Analysis of sequential treatments for hematological diseases by advanced statistical methods

Koster, E.A.S.

Analysis of sequential treatments for hematological diseases by advanced statistical methods

Eva Anne Sophie Koster

ISBN: 978-94-6537-077-4

Images on cover and chapter pages are based on data from swissALTI3D (©swisstopo)
Printed by: Ridderprint | www.ridderprint.nl

The research presented in this thesis was conducted at the department of Hematology of the Leiden University Medical Center and was financially supported by the Landkroon fellowship from the Doelfonds Leukemie of the Bontius Stichting.

Analysis of sequential treatments for hematological diseases by advanced statistical methods

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. S. de Rijcke,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 3 februari 2026
klokke 10.00 uur

door

Eva Anne Sophie Koster
geboren te Leiden
in 1993

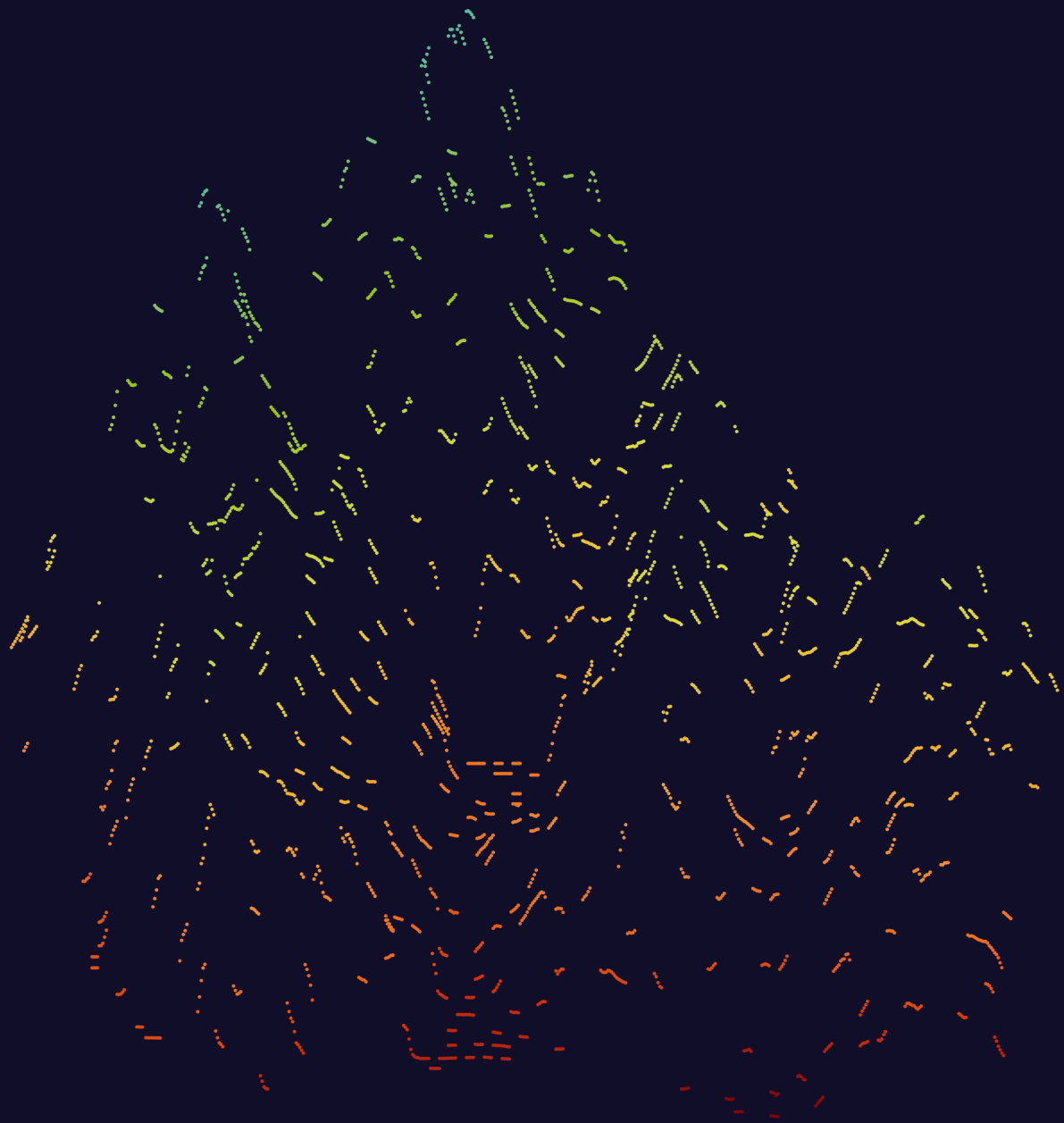
Promotor:	Prof. dr. J.H.F. Falkenburg
Co-promotoren:	Dr. C.J.M. Halkes Dr. L.C. de Wreede
Leden promotiecommissie:	Prof. dr. H. Putter Prof. dr. A.C. Lankester Prof. dr. M.D. Hazenberg (Amsterdam UMC) Dr. J. Versluis (Erasmus MC Cancer Institute)

agnovitque per umbras
obscuram, qualem primo qui surgere mense
aut videt aut vidisse putat per nubila lunam

Aeneis VI, 452-454

en hij herkende haar, verborgen
tussen de schaduwen, zoals men aan het begin van de maand
achter de wolken de maan ziet of denkt te zien opkomen

TABLE OF CONTENTS



1

General introduction

With observational studies researchers can gain valuable knowledge about diseases, treatments and associations between them from real-world data, but they have no control on the interventions or when which measurements are done. This calls for careful assessment of potential biases and optimal use of the available data. Investigating treatment outcomes becomes more challenging when the treatment consists of sequential interventions, more complex outcomes than for instance overall survival are considered, or when risk factors are analyzed that have different effects on different outcomes or are affected by events during follow-up. In these cases, more advanced statistical methods are often required. In this thesis, two complex clinical settings are investigated: allogeneic hematopoietic stem cell transplantation (alloSCT) for patients with acute leukemia and immunosuppressive therapy (IST) for patients with acquired aplastic anemia (AA). Both cases will first be introduced, followed by a brief overview of commonly used methodological approaches, promising advanced approaches and the aims of this thesis.

Allogeneic hematopoietic stem cell transplantation for patients with acute leukemia

The aim of alloSCT for patients with acute leukemia is to eradicate the disease by replacing patient hematopoiesis by donor-derived hematopoiesis and by introducing donor-derived alloreactive T cells that can eliminate the malignant hematopoietic cells of the patient. The latter process is called the Graft-versus-Leukemia (GvL) effect and may result in lifelong immunity against the malignancy.[1] However, if these alloreactive T cells target non-hematopoietic tissues of the patient, Graft-versus-Host-Disease (GvHD) may develop.[2] The success of alloSCT as treatment against acute leukemia depends on establishing sufficient GvL without inducing severe GvHD.

