the discussion.

Fig. 4

New Panel (C)





2.5 - Please provide a table summarizing the baseline characteristics across the different axi-cel studies. If available, a breakdown of prior lines of therapy into chemotherapy-containing versus targeted therapy regimens would be helpful. Additionally, it would be informative to assess the correlation between time from last treatment—particularly cytotoxic chemotherapy—and the proportion of juvenile T cells.

We have included a demographics table for the axi-cel meta-cohort (**Supplementary Table 1**) and the validation meta-cohort (**Supplementary Table 3**), which includes the percentage of subjects who received cytotoxic chemotherapy and/or targeted/biologic regimens in any prior line of therapy. Prior to enrollment, physicians document the patient's previous lines of therapy but do not always provide the treatment schedule of individual drugs within combination chemoimmunotherapy regimens (e.g., R-CHOP). As combination therapies are commonly used in the treatment of NHL, and due to the limited nature of the captured prior regimen data, it is often challenging to obtain the timing of individual agents. Although we could not fully delineate the exact timing of specific drugs within prior combination regimens, we explored the relationship between the number of prior lines of therapy and proportion of naive-like T cells in patients based on whether they did or did not receive a prior targeted or biologic (not including anti-CD20) regimen. Additionally, we assessed whether time to last therapy was associated with naive-like T cell proportion when the last regimen prior to leukapheresis contained a cytotoxic chemotherapy agent.

We evaluated the proportion of naive-like T cells in the product by prior lines of therapy in patients that received a targeted or biologic (beyond anti-CD20 mAb) as a prior regimen (**Supplementary Fig. 1A**) and those who only received chemotherapy or anti-CD20 mAbs (**Supplementary Fig. 1B**) using Kruskal-Wallis for categorical comparisons and spearman correlation for ordinal associations, with prior lines of therapy treated as follows: ≤ 1 line = 1, 2 lines = 2, 3 lines = 3, 4 lines = 4, and ≥ 5 lines = 5. A significant negative correlation was observed between the number of prior lines of therapy and the proportion of naive-like T cells, regardless of targeted or biologic regimen administration. Notably, all patients received a prior chemotherapy-containing regimen, besides one patient from ZUMA-5 who received previous

therapy for DLBCL, not the primary disease of FL, and thus was excluded from the prior regimens analysis.

No correlation was observed between the time from last therapy (if last contained a cytotoxic chemotherapy) and the proportion of naive-like T-cells (**Fig. 2D**). When incorporating time from last cytotoxic chemotherapy-containing regimen into a multivariate linear regression model with the number of prior lines of therapy, age, and sex, the time from last cytotoxic chemotherapy-containing regimen was not significantly associated with PFS (**Fig. 2E**).

Supplementary Table 1

New Table

Variable	ZUMA-1 (N = 225)	ZUMA-5 FL (N = 124)	ZUMA-5 MZL (N = 28)	ZUMA-7 (N = 166)	ZUMA-12 (N = 37)	Overall (N = 580)
Age						
Median age (range)	59 (19-85)	59.5 (34-79)	64 (43-77)	59 (21-80)	60 (23-86)	60 (19-86)
Male						
n(%)	143 (64)	73 (59)	13 (46)	104 (63)	25 (68)	358 (62)
Race, n(%)						
White	190 (84)	115 (93)	25 (89)	136 (82)	30 (81)	496 (86)
Black or African	8 (4)	4 (3)	2 (7)	9 (5)	1 (3)	24 (4)
Hispanic or Latino	26 (12)	6 (5)	3 (11)	8 (5)	2 (5)	45 (8)
Asian	3 (1)	2 (2)	0 (0)	9 (5)	1 (3)	15 (3)
Disease stage, n(%)						
I or II	57 (25)	18 (15)	2 (7)	38 (29)	1 (3)	116 (20)
III or IV	167 (74)	106 (85)	26 (93)	128 (77)	36 (97)	463 (80)
ECOG						
0	104 (46)	78 (63)	15 (54)	91 (55)	15 (41)	303 (52)
1	121 (54)	46 (37)	13 (46)	75 (45)	22 (59)	277 (48)
Study Prognostic Score ^a, n(%)						
Low	126 (56)	70 (56)	8 (29)	94 (57)	6 (16)	304 (52)
High	99 (44)	54 (44)	18 (64)	72 (43)	31 (84)	274 (47)
Prior lines of therapy						
Low (≤1)	6 (3)	3 (2)	0 (0)	166 (100)	37 (100)	212 (37)
Mid (2-3)	150 (67)	74 (60)	18 (64)	0 (0)	0 (0)	242 (42)
High (4-10)	69 (31)	46 (37)	10 (36)	0 (0)	0 (0)	125 (22)
Disease type						
DLBCL	156 (69)	0 (0)	0 (0)	99 (60)	0 (0)	255 (44)
Other (PMBC, TFL, HGBCL, FL, MZL, etc.)	69 (31)	124 (100)	28 (100)	67 (40)	37 (100)	325 (56)
Previous ASCT						
n(%)	172 (76)	30 (24)	4 (14)	0 (0)	0 (0)	96 (17)
CD19 positivity, n(%)						
Negative	16 (7)	9 (7)	0 (0)	12 (7)	6 (16)	43 (7)
Positive	141 (63)	96 (77)	19 (68)	134 (81)	24 (65)	414 (71)
Refractory subgroup ^b						
Primary refractory	6 (3)			121 (73)		
Refractory to 2nd or greater line therapy	155 (69)					
Relapsed to 2nd or greater line therapy	15 (7)					
Relapse post ASCT	49 (22)					
Relapse ≤ 6 months of first-line therapy				26 (16)		
Relapse > 6 and ≤ 12 months of first-line therapy				19 (11)		
Refractory		84 (68)	22 (79)			
Relapsed		40 (32)	6 (21)			
Tumor burden by SPD ^c						
Median (range)	2971 (116-39658)	2790.145 (155.53-20264)	2034 (266-12327)	2508 (181-22892)	2707 (230-26124)	2768.3 (116-39658)
Naive-like T cells (% of viable CD3+ product cells)						
n	225	124	28	166	37	580
Mean (STDEV)	20.1 (14.7)	21.9 (17.2)	25.8 (19.9)	36.2 (17.2)	40.8 (18.3)	26.7 (18.2)
Median (Q1, Q3)	16.4 (8.8, 26.4)	15.4 (9.6, 29.0)	21.5 (10.9, 35.4)	32.3 (23.6, 49.5)	35.0 (28.6, 52.9)	23.1 (12.2, 37.8)
Min, Max	0.4, 76.0	0.6, 81.4	3.3, 76.1	0.1, 74.9	6.5, 80.1	0.1, 81.4
Durable response (Investigator Read)						
n(%)	73 (32)	50 (40)	15 (54)	68 (41)	27 (73)	233 (40)
Prior regimen group ^d						
Radiation	56 (25)					56 (10)
Cytotoxic Chemotherapy-Containing	225 (100)	123 (99)	28 (100)	166 (100)	37 (100)	579 (100)
Targeted or Biologic Regimens	224 (100)	123 (99)	28 (100)	166 (100)	35 (95)	576 (99)
Anti-CD20	224 (100)	123 (99)	27 (96)	166 (100)	35 (95)	575 (99)
Other Targeted or Biologic Regimens	65 (29)	74 (60)	21 (75)	3 (2)		163 (28)

^a Study Prognostic Scores are defined as follows: For ZUMA-1, ZUMA-5, and ZUMA-12, scores of 0–2 are considered low, and 3+ are high. For ZUMA-7, scores of 0–1 are low, and 2+ are high. The International Prognostic Index (IPI) score was used in ZUMA-1 and ZUMA-12. The IPI score used in ZUMA-7 is based on Second-line age-adjusted International Prognostic Index total score (IXRS). Follicular Lymphoma International Prognostic Index (FLIPI) score was used in ZUMA-5.

^b The refractory subgroups in the ZUMA trials are defined as follows: ZUMA-1 includes four categories—Primary refractory, Refractory to 2nd or greater line therapy, Relapsed to 2nd or greater line therapy, and Relapse post ASCT. ZUMA-5 categorizes patients as either Refractory or Relapsed. ZUMA-7 defines refractory subgroups based on the derived response to first-line therapy at randomization, with three categories—Primary refractory, Relapse within 6 months of first-line therapy, and Relapse between 6 and 12 months after first-line therapy. ZUMA-12 does not include refractory subgroup information.

^c Investigator determined SPD for all subjects was used when available. If unavailable, Central SPD was used.

^d One ZUMA-5 subject received previous therapy for DLBCL, not for the primary disease of FL, thus the total number of subjects in the meta-cohort for prior regimens is 579.

Variable	ZUMA-2 (N = 154)	ZUMA-6 (N = 23)	ZUMA-14 (N = 26)	Overall (N = 203)
Age				
Median age (range)	65 (38-82)	56 (43-71)	62 (38-81)	64 (38-82)
Male				
n(%)	124 (81)	13 (57)	14 (54)	151 (74)
Race, n(%)				
White	140 (91)	18 (78)	19 (73)	177 (87)
Black or African	3 (2)	0 (0)	0 (0)	3 (1)
Hispanic or Latino	16 (10)	0 (0)	4 (15)	20 (10)
Asian	1 (0)	4 (17)	3 (12)	8 (4)
Disease stage, n(%)				
I or II	22 (14)	5 (22)	5 (19)	32 (16)
III or IV	132 (86)	18 (78)	21 (81)	171 (84)
ECOG				
0	103 (67)	13 (57)	12 (46)	128 (63)
1	51 (33)	10 (43)	14 (54)	75 (37)
Study Prognostic Score ^a, n(%)				
Low		12 (52)	26 (100)	
High		11 (48)	0 (0)	
Prior lines of therapy				
Low (≤1)	69 (45)	4 (17)	5 (19)	78 (38)
Mid (2-3)	59 (38)	11 (48)	21 (81)	91 (45)
High (4-10)	26 (17)	8 (35)	0 (0)	34 (17)
Disease type				
MCL	154 (100)			
DLBCL		23 (100)	22 (85)	
Other (HGBCL, PMBCL, etc.)			4 (15)	
Previous ASCT				
n(%)	69 (45)	4 (17)	2 (8)	75 (37)
CD19 positivity, n(%)				
Negative	9 (6)	4 (17)		
Positive	113 (73)	17 (74)		
Refractory subgroup ^b				
Primary refractory		4 (17)	5 (19)	
Refractory to last MCL therapy	39 (25)			
Refractory after ASCT			2 (8)	
Relapsed post last MCL therapy	86 (56)			
Refractory to 2nd or greater line therapy		15 (65)	19 (73)	
Relapse post ASCT	29 (43)	4 (17)		
Tumor burden by SPD ^c				
Median (range)	1930 (0-58968)	2309 (180-18381.82)	2225.5 (187-14884)	2059.5 (0-58968)
Naïve-like T cells (% of viable CD3+ product cells)				
n	154	23	26	203
Mean (STDEV)	32.2 (19.1)	26.44 (22.33)	30.4 (17.5)	31.3 (19.3)
Median (Q1, Q3)	28.9 (19.0, 45.6)	19.5 (9.6, 33.4)	28.4 (20.1, 37.3)	27.7 (18.3, 44.2)
Min, Max	0.3, 80.7	4.9, 84.7	8.6, 83.0	0.3, 84.7
Durable response (Investigator Read)				
n(%)	85 (55)	7 (30)	11 (42)	103 (51)
Prior regimen group				
Radiation	16 (10)			16 (8)
Cytotoxic Chemotherapy-Containing	154 (100)	23 (100)	26 (100)	203 (100)
Targeted or Biologic Regimens	154 (100)	23 (100)	26 (100)	203 (100)
Anti-CD20	154 (100)	23 (100)	26 (100)	203 (100)
Other Targeted or Biologic Regimens	74 (48)	4 (17)	5 (19)	83 (41)

^a Study Prognostic Scores are defined as follows: scores of 0–2 are considered low, and 3+ are high. Mantle Cell Lymphoma International Prognostic Index (MIPI) was used in ZUMA-2, International Prognostic Index (IPI) in ZUMA-6, and age-adjusted IPI in ZUMA-14

^b The refractory subgroups in the ZUMA trials are defined as follows: ZUMA-2 includes three categories—Refractory to last MCL therapy, Relapsed post last MCL therapy, and Relapse post ASCT. ZUMA-6 includes three categories—Primary refractory, Refractory to 2nd or greater line therapy, and Relapse Post ASCT. ZUMA-14 includes three categories—Primary refractory, Refractory to 2nd or greater line therapy, and Refractory after ASCT.

^c Investigator determined SPD for all subjects was used when available. If unavailable, Central SPD was used.

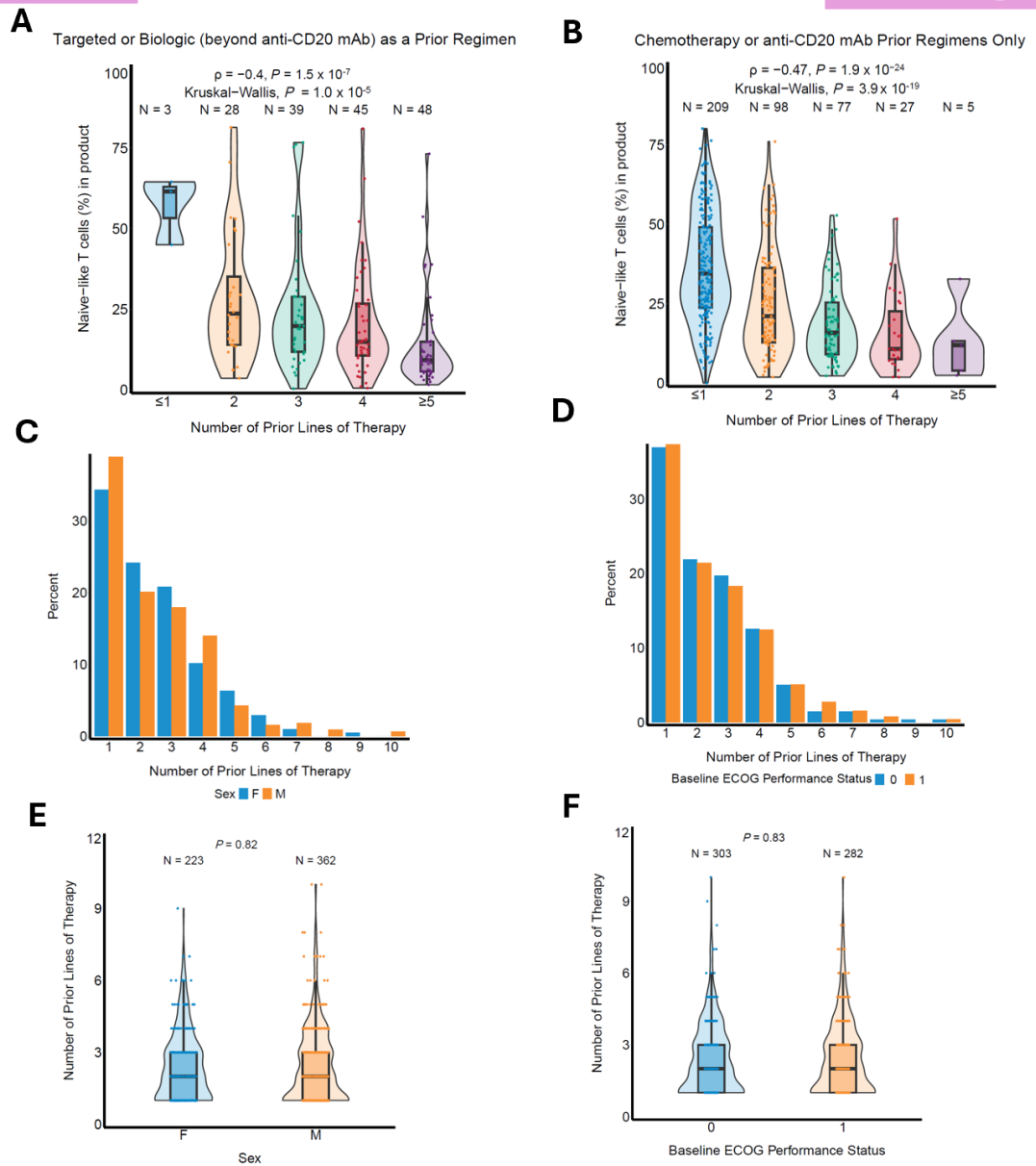
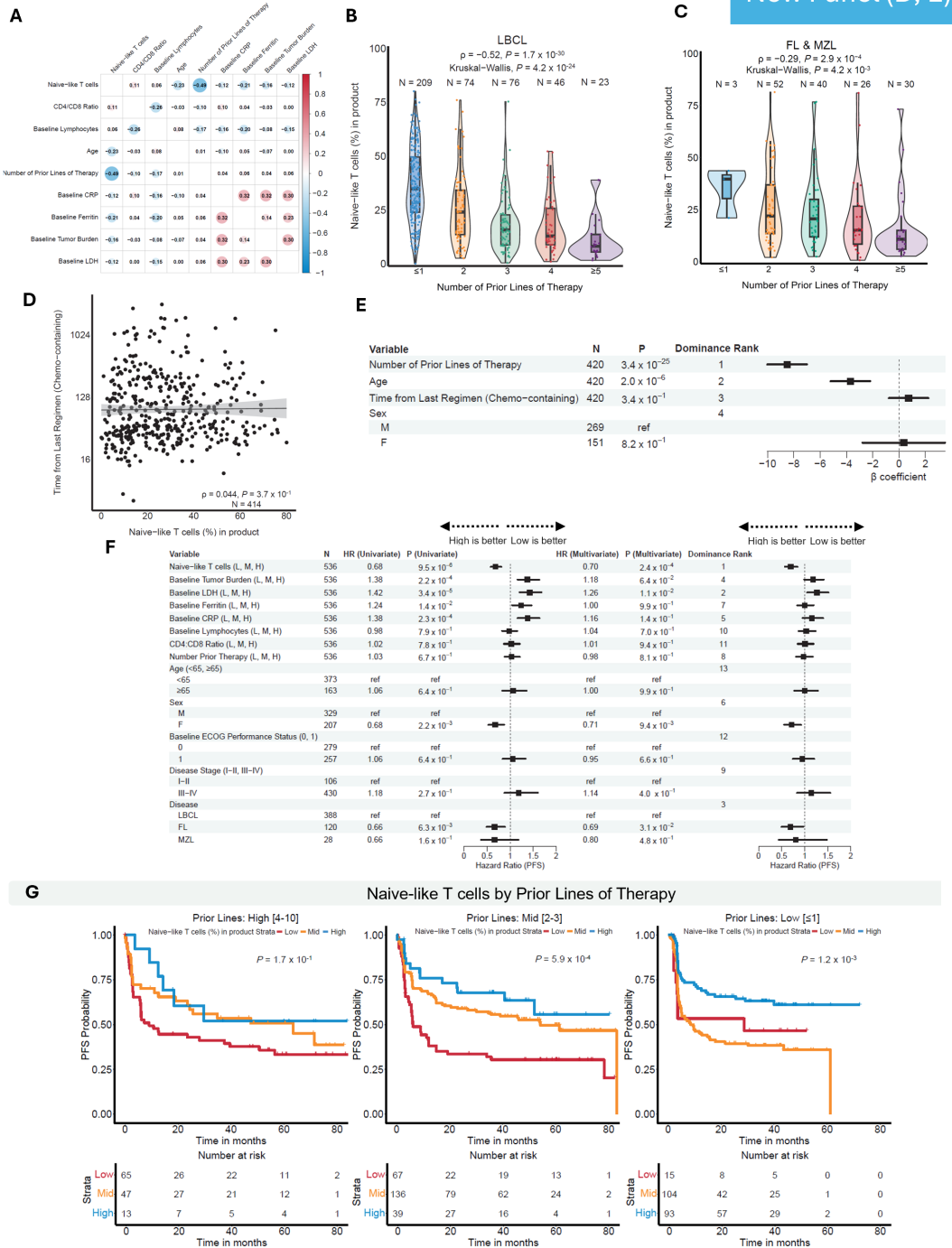