Both the GvL effect and GvHD result from an alloimmune response of donor-derived T cells recognizing a 'nonself' peptide/HLA complex on patient cells: the peptide, HLA molecule or both are not present in the donor due to genetic differences.[2] Vice versa, patient-derived T cells recognizing nonself peptide/HLA complexes on donor-derived hematopoietic cells can cause graft rejection. Since all nucleated cells present HLA and any peptide presented in nonself HLA may provoke an alloimmune response, patients with an HLA-mismatched donor have high risks of developing severe GvHD and graft rejection.[3] Therefore, HLA-matched donors are generally preferred. In the setting of fully HLA-matched alloSCT, alloimmune T-cell responses are directed against immunogenic nonself peptides presented in self HLA. These peptides are called minor histocompatibility antigens (MiHAs). Due to the high genetic diversity in peptides and T-cell receptors, the patient and donor may have hundreds of different MiHA-specific T cells, which are, in theory, capable of graft rejection, GvL and/or GvHD, depending on the origin of these T cells and the tissue expression of the MiHAs.[2] Patients with an HLA-matched unrelated donor (UD) have about twice as many disparities as patients with an HLA-matched sibling donor (called related donor, RD).[4] They have generally higher risks of GvHD and graft rejection, and a stronger GvL effect leading to a lower relapse risk.[5]

The alloSCT strategy consists of several steps: the conditioning of the patient, infusion of the graft, interventions to prevent severe GvHD and, in some strategies, interventions to improve the GvL effect. All steps influence the lymphohematopoietic status and

recovery of the patient and the risks of graft rejection, infections, GvHD, relapse and mortality.

Conditioning, graft infusion and lymphohematopoietic recovery

In the case of alloSCT for acute leukemia, the conditioning regimen has three aims: to make space for the hematopoietic stem cells (HSCs) of the donor, to reduce the tumor burden, and to prevent graft rejection. The first two aims are achieved by eliminating most HSCs of the patient, both healthy and malignant, and the third aim by suppressing the patient's immune system, including the MiHA-specific T cells. This is done by administering chemotherapy with or without irradiation and/or antibody therapy against immune cells during the days before graft infusion. The myelosuppressive potency of the conditioning regimen determines how many HSCs of the patient survive the conditioning regimen to compete with the donor HSCs and how many leukemia cells survive this stage.[6]

After the conditioning, the stem cell graft from the donor is infused. Unmanipulated grafts consist of HSCs and immune cells. The donor HSCs home to the patient's bone marrow (BM), where they compete with the surviving patient-derived HSCs to repopulate the BM.[7] This competitive repopulation often leads to a state of mixed chimerism (MC): both patient- and donor-derived HSCs are present. Both populations produce immune cells. The recovery of innate immune cells closely follows the BM repopulation, as they have a relatively high turnover. In contrast, *de novo* generation of T cells requires a functioning thymus, and it takes months before *de novo* T cells appear after alloSCT. The early T-cell recovery depends on homeostatic proliferation (i.e., expansion of mature memory/effector T cells in a lymphopenic environment) of both the patient T cells that survived the conditioning regimen and the donor T cells that were present in the graft, and the expansion of T cells in response to antigens they encounter.[8]

T-cell alloreactivity and the need for GvHD prophylaxis

From the moment the graft is infused, alloreactive T cells encounter nonself antigens: any surviving patient-derived alloreactive T cells may encounter nonself antigens on hematopoietic cells of the donor while infused donor alloreactive T cells may encounter nonself antigens on hematopoietic and non-hematopoietic cells of the patient. In HLA-matched alloSCT, alloreactive T cells are usually naïve T cells: they have never encountered the antigen before and need costimulatory signals to become appropriately activated.[2] These can be given by professional antigen presenting cells (APCs) such as dendritic cells. Around the time of alloSCT, these APCs become activated in several ways: the conditioning regimen causes tissue damage leading to release of danger signals, the destruction of epithelial cells allows translocation of microbial products over the intestinal barrier, and infections and viral reactivations occur often due to the low immunity of the patient.[9,10] Without intervention, the resulting activation of APCs would lead to massive activation of alloreactive T cells. Therefore, to prevent lethal GvHD (and graft rejection) patients usually receive systemic immunosuppression for several months after alloSCT. During this period, the tissue damage and the epithelial barrier are repaired and initial immunological recovery takes place, providing some protection against infections. Moreover, the patient-derived professional APCs are gradually replaced by donor-derived APCs, which are less likely to induce a strong alloimmune response by donor-derived T cells.[11]

T-cell depletion and posttransplant cyclophosphamide to reduce the risk of GvHD

Even if GvHD prophylaxis is used, the allo-immunological pressure after HLA-matched alloSCT is considerable: about a third of the patients develop GvHD and GvHD is the main cause of non-relapse mortality (NRM).[5] To reduce the risk of severe GvHD, T-cell depletion (TCD) can be applied. With *ex vivo* TCD, the graft is manipulated by selecting certain cell subsets (e.g., CD34+ selection by immunomagnetic procedures) or by removing certain cell subsets (e.g., CD52+ immune cells by alemtuzumab, depletion of (subsets of) T cells by immunomagnetic procedures).[12,13] With *in vivo* TCD, patients receive alemtuzumab and/or anti-thymocyte globulin (ATG) intravenously. While TCD can effectively reduce the risk of GvHD[14], some studies have shown an increase in the risks of relapse and infections.[15-17] These studies demonstrate the downside of TCD: also the alloreactive T cells responsible for the GvL effect and the non-alloreactive T cells responsible for the protection against viruses are affected.

Another method to reduce the GvHD risk is posttransplant cyclophosphamide (PTCY): patients receive an unmanipulated graft, followed by cyclophosphamide and start of GvHD prophylaxis a few days later, when the alloreactive T cells have been activated but before they start eliminating their target cells. Cyclophosphamide affects mostly activated T cells, leading to preferential recovery of regulatory T cells and non-alloreactive T cells.[18,19] While this leads to a better protection against infections compared to TCD[20], the GvL effect is still suppressed and the risk of relapse remains higher compared to non-TCD alloSCT[21].

Donor lymphocyte infusions to boost the GvL effect

To boost the GvL effect after alloSCT, additional alloreactive donor T cells may be given to the patient. This can be done by the administration of unmodified donor lymphocyte infusions (DLI), which contain alloreactive and non-alloreactive T cells and other immune cells. The higher the T-cell dose the more effective and potentially toxic the DLI, i.e., the stronger the GvL effect and the higher the risk of severe GvHD. Therefore, the dose depends partly on the indication of the DLI.[22] Firstly, DLI can be given therapeutically to patients with a relapse after alloSCT, often in combination with chemotherapy to reduce the tumor burden. Establishment of a strong alloimmune response is vital for these patients. Therefore, the DLI dose is relatively high, and an increased risk of inducing GvHD is accepted. While some patients with overt relapse can be rescued with this treatment, the majority dies of relapse (insufficient GvL effect) or GvHD (too strong alloimmune response).[23-25] Secondly, DLI can be administered preemptively to patients with MC or minimal residual disease (MRD), which may be signs of an impending relapse. In this case, there is more time for awaiting the effect of DLI, and the starting dose is lower. Based on the persistence of MC and/or MRD, subsequent DLIs are given over time with increasing dose.[22] The goal of this gradual dose escalation is to achieve a sufficient GvL effect with the lowest dose possible, thereby minimizing the risk of inducing severe GvHD. Lastly, DLI can be administered prophylactically to all patients without GvHD, i.e., to all patients without a sign of alloreactivity. As relapses may occur without any foreboding signs, one may choose to administer prophylactic DLI to boost the GvL effect even in the absence of MC and/or MRD to minimize the risk of relapse. The dose is usually comparable to that of

preemptive DLI, as these patients do not have a relapse yet and the risk of inducing severe GvHD should be minimized.

The alloreactive potential of DLI also depends on the genetic disparity and the presence of pro-inflammatory conditions that gradually diminish after alloSCT. Therefore, the DLI dose is also determined by donor type and timing after alloSCT: patients with an UD often receive a lower dose than patients with a RD, and earlier DLI are given at a lower dose than later DLI.[22] However, despite adjusting the dose to donor type and timing, alloreactivity by DLI is highly variable: some patients succumb to severe GvHD, while others do not show any sign of GvHD and GvL and may relapse.

Combining interventions to optimize the balance between GvHD and GvL

Some alloSCT strategies combine TCD or PTCY with prophylactic DLI to improve the balance between GvHD and GvL. The idea is to perform the alloSCT in two steps: to first introduce donor-derived hematopoiesis with minimal risk of severe GvHD, and then introduce donor-derived immunity to establish a sufficient GvL effect. Because the second step occurs after the initial recovery has taken place, the alloreactive T cells arrive in a less pro-inflammatory environment, leading to a lower risk of GvHD compared to if they had been infused directly after the conditioning. The strategy relies on the antitumor effect of the conditioning itself to control the leukemia until the DLI can be given. After the prophylactic DLI, preemptive DLI can be given if the patient still has MC or MRD.

The complex dynamics of lymphohematopoietic recovery and clinical events after alloSCT and DLI

Disentangling the effects of the different factors, mechanisms and interventions on the recovery after alloSCT is challenging. Patient factors, donor factors, conditioning intensity, the use and type of GvHD prophylaxis, TCD and/or PTCY all influence the competitive repopulation, homeostatic proliferation and/or allo-immunological pressure after alloSCT, determining the sizes of the emerging patient- and donor-derived lymphohematopoietic cell populations. The patient- and donor-derived populations can coexist or one population can eliminate the other via an alloimmune response, leading to graft rejection or a GvL effect. The latter may be accompanied by GvHD. While immunity is low, patients have a high risk of infections, which lead to a proinflammatory environment stimulating alloimmune responses. Posttransplant interventions such as DLI further complicate the dynamics. The potency of each DLI depends on many factors, i.e., patient- and donor-related factors, the DLI product and the conditions at the time of DLI. The DLI itself may affect the lymphohematopoietic recovery, cause GvHD and temporarily increase the mortality risk. The GvL effect of the DLI is often hard to quantify: therapeutic DLI are usually combined with other treatments, conversion from MC to FDC may also have occurred without preemptive DLI and for patients with FDC and no MRD receiving prophylactic DLI, it's impossible to say what would have happened if no DLI had been administered. Capturing these dynamics and estimating the effects of risk factors is complex. Some of the commonly used methodological approaches and more advanced approaches will be explained in the methodological section of the introduction.

Immunosuppressive treatment for patients with acquired aplastic anemia

AA is a hematological disease characterized by a hypocellular BM and hematopoietic failure leading to pancytopenia. Without treatment, patients may succumb to anemia, bleeding or infections. In the majority of the cases, acquired AA seems to be caused by an autoimmune reaction against hematopoietic cells.[26,27] There are two main treatment options: replacing the patient hematopoiesis and immunity by alloSCT or suppressing the autoimmune reaction by IST. AlloSCT leads to rapid and enduring hematopoietic recovery at the risk of transplant-related morbidity and mortality, mostly because of GvHD.[28] Therefore, currently alloSCT is only recommended for patients of 40 years or younger who have a suitable HLA-matched RD. The majority of adult patients with AA are treated with an IST regimen based on ATG and ciclosporin.[29] This treatment has moderate side effects compared to alloSCT but is less effective: only two-third of the patients respond, often only partially: these patients become transfusion-independent but their blood counts remain low.[30] Improvement of hematopoiesis after IST can take six months or even longer, as the autoimmune response first needs to be sufficiently suppressed after which the few surviving HSCs need time to repopulate the BM. During this period, patients remain at risk for bleeding and life-threatening infections due to their pancytopenia. After achievement of a response, the IST is tapered with the aim to stop. However, 30% of the responders develop relapse of the disease, requiring to restart or increase the dose of the IST, or even proceed to alloSCT.[31] Additionally, patients with AA receiving IST often have clonal evolution of hematopoietic cells, which may eventually lead to other BM diseases, most importantly acute myeloid leukemia (AML) and paroxysmal nocturnal hemoglobinuria (PNH).[31]

As older patients may have a higher risk of treatment-related toxicity and mortality and a lower likelihood of achieving a response, there is still debate whether ATG-based IST should be the first-line treatment of choice for patients aged 60 or older, instead of a less intensive treatment.[29,32-34] The arguments in this discussion are usually based on short-term response descriptions, overall survival and cumulative incidences of different types of failures. However, these estimates do not give a good overview of the likelihood of treatment success over time for several reasons. IST patients often need to be treated for months before a response becomes visible, while those who respond remain at high risk for several failure types: recurrence of the disease, development of another BM disease and death due to the complications of cytopenia or due to treatment toxicity. Some failures can be reversed by change of treatment (e.g., recurrence of disease can be treated by increasing the IST dose, restarting the IST, starting other IST or alloSCT). Patients can also experience different types of failure over time. For instance, a patient may first show a response, then relapse, develop AML and die. Moreover, some failure types are less severe than others: recurrence of the disease is less severe than development of AML. A single 'treatment success' measure capturing these highly dynamic outcome possibilities would be valuable for the evaluation of the treatment toxicity and efficacy over time. However, the estimation of such an endpoint is a challenge, as will be explained in the methodological section.

Methodological challenges and common methodological approaches

Measuring treatment outcomes

To assess treatment success after alloSCT for acute leukemia GvHD-relapse-free survival (GRFS) is often used: the probability of surviving without experiencing any clinically relevant GvHD or relapse.[35] A limitation of composite endpoints like GRFS is that they do not give any information on the reason of failure, in this case relapse (not enough GvL), GvHD (too strong alloimmune response) or death without relapse and GvHD, often related to treatment toxicity. Not all failures are equally severe and risk factors often have different effects on different components of the composite endpoint (for instance, having an unrelated donor decreases the risk of relapse but increases the risk of GvHD). Another limitation of composite endpoints is that subsequent events are ignored. For instance, GRFS does not consider that GvHD can resolve, and that patients who developed GvHD also likely established a GvL effect protecting them from relapse. Their prognosis may even be better than that of patients who never developed GvHD. Therefore, considering GvHD as definitive treatment failure seems too strict. To overcome these limitations, GRFS is often reported together with relapse-free survival, overall survival and cumulative incidences of GvHD, relapse and non-relapse mortality separately. The reader has to combine the results of all these analyses to obtain a full picture of the clinical recovery.