Fig. 2



2.6. - The observation that the correlation between the number of prior regimens and the proportion of juvenile T cells is stronger in the final product than in the apheresis is noteworthy and highlights the limitations of phenotyping alone in assessing T-cell state and function.

We agree and thank the reviewers for this comment. Indeed, the paired analysis of apheresis and infusion products from 235 patients represents a key and unique strength of our study, allowing us to uncover this important finding. In response to the reviewer's comments (previous

points), the revised manuscript now includes an expanded CyTOF-based characterization of the infusion product, providing additional insights into the functional orientation of T-cell subsets, particularly naive-like T cells. And we agree with the reviewer that functional investigations (which are undergoing in a much smaller proportion of samples and that are outside the scope of this *Analysis* manuscript) can yield additional important observations and provide a novel insights on the biological mechanisms underlying these associations.

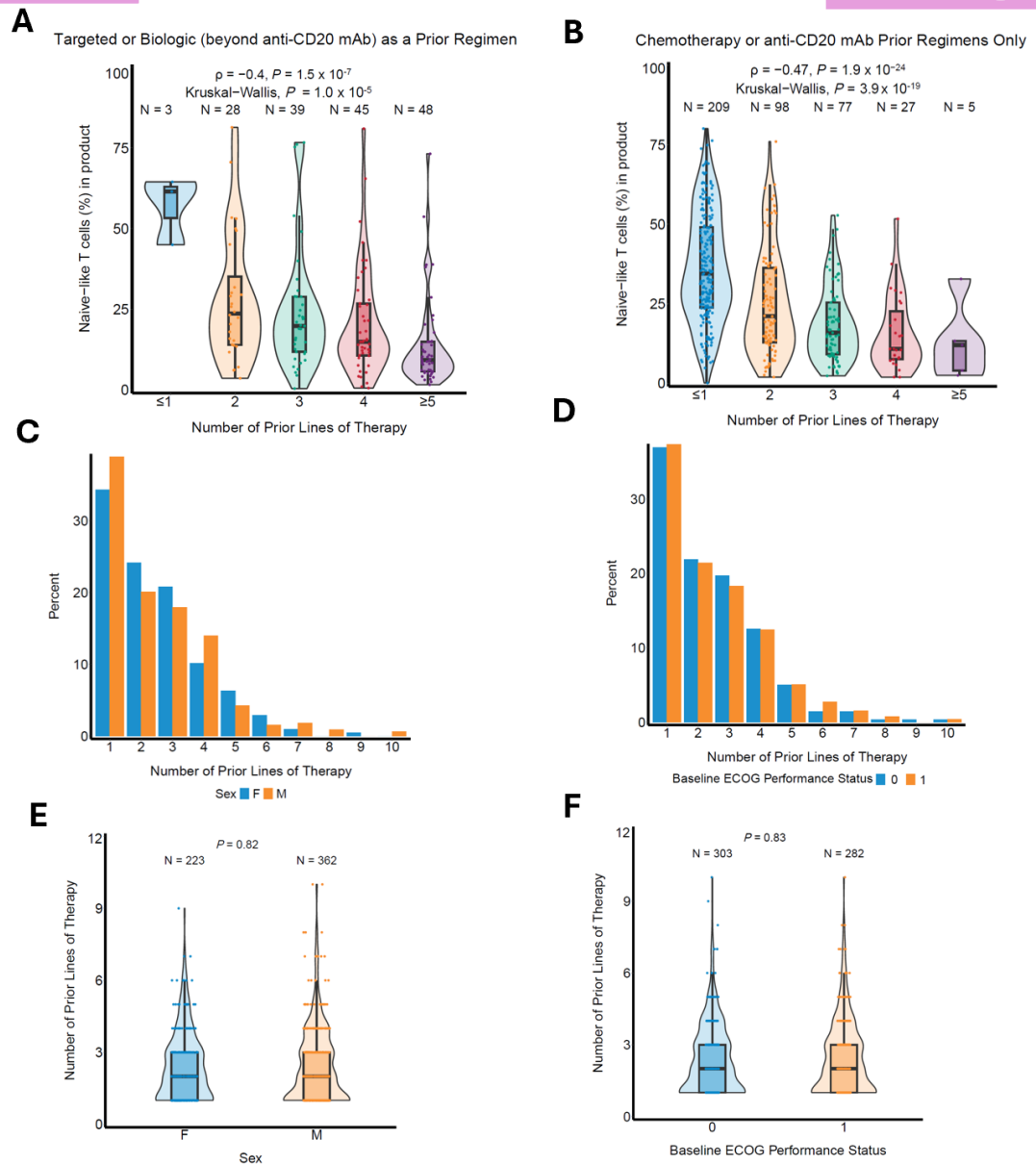
Minor comments

2.7. For Fig 1F, Extended Data Fig 2B, and Extended Data Fig 3B-C, please also provide ANOVA and/or Kruskal-Wallis test results.

We included Kruskal Wallis P values in all these panels and all the new similar plots that were included in the revised manuscript (violin plots comparing 3 groups).

2.8 Were there any differences in the median number of prior lines of therapy between males and females, or between patients with ECOG 0 versus 1?

As requested, we evaluated the relationship between prior lines of therapy and sex (male or female; **Supplementary Fig. 1C**) or baseline ECOG performance status (0 or 1; **Supplementary Fig. 1D**) and did not observe clear trends for differences in the number of prior lines of therapy within categorizations of either feature. Moreover, we performed Wilcoxon rank sum test to assess the statistical association between prior lines of therapy and sex (**Supplementary Fig. 1E**) or ECOG (**Supplementary Fig. 1F**) and saw no significant differences within either feature.