Analyzing the outcome of IST for aplastic anemia faces similar problems. In this setting, assessment of treatment success is usually based on overall survival and the recovery of the blood cell counts: (partial) recovery indicates (partial) disease response. As mentioned before, the timing of the disease response is variable, with most responses occurring within 3 months but some also beyond 6 months[30], and the response can be lost. The probability of reaching a response over time can be shown by cumulative incidence curves, but these give no information on what happened after achievement of a response. This is a major limitation in a setting where even after a response, failure often occurs. Therefore, usually the current response at certain times (most often 6 months) is reported. Aside from only giving information at one time point, these descriptive analyses give no information on temporary responses before this time point. Peffault de Latour et al. provided information on loss of response between 3 and 6 months in the table legends[30], but most often the temporary responses are not described at all. Prabahran et al.[33] reported outcomes of subsequent events by showing overall survival curves and risks of relapse and clonal evolution from different stages: from start of IST, from the 6-month response, after relapse and after second-line alloSCT. While this approach allows to zoom in on certain phases of the treatment, it requires a multitude of analyses (for each phase and outcome measure) and figures with different timescales. For each analysis, the reader needs to consider who is at risk (e.g., only those who have a response at 6 months) and the time between the start of the study and the start of the analysis. Combining all information to obtain a full picture of the recovery is challenging and becomes even impossible if some analyses start at the time of an intermediate event instead of a fixed time since start of the main analysis.

Investigating the effects of events and biomarkers during follow-up

Risk factors can be categorized into baseline risk factors, known at or before the start of the treatment, and time-dependent risk factors which can change after treatment, such

as infection. Especially the effect of the latter can be complex to estimate correctly. The most important rule is to only use information known at the present or past to predict the future. Otherwise, immortal time bias will occur. Immortal time refers to a period in which the event of interest (e.g., death) cannot occur due to the design of the analysis. Bias occurs for example when responders are defined by having a response during follow-up and are compared with non-responders from the start of treatment. The patients in the first group cannot die until they have achieved a response (otherwise, they would not have been selected for this group), while those in the latter group can. Thus, considering the response during follow-up as known at baseline favors the first group in this case. Even though this problem was already recognized in 1983[36], this is still a commonly made mistake. For instance, Zhou et al. found chronic GvHD to be the strongest predictor in their prognostic model for longer survival in patients with chronic myelomonocytic leukemia receiving alloSCT, but did not take into account that patients needed to survive for at least a few months before they could develop chronic GvHD.[37] The non-GvHD group could die from day 1, leading to an overestimation of the effect of chronic GvHD on survival. In the setting of IST for AA, two studies aimed to show the impact of experiencing relapse, PNH and AML on survival, but did not consider this bias. In their figures, at the beginning of follow-up the 'no event' group has temporarily lower survival compared to the groups with an adverse event.[38,39]

There are several relatively commonly used approaches to prevent this bias. In intention-to-treat analyses, groups are defined at baseline based on the treatment they are *intended* to receive instead of who actually received it. However, this method can only be used in settings where treatment allocation is known at time of start, it cannot be used for clinical developments like GvHD, and it usually attenuates outcome differences between groups. This attenuation occurs because often the treatment group contains some patients who actually did not receive the treatment and vice versa: the groups become more similar and the differences in outcome often smaller than if all patients could have been allocated to the correct group. Landmark analyses at certain time points after start only include the patients who are still at risk for the event of interest and split the group based on events that occurred until the landmark time. Immortal time bias is prevented while at least some of the information during follow-up can be used to define the groups, but information of patients who already had the event of interest before the landmark time is lost. Moreover, there is often no clear optimal landmark time: earlier landmark times include more patients but the groups may still contain a considerable number of patients that have the group-defining event after the landmark time, while later landmark times throw away more information. Often, multiple landmark times are chosen, but this may require correction for multiple testing. A method that can include all patients and event data is the time-dependent Cox proportional hazards model. In this model, covariates can change their value over time, assuming that the values of the covariates are constant until the next observation. For clinical events this is often acceptable – for treatments the exact starting time is known and the time of events like relapse are defined as the day of observing the relapse – but this may be a problem when analyzing biomarkers such as MRD markers and lymphocyte counts. Their values can change significantly in between two measurements, which may relate to the development of events such as relapse and death. Other limitations of the time-dependent Cox model are that it does not consider measurement error and that no absolute risks can be calculated since the probability of the intermediate event is not modelled explicitly.

Promising advanced methodological approaches

Multi-state model to capture complex sequences of events

The main limitation of GRFS as an endpoint can be overcome by including recovery after GvHD and calculating current GvHD-relapse-free survival (cGRFS): the probability of being alive without relapse and currently not having GvHD.[40,41] This can be done by multi-state models, which capture sequences of events, allowing to keep track of the clinical trajectories of patients in detail. In a multi-state model patients move between states at the occurrence of clinical events or treatments. Transitions define which routes between states are allowed. Figure 1 shows the structure of a multi-state model incorporating GvHD, relapse and death. The most common multi-state model is the time-inhomogeneous Markov model, which assumes that the hazard of making a certain transition only depends on the current state and the time since the start of the analysis.

Another advantage of the multi-state model is that the effects of risk factors can be modelled on each of the transitions separately, usually by means of transition-specific Cox proportional hazards models. For example, in the model of Figure 1, donor type and conditioning intensity are likely relevant for all transitions, while disease risk only needs to be modelled for the transitions to Relapse. Each transition hazard zooms in on a specific part of the process and all this information needs to be combined to get a full picture of the recovery. The model does this by using all transition hazards to calculate the probability of being in a certain state or set of states. This can be done non-parametrically (without taking any risk factors into account) or semi-parametrically by considering transition-specific Cox models for one or more transitions. The latter allows to show the clinical impact of the risk factors on different outcome measures, such as cGRFS (probability of being in 'Alive without relapse/GvHD') and relapse-free survival (probability of being in 'Alive without relapse/GvHD' or GvHD) in Figure 1.