- 2.9 Please clarify in the Results section that the AUC refers to the period from day 0 to day 28 following CAR T-cell infusion.

This has been clarified in the revised manuscript: AUC refers to the period from day 0 to day 28 following CAR T-cell infusion (Page 9, line 235-236).

Reviewer #3 (Remarks to the Author)

This paper presents a multivariate analysis of 580 patients across 5 multi-center clinical trials who have received axi-cel, the CD19.28z CART product of KITE-Gilead, for a range of NHL subtypes. Multiple studies to identify correlates of response and response duration have been published, highlighting the importance of T-cell subsets in the apheresis product and the infusion product (going back to 2014) as well as disease burden, baseline inflammation (of the patient), prior lines of therapy and CAR T expansion and persistence. This analysis studies all of these parameters in a much larger group of patients with a median follow-up of 41.9 months.

They find that the frequency of what they call “juvenile” cells, phenotypically CCR7+CD45RA+, a population that contains both naive and T memory stem cells – cells with very different characteristics- to be the factor most positively associated with tumor response and duration in both univariate and multivariate analyses. Other factors such as baseline tumor burden, baseline inflammatory markers are also significant, but less so. The correlation held true across all five trials. These findings are not new or surprising, but are very clear and well presented in this manuscript. There was a correlation between the number of juvenile T-cells in the apheresis product and the infusion product, but their presence in the product was more significant.

We thank the reviewer for the thoughtful assessment of our study and for recognizing the value of the large integrated dataset and extended clinical follow-up. We appreciate the acknowledgment that our findings are clearly presented and consistent across trials. The reviewer’s comments helped us further strengthen the manuscript by expanding the discussion of potential mechanisms, manufacturing strategies, and prompted as to include additional phenotypic analyses of the product.

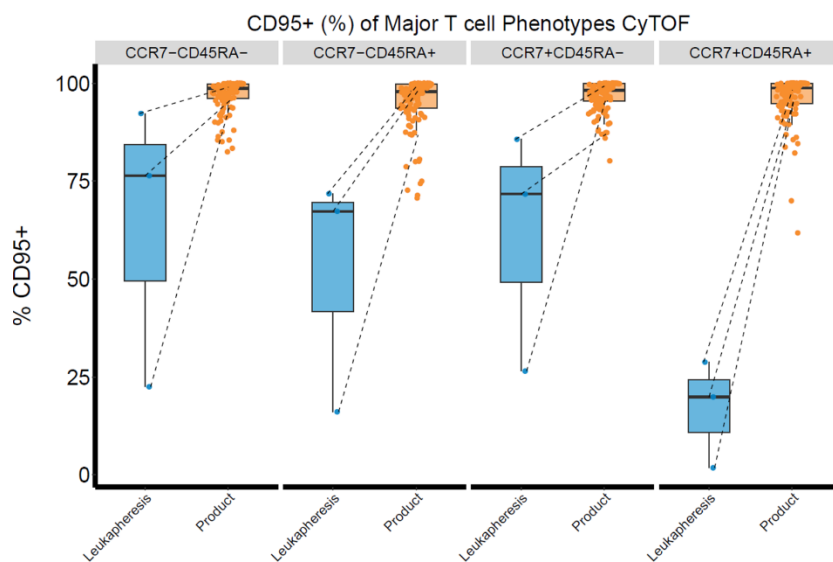
Specific comments

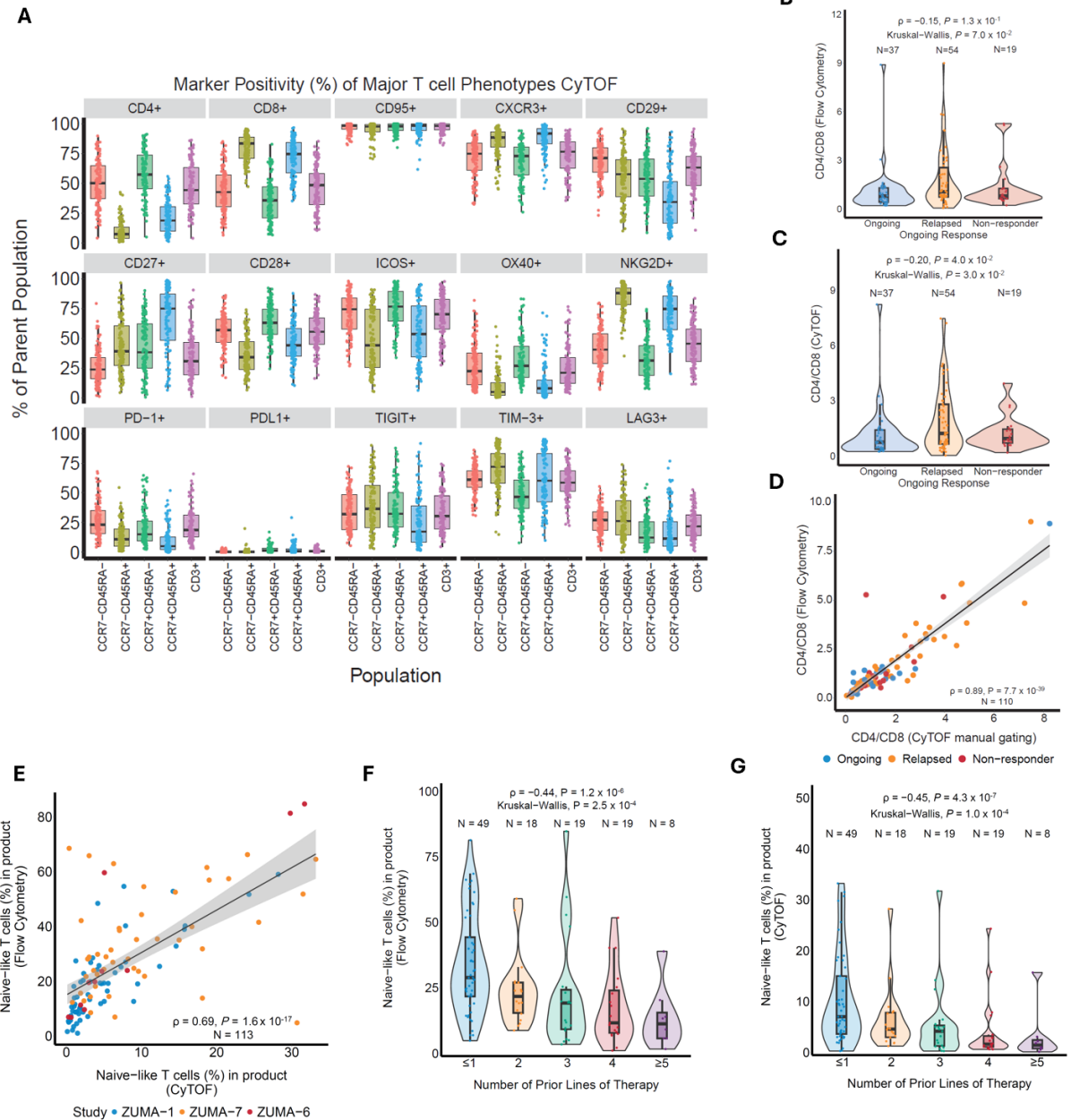
3.1 - You remark that the positive effect of a high frequency of juvenile T-cells is actionable, and refer to two clinical trials, but you do not describe how more juvenile T-cells could be obtained or preserved. You should discuss the many other studies that have attempted to increase the juvenile phenotype and other phenotypes.

We thank the reviewer for the important remarks. We have included the following text in the manuscript, “Enrichment of naive/memory-stem T cells in the product by CD62L selection (Sabatino et. al., Blood, 2016; Larson et. al., Cancer Discovery, 2023), CD45RO depletion (Jain, et. al., Cytotherapy, 2024), cytokine modulation (Zeng et. al., Journal of Experimental Medicine, 2005; Xu et. al., Blood, 2014; Gomez-Eerland et. al., Hum Gene Ther Methods, 2014; Sabatino et. al., Blood, 2016; Gattinoni, et. al., Cell, 2026), or rapid manufacturing (Murakami, et. al., Blood, 2024; Chen, et. al., Trends in Biotechnology, 2025) are strategies that could be implemented to improve product fitness and response rates. Aside from manufacturing modifications, our data suggests that CAR T cell treatment in earlier lines of therapy should result in higher proportions of naive/memory-stem cell T cells in the apheresis and product, resulting in improved outcomes.”

3.2 - It is important to know if naive or TSCM’s should be preserved. This distinction could have added innovation to this otherwise impressive, but descriptive work.

To differentiate naive T cells from TSCM, we assessed CD95 (Fas), a marker associated with stem-cell memory and T-cell activation, using manually gated CyTOF data from ZUMA-1, ZUMA-6, and ZUMA-7 product samples. CyTOF-based CD95 gating was performed on apheresis material from three (3) ZUMA-6 subjects as well as the full cohort of product samples analyzed (N = 116), allowing for evaluation of CD95 positive T cells pre- and post-manufacturing. Notably, we observed that CD95 levels were low in CCR7+CD45RA+ cells in the 3 tested apheresis samples, while regardless of CCR7 and CD45RA status in the product, we saw a high percentage of CD95 (**Extended Data Fig. 9A**). During axi-cel manufacturing, exogenous activation, cytokine stimulation, and expansion steps may drive CD95 upregulation. These results suggest that CD95 expression in the product may be reflective of manufacturing-induced activation rather than intrinsic memory phenotype, making it unsuitable for distinguishing naive from stem-cell memory T cells. The below panel on CD95 in 3 apheresis samples is provided for reviewers' consideration but has not been included in the revised manuscript in view of the very limited sample size of the apheresis CyTOF data.





3.3 - Extended Data Figure 1 could have been better explained for non-statisticians

The text has been updated to the following, “While the proportion of naive/stem-cell memory T cells (CCR7+CD45RA+) inversely correlated with the number and proportion of effector memory (EM) T cells (CCR7-CD45RA-) (Spearman’s $\rho = -0.79$, and -0.86 , respectively, **Extended Data Fig. 1**), the strength of the positive relationship between the proportion of naive/stem-cell memory T cells and PFS was greater than that of the negative relationship between the proportion or number of EM T cells and PFS (**Fig. 1B**).”

The figure legend has been updated to the following, “Spearman correlation heatmap for the pairwise comparison of each axi-cel product phenotype against each other in the 580 patients included in the meta-cohort. The spearman correlation coefficient (ρ) is provided for each pairwise comparison with the color capturing the directionality of correlation (positive correlation coefficient in red and negative correlation in blue), and the circle size capturing the absolute value of the correlation coefficient. All attributes were measured in all 580 samples. *N* indicates the number of infused cells for each population, and % indicates the proportion relative to the parent population (total cells for T cells, and T cells for the other populations). CM, central memory (CCR7+CD45RA-); EM, effector memory (CCR7-CD45RA-); EMRA, terminally differentiated effector memory (CCR7-CD45RA+); N/SCM, naive/memory stem cell (naive-like T cells (CCR7+CD45RA+)).”

E. D. Fig. 1

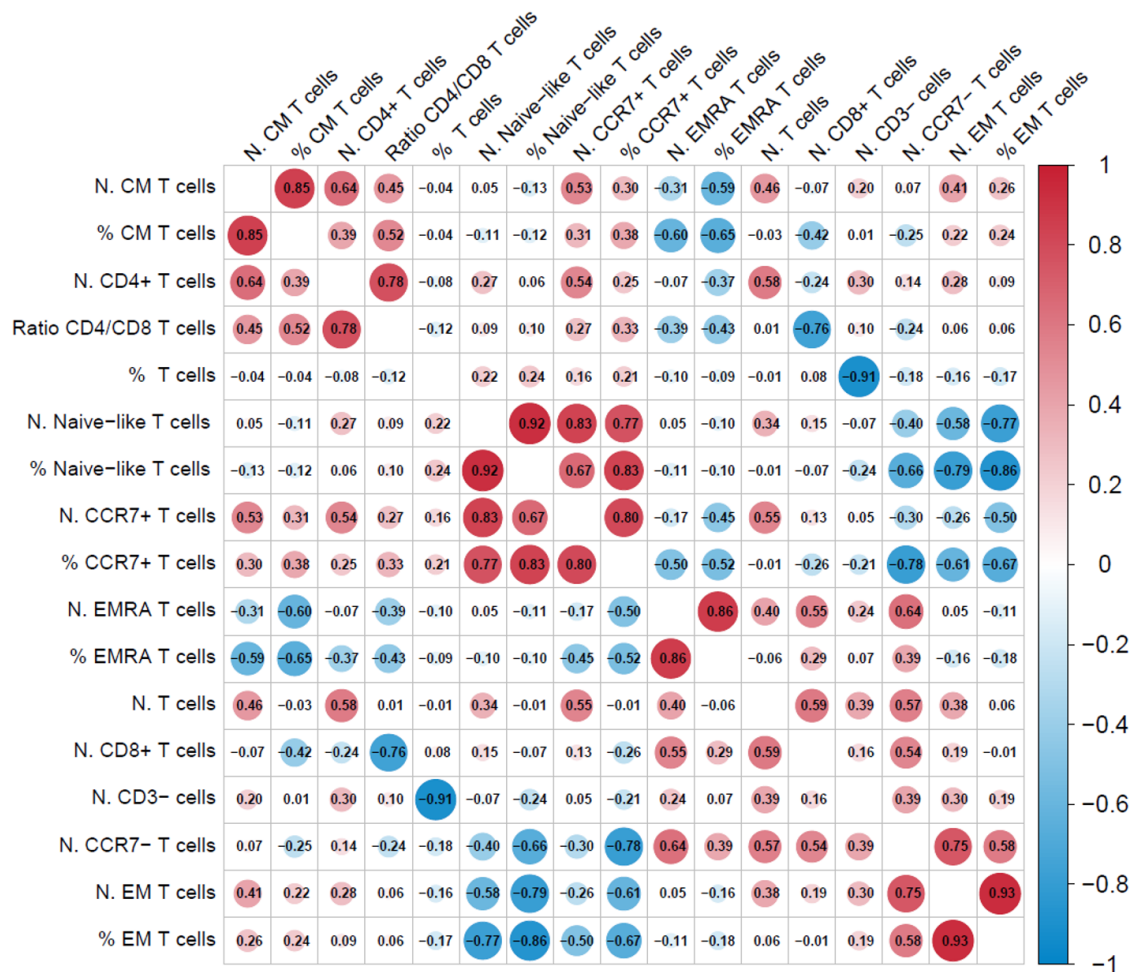
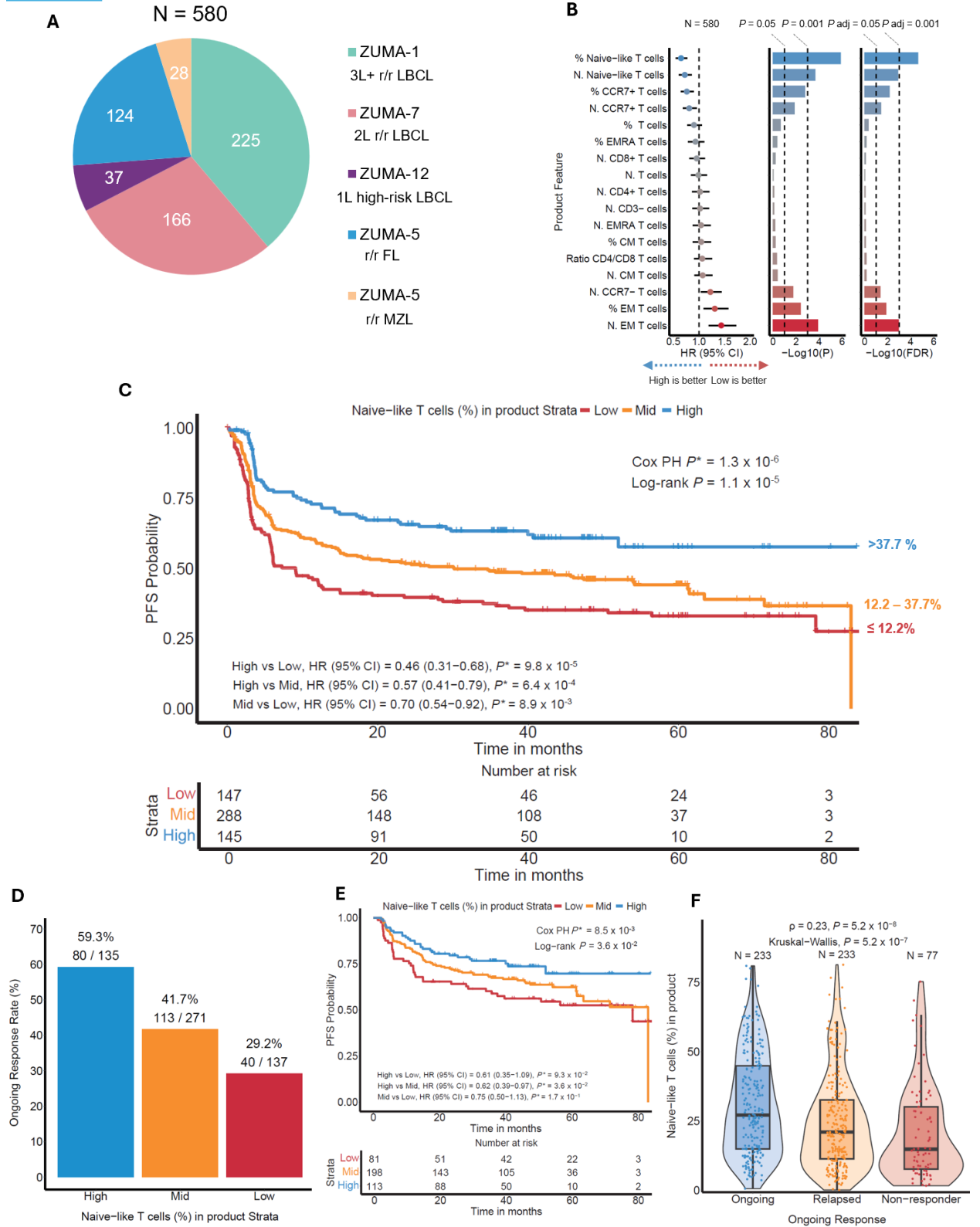