In conclusion, the multi-state framework can overcome all described limitations of the composite endpoint: it keeps track of which failures and recoveries occur, captures sequences of events, enables to investigate the effects of risk factors on different

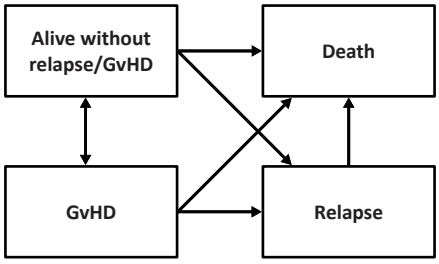


Figure 1. Example of a multi-state model. The boxes represent states, the arrows transitions. All patients start in the starting state 'Alive without relapse/GvHD'. From here, they move through the model at the occurrence of events. The other states can be absorbing (impossible to leave, e.g. Death) or intermediate (possible to leave, e.g. GvHD and Relapse). From Relapse, patients can only move to Death: achievement of remission after relapse is not considered. In contrast, patients in the state GvHD can return to the state 'Alive without relapse/GvHD' if their symptoms disappear. If they relapse during GvHD, they move to the state Relapse: this model considers relapse a more important failure than GvHD.

components of the recovery process and can translate them to the total impact on clinically relevant outcome measures. Despite appearing to be the ideal framework for analyzing complex recovery patterns of patients with hematological diseases, a recent systematic review by Bonneville et al. on studies reporting multivariable Cox proportional hazards models in the setting of malignant hematological diseases showed that only 2 of the 299 included papers involved a multi-state model.[42] This is likely due to the requirement of high-quality clinical data and sufficient clinical, biological and statistical knowledge to translate a clinical research question into a multi-state model. Multi-state modelling demands careful choices in which clinical events are relevant, which transitions are allowed and which risk factors should be modelled in which way for which transitions.

Joint model to investigate effects of biomarkers

The main limitations of the time-dependent Cox model for analyzing biomarkers such as MRD markers and lymphocyte counts are the assumptions that the measurement values are constant between visits, that there is no measurement error, and that the availability of the measurements is not related to the failure status.[43] The latter indicates that the biomarker needs to be exogeneous, which is by definition untrue. The joint model does not depend on these assumptions.[44] It captures biomarkers and clinical events simultaneously by linking two submodels, one for the longitudinal measurements and one for the risks of the clinical events, via an association structure (Figure 2). This allows to model the measurement trajectories over time (which are not yielded by the time-dependent Cox model) while appropriately accounting for both the heterogeneity in subject-specific trajectories and measurement error, and enables the estimation of an association between the longitudinal measurements and the risks of clinical events.

While joint models seem to be the method of choice for analyzing the impact of biomarkers on survival outcomes, they have been applied rarely in the field of hematology.[45,46] As for multi-state models, their disuse is likely due to the required clinical, biological and statistical knowledge to correctly specify the model.

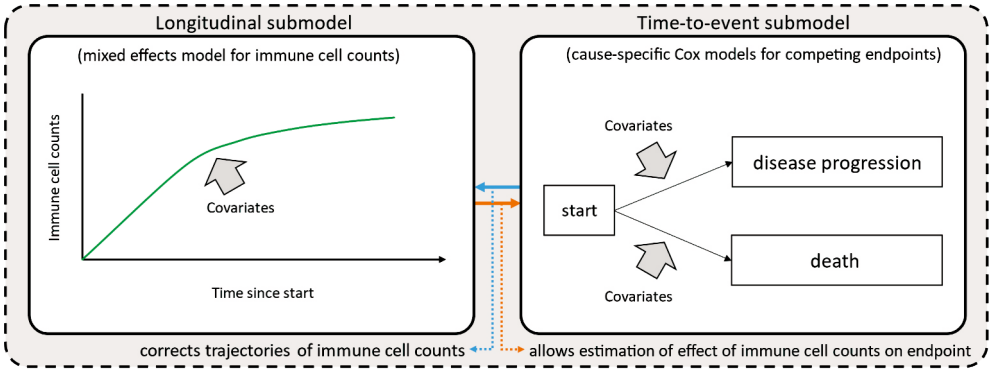


Figure 2. Example of a joint model. The model consists of two components: a longitudinal submodel for the immune cell counts and a time-to-event submodel for the clinical events, which are linked together via an association structure.

Aim of the thesis

The aim of this thesis is to investigate how careful selection of a specific setting and application of advanced statistical methodology such as multi-state and joint models can be used to investigate complex mechanisms or research questions in observational studies in the field of hematology. **Chapters 2-5** investigate the lymphohematopoietic and clinical recovery after alloSCT for acute leukemia and **Chapter 6** investigates the multi-step treatment of and recovery after IST for AA.

In **Chapter 2**, we aim to investigate how selection of a specific alloSCT strategy, TCD alloSCT followed by prophylactic or preemptive DLI, can be used to disentangle the effects of competitive repopulation and allo-immunological pressure on the patient- and donor-derived lymphohematopoietic recovery. The reduction of allo-immunological pressure early after alloSCT by the TCD will provide us an opportunity to investigate the impact of the conditioning intensity on the competitive repopulation in the absence of strong allo-immunological pressure. By selecting a cohort with different DLI strategies, prophylactic versus preemptive DLI and with different starting times of first DLI based on the anticipated relapse risk, we will be able to study the impact of introducing allo-immunological pressure after the competitive repopulation has taken place.

In **Chapter 3**, we will dive deeper into the immune cell kinetics after alloSCT and aim to investigate the complex associations between immune cell kinetics and alloreactivity by using joint modeling. Also in this case, we will use a setting of TCD alloSCT followed by DLI to study the impact of DLI on the immune cell kinetics. The joint model framework will also enable us to estimate the impact of the number of circulating immune cells on the risks of GvHD and relapse. However, we will need to take into account that the actual administration of DLI not only depends on the treatment plan, but also on the clinical circumstances, which may influence the immune cell counts. To take this properly into account, we will perform an intention-to-treat analysis.

In **Chapter 4**, we will focus on the clinical outcomes after DLI and aim to identify factors that influence the alloreactivity of DLI, taking into account the dynamic nature of GvHD, which can lead to death, resolve, and decrease the risk of relapse, by using a multi-state model. We will investigate the effects of conditioning intensity, donor type, presence of patient-derived APCs in the BM, lymphopenia and viral infections in relation to the timing and dose of the DLI. Using the multi-state framework the clinical relevance of any found associations will be demonstrated by assessing the impact of these risk factors on different outcomes after DLI, such as cGRFS.

In **Chapter 5**, we aim to investigate how the transplantation strategy affects the alloreactivity of DLI by considering a different clinical setting, PTCY alloSCT followed by DLI, than in the previous chapters. By keeping the patient selection and interventions after alloSCT similar, the impact of the transplantation strategies on the conditions at the time of DLI and the alloreactivity of DLI can be investigated. We will assess chimerism conversion after DLI and compare this with the results of **Chapter 2**, and compare the DLI conditions and risk of DLI-induced GvHD with the results of **Chapter 4**.

In **Chapter 6**, we move to the AA setting and aim to investigate whether and how the multi-state framework can be used to develop a dynamic measure of "treatment success" that can better capture the complex clinical recovery and failure patterns of patients with

AA receiving IST compared to conventional analysis approaches. We will use the model to evaluate treatment outcome in different age groups. The multi-state framework will also allow us to investigate the effects of risk factors such as age and the presence of a GPI-deficient cell clone on different phases of the recovery and assess their impact on overall treatment outcomes.