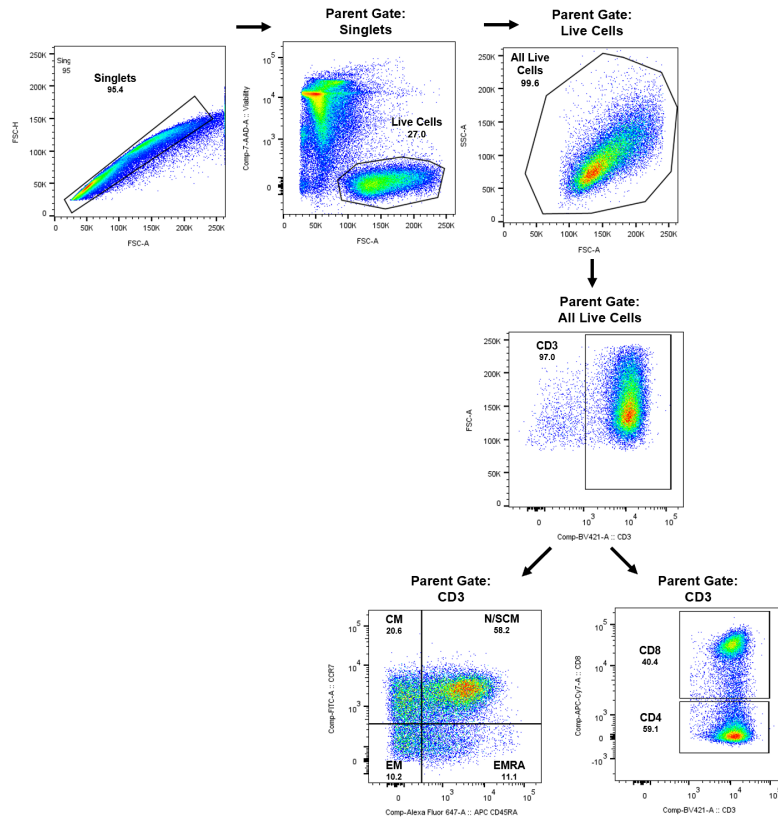
Fig. 1



3.4 - Was Supplementary Figure 1 describing the apheresis product or the infusion product

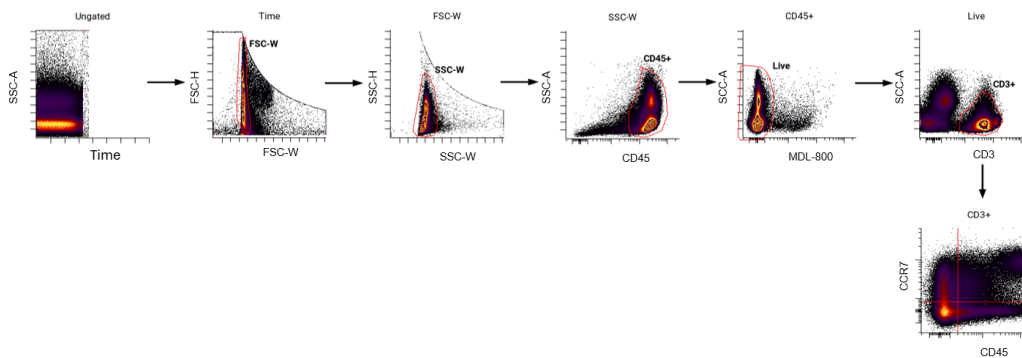
Supplementary Fig. 7 describes the infused product gating strategy, with analysis performed at Kite Pharma. Apheresis T cell phenotypes were calculated as a percentage of viable CD45+CD3+ singlet cells and identified by flow cytometry using a validated assay at CellCarta. An example of the apheresis gating strategy has been included as **Supplementary Fig. 8**, and the gating strategy has been described in the figure legend and Methods section.

Supplementary Fig. 7



Supplementary Fig. 8

New Figure



3.5 - There is only a weak correlation between CART expansion and AUC with juvenile T-cells. Could this be influenced by the copy# per cell. Could you supply the (median and range), since the PCR does not necessarily reflect the number of circulating cells.

This manuscript quantifies CAR T cell peak and AUC as cells/ μ L of blood using qPCR. Copy # per cell is not measurable because the qPCR assay did not include a quantitative standard curve which would allow for derivation of copies/ μ g of DNA.

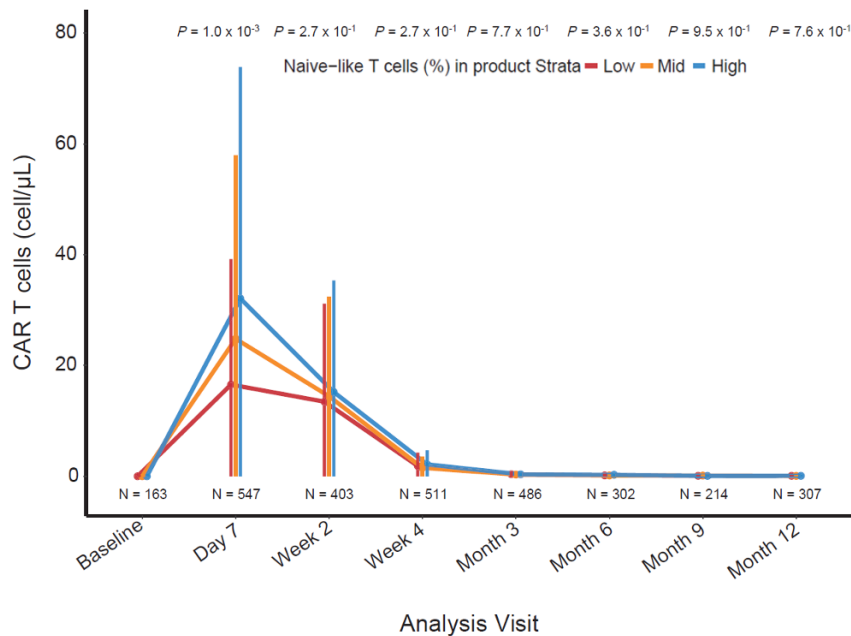
As previously described in Kochenderfer et al. (2012), DNA was extracted from PBMCs collected before treatment and at multiple time-points after treatment for each patient. Serial 1:5 dilutions were performed using DNA from the infused T cells (product) of each patient into pretreatment DNA prepared from leukocytes of the same patient, and standard curves were generated by performing qPCR on this DNA. The percent transduction was determined by flow cytometry as the percentage of infused T cells that expressed the anti-CD19 CAR, thereby assuming only infused T cells with surface CAR expression detected contained the CAR gene. This assumption likely underestimates the actual number of cells containing the CAR gene because all cells containing the CAR gene might not express the CAR protein on the cell surface. The percentage of PBMCs that contained the CAR gene at each time-point was determined by comparing the qPCR results obtained with DNA of PBMCs from each time-point to the qPCR results obtained from each patient's infused T cell standard curve. To estimate the number of CAR T cells/ μ L, the proportion of CAR+ cells in PBMCs was multiplied with the number of mononuclear cells in blood.