In **Chapter 7**, the results of this thesis will be summarized and discussed in the light of the current literature and other methodological approaches.

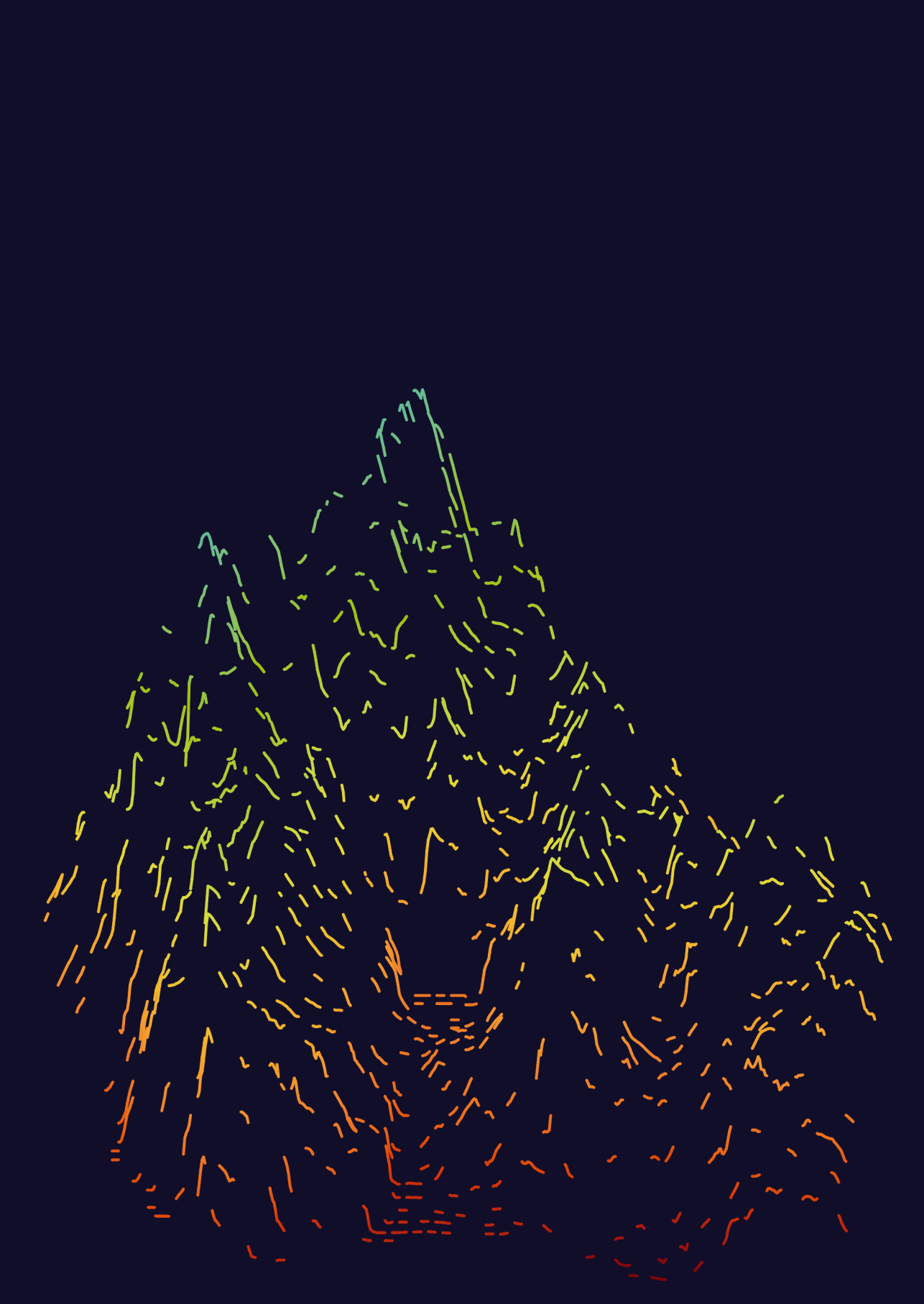
REFERENCES

1. Horowitz, M., Gale, R., Sondel, P. et al. Graft-versus-leukemia reactions after bone marrow transplantation. Blood 75, 555-562 (1990). doi: 10.1182/blood.V75.3.555.555
2. Falkenburg, J. H. F. & Jedema, I. Graft versus tumor effects and why people relapse. Hematology. American Society of Hematology. Education Program 2017, 693-698 (2017). doi: 10.1182/asheducation-2017.1.693
3. Beatty, P. G., Clift, R. A., Mickelson, E. M. et al. Marrow transplantation from related donors other than HLA-identical siblings. N Engl J Med 313, 765-771 (1985). doi: 10.1056/nejm198509263131301
4. Martin, P. J., Levine, D. M., Storer, B. E. et al. Genome-wide minor histocompatibility matching as related to the risk of graft-versus-host disease. Blood 129, 791-798 (2017). doi: 10.1182/blood-2016-09-737700
5. Shouval, R., Fein, J. A., Labopin, M. et al. Outcomes of allogeneic haematopoietic stem cell transplantation from HLA-matched and alternative donors: a European Society for Blood and Marrow Transplantation registry retrospective analysis. Lancet Haematol 6, e573-e584 (2019). doi: 10.1016/s2352-3026(19)30158-9
6. Storb, R. & Sandmaier, B. M. Nonmyeloablative allogeneic hematopoietic cell transplantation. Haematologica 101, 521-530 (2016). doi: 10.3324/haematol.2015.132860
7. Quesenberry, P. J., Colvin, G. & Abedi, M. Perspective: fundamental and clinical concepts on stem cell homing and engraftment: a journey to niches and beyond. Exp Hematol 33, 9-19 (2005). doi: 10.1016/j.exphem.2004.10.012
8. Elfeky, R., Lazareva, A., Qasim, W. & Veys, P. Immune reconstitution following hematopoietic stem cell transplantation using different stem cell sources. Expert review of clinical immunology 15, 735-751 (2019). doi: 10.1080/1744666x.2019.1612746
9. Ghimire, S., Weber, D., Mavin, E. et al. Pathophysiology of GvHD and Other HSCT-Related Major Complications. Front Immunol 8, 79 (2017). doi: 10.3389/fimmu.2017.00079
10. Ferrara, J. L., Levine, J. E., Reddy, P. & Holler, E. Graft-versus-host disease. Lancet 373, 1550-1561 (2009). doi: 10.1016/s0140-6736(09)60237-3
11. Chakraverty, R. & Sykes, M. The role of antigen-presenting cells in triggering graft-versus-host disease and graft-versus-leukemia. Blood 110, 9-17 (2007). doi: 10.1182/blood-2006-12-022038
12. Saad, A. & Lamb, L. S. Ex vivo T-cell depletion in allogeneic hematopoietic stem cell transplant: past, present and future. Bone Marrow Transplant 52, 1241-1248 (2017). doi: 10.1038/bmt.2017.22
13. Barge, R. M., Starrenburg, C. W., Falkenburg, J. H. et al. Long-term follow-up of myeloablative allogeneic stem cell transplantation using Campath "in the bag" as T-cell depletion: the Leiden experience. Bone Marrow Transplant 37, 1129-1134 (2006). doi: 10.1038/sj.bmt.1705385
14. Busca, A. & Aversa, F. In-vivo or ex-vivo T cell depletion or both to prevent graft-versus-host disease after hematopoietic stem cell transplantation. Expert Opin Biol Ther 17, 1401-1415 (2017). doi: 10.1080/14712598.2017.1369949
15. Soiffer, R. J., Lerademacher, J., Ho, V. et al. Impact of immune modulation with anti-T-cell antibodies on the outcome of reduced-intensity allogeneic hematopoietic stem cell transplantation for hematologic malignancies. Blood 117, 6963-6970 (2011). doi: 10.1182/blood-2011-01-332007
16. Bacigalupo, A. Antilymphocyte/thymocyte globulin for graft versus host disease prophylaxis:

efficacy and side effects. Bone Marrow Transplant 35, 225-231 (2005). doi: 10.1038/sj.bmt.1704758
17. Soiffer, R. J., Kim, H. T., McGuirk, J. et al. Prospective, Randomized, Double-Blind, Phase III Clinical Trial of Anti-T-Lymphocyte Globulin to Assess Impact on Chronic Graft-Versus-Host Disease-Free Survival in Patients Undergoing HLA-Matched Unrelated Myeloablative Hematopoietic Cell Transplantation. J Clin Oncol 35, 4003-4011 (2017). doi: 10.1200/jco.2017.75.8177
18. Nunes, N. S. & Kanakry, C. G. Mechanisms of Graft-versus-Host Disease Prevention by Post-transplantation Cyclophosphamide: An Evolving Understanding. Front Immunol 10, 2668 (2019). doi: 10.3389/fimmu.2019.02668
19. Zhao, C., Bartock, M., Jia, B. et al. Post-transplant cyclophosphamide alters immune signatures and leads to impaired T cell reconstitution in allogeneic hematopoietic stem cell transplant. Journal of Hematology & Oncology 15, 64 (2022). doi: 10.1186/s13045-022-01287-3
20. Montoro, J., Roldán, E., Piñana, J. L. et al. Ex vivo T-cell depletion vs post-transplant cyclophosphamide, sirolimus, and mycophenolate mofetil as graft-vs-host disease prophylaxis for allogeneic hematopoietic stem cell transplantation. European journal of haematology 106, 114-125 (2021). doi: 10.1111/ejh.13529
21. Nagler, A., Labopin, M., Dholaria, B. et al. Graft-versus-Host Disease Prophylaxis with Post-Transplantation Cyclophosphamide versus Cyclosporine A and Methotrexate in Matched Sibling Donor Transplantation. Transplant Cell Ther 28, 86.e81-86.e88 (2022). doi: 10.1016/j.jtct.2021.11.013
22. Falkenburg, J. H. F., Schmid, C., Kolb, H. J. & Kuball, J. in The EBMT Handbook: Hematopoietic Cell Transplantation and Cellular Therapies (eds Anna Sureda et al.) 531-539 (Springer International Publishing, 2024).
23. Schmid, C., Labopin, M., Nagler, A. et al. Donor lymphocyte infusion in the treatment of first hematological relapse after allogeneic stem-cell transplantation in adults with acute myeloid leukemia: a retrospective risk factors analysis and comparison with other strategies by the EBMT Acute Leukemia Working Party. J Clin Oncol 25, 4938-4945 (2007). doi: 10.1200/jco.2007.11.6053
24. Schmid, C., Labopin, M., Nagler, A. et al. Treatment, risk factors, and outcome of adults with relapsed AML after reduced intensity conditioning for allogeneic stem cell transplantation. Blood 119, 1599-1606 (2012). doi: 10.1182/blood-2011-08-375840
25. Eefting, M., von dem Borne, P. A., de Wreede, L. C. et al. Intentional donor lymphocyte-induced limited acute graft-versus-host disease is essential for long-term survival of relapsed acute myeloid leukemia after allogeneic stem cell transplantation. Haematologica 99, 751-758 (2014). doi: 10.3324/haematol.2013.089565
26. Young, N. S., Calado, R. T. & Scheinberg, P. Current concepts in the pathophysiology and treatment of aplastic anemia. Blood 108, 2509-2519 (2006). doi: 10.1182/blood-2006-03-010777
27. Pool, E. S., Luk, S. J., Ijsselsteijn, M. E. et al. Imaging mass cytometry reveals the order of events in the pathogenesis of immune-mediated aplastic anemia. Blood 146, 951-963 (2025). doi: 10.1182/blood.2025028723
28. Zhang, Y., Huo, J., Liu, L. et al. Comparison of Hematopoietic Stem Cell Transplantation Outcomes Using Matched Sibling Donors, Haploidentical Donors, and Immunosuppressive Therapy for Patients With Acquired Aplastic Anemia. Front Immunol 13, 837335 (2022). doi: 10.3389/fimmu.2022.837335
29. Kulasekararaj, A., Cavenagh, J., Dokal, I. et al. Guidelines for the diagnosis and management of adult aplastic anaemia: A British Society for Haematology Guideline. Br J Haematol 204, 784-804 (2024). doi: 10.1111/bjh.19236
30. Peffault de Latour, R., Kulasekararaj, A., Iacobelli, S. et al. Eltrombopag Added to Immunosuppression in Severe Aplastic Anemia. N Engl J Med 386, 11-23 (2022). doi: 10.1056/NEJMoa2109965
31. Young, N. S. Aplastic Anemia. N Engl J Med 379, 1643-1656 (2018). doi: 10.1056/NEJMra1413485
32. Tjon, J. M., de Groot, M. R., Sypkens Smit, S. M. A. et al. Short-term efficacy and safety of antithymocyte globulin treatment in elderly patients with acquired aplastic anaemia. Br J Haematol 180, 459-462 (2018). doi: 10.1111/bjh.14372