The derivation for CAR T cells/ μ L blood incorporates percent (%) transduction and normalization to mononuclear cells, as outlined below:

- $\% \text{CAR+ PBMCs} = \text{CAR19/Actin (\%)} * \text{transduction measured by flow cytometry (\%)}$
 - $\text{CAR19/Actin (\%)} = (\text{CAR concentration/housekeeping gene concentration}) * 100$
 - $\% \text{CAR+ PBMCs}$ normalized to mononuclear cells to yield CAR T cells/ μ L blood

We have incorporated the formula for CAR T cell expansion using qPCR in the Methods section of the manuscript.

Median and median absolute deviation of CAR T cell concentrations at each time point are shown in **Fig. 4C** (reproduced below).



Reviewer #3 (Remarks on code availability):

3.6 - What code?

This comment was not marked as remark to the author and we interpret that was intended for internal clarification during the review process. In the code availability we stated: “All statistical analyses were performed through the R software v.4.2.1 and v.4.2.2 (2022-10-31 ucrt), using predefined packages outlined in the methods” as there was no development of new codes.

Additional notes from the authors

We revised all figures to include high-resolution panels. Figures 1C and 1E were further updated to incorporate study as a covariate. We also added new authors who made substantial contributions to the CyTOF data analysis, validation cohort analysis, and extraction of additional features necessary to address the reviewers’ comments. All newly added authors reviewed and approved the final version of the manuscript.

Reviewer #3 (Remarks to the Author):

The authors have responded extraordinarily to the reviewers' comments and have added much new data, incorporating patients from additional clinical trials and Cytof data. The clinical dataset is exceptional, with 580 patients in the discovery cohort now validated by 203 patients in an independent validation cohort over multiple ZUMA studies with a long median follow-up. The statistical analyses appear rigorous, and the association between naïve-like product T cells and durable clinical benefit is convincing. However, the manuscript remains primarily correlative and confirmative of earlier, smaller studies showing associations between naïve-like cells and PFS and between the frequency of circulating naïve cells and lines of prior chemotherapy. The mechanistic interpretation lacks supporting biological data. The authors favor the hypothesis that PFS is driven by T_{SCM} cells, but since this population cannot be distinguished from naïve cells once activated, this remains speculative. The association between naïve-like T cells and PFS is strong, but remains an association so phrases implying a causal relationship should be modified. The manuscript would be strengthened by a deeper discussion of alternative biological explanations.

We thank the reviewer for their thoughtful comments and for recognizing the substantial additions to the revised manuscript, including the expanded clinical validation cohorts and CyTOF analyses. We additionally appreciate the reviewer's assessment that the association between naïve-like product T cells and durable clinical benefit is robust across the discovery and validation datasets.

We agree that the current study is primarily observational in nature and that the reported text regarding the results should be cautious about declaring causality. Accordingly, we have revised the manuscript throughout to use more conservative language when describing these findings. Specifically, we have replaced terms that could imply causation (e.g., "influence," "impact," "drive") with language that more appropriately reflects the associative nature of the results (e.g., "relationship," "association," "correlation," and "linked to"). We have also revised statements discussing naïve vs T_{SCM} cells to help clarify that the potential contributing population(s) which produce the naïve-like T cell product phenotype remains undetermined.

The following specific updates have been included in the Results and Discussion sections to address the inability to distinguish naïve from T_{SCM} in the CAR T product:

"Because both naïve T cells²⁴ and stem cell memory T (T_{SCM}) cells²⁵ express CCR7 and CD45RA²⁶, and CAR T cell products such as axi-cel are generated through polyclonal stimulation during manufacturing, this subset has been variably described as either T_{SCM}-like^{27,28} or naïve-like T cells¹⁹ in CAR T cell products. Here, we conservatively refer to this population (CCR7+ CD45RA+) as naïve-like T cells."

"Cluster 1 also displayed high expression of CXCR3, which is expressed by T_{SCM} cells but has also been reported in a subset of naïve T cells with enhanced effector differentiation potential (Fig. 6D, 6F)³⁵. Although the T_{SCM} marker CD95^{26,27,34} was highly expressed in cluster 1, it was also broadly expressed across other clusters, likely driven by the exogenous activation of the product during manufacturing."

“Although CD95 was broadly expressed across product cells, the absence of a comprehensive T_{SCM} marker panel (including CD45RO and other markers^{34,40}) limits the ability to definitively identify a genuine T_{SCM} population, and it remains possible that T_{SCM} cells were enriched within this cluster.”

“Future studies should better define and characterize the subpopulation responsible for the improved therapeutic effects.”

We have additionally provided updates to the Discussion regarding alternative biological explanations:

“These findings suggest that the favorable prognostic value of CCR7+CD45RA+ cells may arise from their ability to capture a population of less differentiated T cells characterized by greater functional fitness and limited exhaustion features. Alternatively, they may serve as a surrogate marker for the rarer T_{SCM} population which has demonstrated potent antitumor activity but is typically present at low frequencies in CAR T cell products under standard manufacturing conditions^{27,40}. Although CD95 was broadly expressed across product cells, the absence of a comprehensive T_{SCM} marker panel (including CD45RO and other markers^{34,40}) limits the ability to definitively identify a genuine T_{SCM} population, and it remains possible that T_{SCM} cells were enriched within this cluster.”

Specific comments

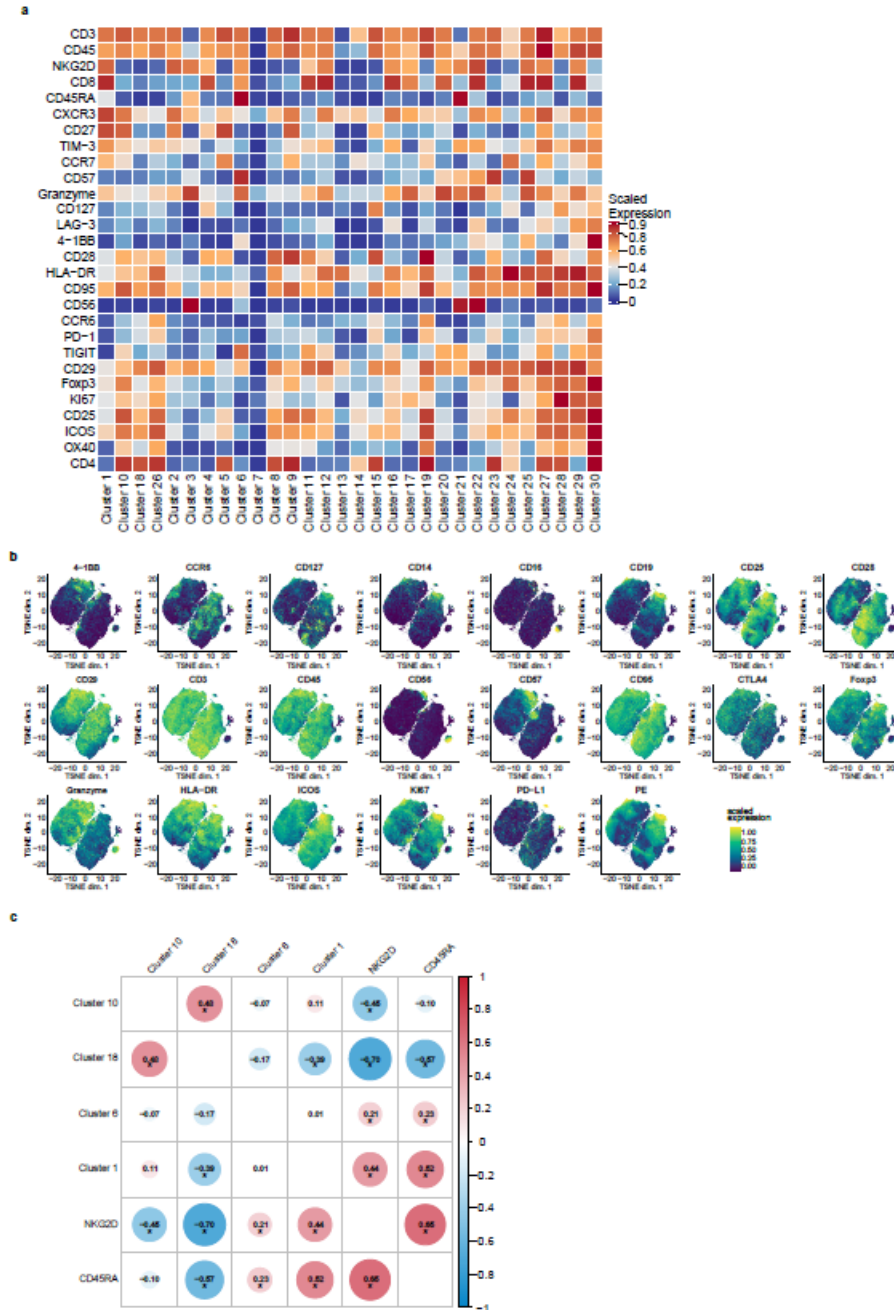
1) The Cytof data could be further exploited to identify pathways that distinguish Cluster 1. Does CXCR3 remain after adjusting for CCR7, are these cells metabolically distinct? Are the naive-like cells polarized towards TH1/2/17 etc.

We thank the reviewer for this thoughtful suggestion and agree that the CyTOF dataset contains a rich set of phenotypic information that can be used to further understand Cluster 1 and other T cell populations.

As presented in the manuscript, the phenotypic characteristics distinguishing the identified clusters are summarized in **Figure 6D**, **Figure 6F**, and **Supplementary Figure 6** (provided below), which provide marker-level comparisons across clusters. In addition, the Results section includes a detailed description of the marker expression patterns that distinguish these populations, including markers which appear as more distinctly expressed or absent within the CD8+ naive-like cluster 1. Given that the majority of the cells in this cluster are CD8+, the polarization towards helper phenotypes such as TH1/2/17 was not evaluated.

We also note that the elastic net analysis presented in **Figure 6E** was specifically designed to evaluate the relative contribution of clusters and phenotypic markers in a multivariable setting. By incorporating all measured clusters and markers into a single model, this analysis identifies features that retain independent associations with progression-free survival when accounting for the effects of other populations or markers. As such, the model provides context regarding the importance of individual markers and cell populations/clusters which extend beyond those identified in the univariable analyses. Additionally, we have provided a summary of correlation results related to the features maintained in the elastic net model to assist with delineation of relationships across features in **Supplementary Figure 6C**.

We agree that focused analyses of specific signaling pathways, metabolic programs, lineage polarization states, or the independence of individual markers such as CXCR3 relative to CCR7 could provide further biological insight. However, these questions extend beyond the primary objectives of the current study, which were to identify and validate cellular features in CAR T cell products which are associated with clinical outcomes across large independent cohorts. We have also added language to the Discussion to ensure the descriptive and associative nature of the CyTOF findings and the need for future studies to further define the biological mechanisms underlying these observations are highlighted. We have also added language to the Discussion to ensure the descriptive and associative nature of the CyTOF findings, and the need for future studies to further define the biological mechanisms underlying these observations, are highlighted.



2) The CyTOF analysis of the apheresis samples should have allowed you to distinguish naïve T cells from T memory stem cells (T_{SCM}), for example using CD95, CD122, CXCR3, or other markers. If so, was the association with progression-free survival driven predominantly by naïve T cells, T_{SCM}, or both?

We thank the reviewer for this important question. We agree that markers such as CD95, CD122, CXCR3, and related phenotypic features can be informative for distinguishing naïve T cells from T memory stem cells (T_{SCM}) under appropriate experimental conditions. However, the CyTOF data in this study do not permit a robust assessment of whether the observed association with

progression-free survival is driven predominantly by naive T cells, T_{SCM} cells, or both from the apheresis material. Specifically, only three (3) apheresis samples were available for CyTOF analysis. While these 3 samples were useful for evaluating broad phenotypic changes associated with the manufacturing process (such as CD95 expression in apheresis and product material, which was shared with the reviewer previously), the sample size is insufficient to establish meaningful relationships with progression-free survival or to perform reliable correlations between leukapheresis and product populations.

3) Line 118: Your prediction that CD45RA+ CCR7+ naïve-like cells more closely resemble T_{SCMs} is unlikely to be correct. T_{SCMs} represent less than 2% of all T-cells in healthy individuals and would be expected to maintain a similar frequency during the 14 days expansion, unless they expanded preferentially.

We thank the reviewer for this point and agree that T_{SCM} cells represent a relatively small fraction of circulating T cells in healthy individuals and are likely present at similar frequency in CAR T products with standard manufacturing. We would like to clarify that our statement was not intended to imply that the CD45RA+CCR7+ product population represents, or is derived predominantly from, T_{SCM} cells; rather, our intent was to highlight that the manufacturing process alters the phenotype of the T cell populations which leads to products which have atypical marker expression that more closely resemble classical circulating T_{SCM} rather than naive T cells. From a nomenclature/terminology perspective, in their recent paper, Gattinoni and colleagues (who initially characterized this subset in humans) labeled the CCR7+CD45RA+ population within the product as T_{SCM}-like, “TSCM were defined as CCR7+CD45RA+CD45RO- CD95+; TCM as CCR7+CD45RA-; TEM as CCR7-CD45RA-; TTE as CCR7- CD45RA+; TSCM-like cells as CCR7+CD45RA” ..from Gattinoni, L., et al. *Distinct in vivo dynamics of donor-derived stem cell memory CAR T cells post-allogeneic HSCT relapse*. *Cell* (2026). [Ref 27 in the present manuscript].

As such, the following text is included in the Results and Discussion sections:

“Because both naive T cells²⁴ and stem cell memory T (T_{SCM}) cells²⁵ express CCR7 and CD45RA²⁶, and CAR T cell products such as axi-cel are generated through polyclonal stimulation during manufacturing, this subset has been variably described as either T_{SCM}-like^{27,28} or naive-like T cells¹⁹ in CAR T cell products. Here, we conservatively refer to this population (CCR7+ CD45RA+) as naive-like T cells.”

“Cluster 1 also displayed high expression of CXCR3, which is expressed by T_{SCM} cells but has also been reported in a subset of naive T cells with enhanced effector differentiation potential (Fig. 6D, 6F)³⁵. Although the T_{SCM} marker CD95^{26,27,34} was highly expressed in cluster 1, it was also broadly expressed across other clusters, likely driven by the exogenous activation of the product during manufacturing.”

“Although CD95 was broadly expressed across product cells, the absence of a comprehensive T_{SCM} marker panel (including CD45RO and other markers^{34,40}) limits the ability to definitively identify a genuine T_{SCM} population, and it remains possible that T_{SCM} cells were enriched within this cluster.”

4) Can you comment from the literature whether naïve-like cells are superior in functional assays, for example, in serial killing?

We thank the reviewer for this insightful comment. Functional assays such as serial killing were not performed in the current study, but the broader literature generally supports the concept that less differentiated T cell populations, including naïve-like/stem memory-associated T cells, possess biological properties that are favorable for adoptive cellular therapy.

Specifically, previous studies have demonstrated that naïve-like and T_{SCM}-like CAR T cells cell exhibit enhanced proliferative capacity, greater self-renewal potential, improved persistence in mouse rechallenge assays, and superior antitumor activity compared with more differentiated T cell populations (Sommermeyer et al., 2015; Sabatino et al., 2016; Arcangeli et al., 2022). These characteristics are thought to contribute to sustained tumor control and durable responses following CAR T cell therapy.

With respect to serial killing specifically, we are not aware of any studies which have analyzed naïve-like or T_{SCM}-like CAR T cells as compared with other product phenotypes. However, the extent to which serial killing specifically accounts for the clinical associations observed in our study remains uncertain. Our data does not address this question directly, but we do provide the self-renewal hypothesis in the Discussion as a possible explanation regarding the observed results. Moreover, in the Discussion we acknowledge that the favorable associations observed in patients may reflect biological properties that have been ascribed to less differentiated T cell populations in prior studies.

5) The writing has become a little repetitive due to the additional data

We thank the reviewer for this helpful observation. We agree that the addition of validation cohorts, CyTOF analyses, and expanded clinical results increased the complexity of the manuscript and led to some repetitive narratives. In response, we have carefully revised the text to improve clarity and conciseness, removing repetitive descriptions and consolidating overlapping concepts where possible. These revisions reduced the manuscript length from more than 6,000 words to fewer than 5,000 words while preserving the key findings and supporting context. We believe these changes have improved the overall readability and flow of the manuscript.

References

Gattinoni, L., et al. Distinct in vivo dynamics of donor-derived stem cell memory CAR T cells post-allogeneic HSCT relapse. *Cell* (2026).

Sommermeier, D., Hudecek, M., Kosasih, P.L. et al. Chimeric antigen receptor-modified T cells derived from defined CD8⁺ and CD4⁺ subsets confer superior antitumor reactivity in vivo. *Leukemia* 30, 492–500 (2016).

Sabatino, M., Hu, J., Sommariva, M. et al. Generation of clinical-grade CD19-specific CAR-modified CD8⁺ memory stem cells for the treatment of human B-cell malignancies. *Blood* 128, 519–528 (2016).

Arcangeli, S., Bove, C., Mezzanotte, C. et al. CAR T cell manufacturing from naive/stem memory T lymphocytes enhances antitumor responses while curtailing cytokine release syndrome. *J. Clin. Invest.* 132, e150807 (2022).