33. Prabahran, A., Durrani, J., Coelho-Da Silva, J. et al. Safety and efficacy of immunosuppressive therapy for elderly patients with severe aplastic anaemia. Br J Haematol 205, 1170-1179 (2024). doi: 10.1111/bjh.19648
34. Fattizzo, B., Gurnari, C., Giammarco, S. et al. Elderly Patients With Aplastic Anemia: Treatment Patterns and Outcomes in the Real World. Am J Hematol 100, 584-591 (2025). doi: 10.1002/ajh.27611
35. Holtan, S. G., DeFor, T. E., Lazaryan, A. et al. Composite end point of graft-versus-host disease-free, relapse-free survival after allogeneic hematopoietic cell transplantation. Blood 125, 1333-1338 (2015). doi: 10.1182/blood-2014-10-609032
36. Anderson, J. R., Cain, K. C. & Gelber, R. D. Analysis of survival by tumor response. Journal of Clinical Oncology 1, 710-719 (1983). doi: 10.1200/jco.1983.1.11.710
37. Zhou, J. Y., Wang, S., Yuan, H. L. et al. Impact of a novel prognostic model on allogeneic hematopoietic stem cell transplantation outcomes in patients with CMML. Am J Hematol 98, 1394-1406 (2023). doi: 10.1002/ajh.26999
38. Patel, B. A., Groarke, E. M., Lotter, J. et al. Long-term outcomes in patients with severe aplastic anemia treated with immunosuppression and eltrombopag: a phase 2 study. Blood 139, 34-43 (2022). doi: 10.1182/blood.2021012130
39. Chattopadhyay, S., Lionel, S., Selvarajan, S. et al. Relapse and transformation to myelodysplastic syndrome and acute myeloid leukemia following immunosuppressive therapy for aplastic anemia is more common as compared to allogeneic stem cell transplantation with a negative impact on survival. Annals of Hematology 103, 749-758 (2024). doi: 10.1007/s00277-024-05621-2
40. Solomon, S. R., Sizemore, C., Zhang, X. et al. Current Graft-versus-Host Disease-Free, Relapse-Free Survival: A Dynamic Endpoint to Better Define Efficacy after Allogenic Transplant. Biol Blood Marrow Transplant 23, 1208-1214 (2017). doi: 10.1016/j.bbmt.2017.02.022
41. Eefting, M., de Wreede, L. C., Halkes, C. J. et al. Multi-state analysis illustrates treatment success after stem cell transplantation for acute myeloid leukemia followed by donor lymphocyte infusion. Haematologica 101, 506-514 (2016). doi: 10.3324/haematol.2015.136846
42. Bonneville, E. F., Schetelig, J., Putter, H. & de Wreede, L. C. Handling missing covariate data in clinical studies in haematology. Best Practice & Research Clinical Haematology 36, 101477 (2023). doi: 10.1016/j.beha.2023.101477
43. Baart, S. J., van der Palen, R. L. F., Putter, H. et al. Joint Modeling of Longitudinal Markers and Time-to-Event Outcomes: An Application and Tutorial in Patients After Surgical Repair of Transposition of the Great Arteries. Circ Cardiovasc Qual Outcomes 14, e007593 (2021). doi: 10.1161/circoutcomes.120.007593
44. Rizopoulos, D. (Boca Raton: Chapman and Hall/CRC, 2012).
45. Huang, A., Chen, Q., Fei, Y. et al. Dynamic prediction of relapse in patients with acute leukemias after allogeneic transplantation: Joint model for minimal residual disease. International Journal of Laboratory Hematology 43, 84-92 (2021). doi: 10.1111/ijlh.13328
46. Tang, X., Alatrash, G., Ning, J. et al. Increasing chimerism after allogeneic stem cell transplantation is associated with longer survival time. Biol Blood Marrow Transplant 20, 1139-1144 (2014). doi: 10.1016/j.bbmt.2014.04.003

1



2

Competitive repopulation and allo-immunological pressure determine chimerism kinetics after T-cell depleted allogeneic stem cell transplantation and donor lymphocyte infusion

Eva A.S. Koster, Peter A. von dem Borne, Peter van Balen,
Esther H.M. van Egmond, Erik W.A. Marijt, Sabrina A.J. Veld, Inge Jedema,
Tjeerd J.F. Snijders, Daniëlle van Lammeren, Hendrik Veelken,
J.H. Frederik Falkenburg, Liesbeth C. de Wreede, Constantijn J.M. Halkes

Transplant Cell Ther. 2023 Apr;29(4):268.e1-268.e10.

ABSTRACT

After allogeneic stem cell transplantation (alloSCT), patient-derived stem cells that survived the pretransplant conditioning compete with engrafting donor stem cells for bone marrow (BM) repopulation. In addition, donor-derived alloreactive T cells present in the stem cell product may favor establishment of complete donor-derived hematopoiesis by eliminating patient-derived lymphohematopoietic cells. T-cell depleted alloSCT with sequential transfer of potentially alloreactive T cells by donor lymphocyte infusion (DLI) provides a unique opportunity to selectively study how competitive repopulation and allo-immunological pressure influence lymphohematopoietic recovery. This study aimed to determine the relative contribution of competitive repopulation and donor-derived anti-recipient allo-immunological pressure on the establishment of lymphohematopoietic chimerism after alloSCT. In this retrospective cohort study of 281 acute leukemia patients treated according to a protocol combining alemtuzumab-based T-cell depleted alloSCT with prophylactic DLI, we investigated engraftment and quantitative donor chimerism in the BM and immune cell subsets. DLI-induced increase of chimerism and development of Graft-versus-Host-Disease (GvHD) were analyzed as complementary indicators for donor-derived anti-recipient allo-immunological pressure. Profound suppression of patient immune cells by conditioning sufficed for sustained engraftment without necessity for myeloablative conditioning or development of clinically significant GvHD. Although 61% of the patients without any DLI or GvHD showed full donor chimerism (FDC) in the BM at 6 months after alloSCT, only 24% showed FDC in the CD4+ T-cell compartment. In contrast, 75% of the patients who had received DLI and 83% of the patients with clinically significant GvHD had FDC in this compartment. In addition, 72% of the patients with mixed hematopoiesis receiving DLI converted to complete donor-derived hematopoiesis, of whom only 34% developed clinically significant GvHD. Our data show that competitive repopulation can be sufficient to reach complete donor-derived hematopoiesis, but that some allo-immunological pressure is needed for the establishment of a completely donor-derived T-cell compartment, either by the development of GvHD or by administration of DLI. We illustrate that it is possible to separate the Graft-versus-Leukemia effect from GvHD, as conversion to durable complete donor-derived hematopoiesis following DLI did not require induction of clinically significant GvHD.

INTRODUCTION

The goal of allogeneic hematopoietic stem cell transplantation (alloSCT) in patients with hematological malignancies is to eradicate the disease by replacing patient hematopoiesis with donor-derived hematopoiesis and introducing donor alloreactive T cells capable of eliminating residual malignant cells. After alloSCT, patient hematopoietic stem cells (HSCs) that survived the pretransplant conditioning compete with engrafting donor HSCs for bone marrow (BM) repopulation.[1] Patient-derived alloreactive T cells may reject the graft[2], but donor engraftment can be supported by alloreactive donor-derived T cells recognizing nonself antigens on patient immune cells.[3,4] These alloreactive donor T cells can further eliminate patient HSCs and residual malignant cells and provide lasting immune surveillance against the malignancy, the Graft-versus-Leukemia (GvL) effect. However, when non-hematopoietic tissues of the patient are recognized, Graft-versus-Host-Disease (GvHD) may develop.[5,6]

Allo-immune responses are induced by presentation of antigens to functional alloreactive T cells. To become properly activated, naïve alloreactive T cells require costimulatory signals from activated professional antigen-presenting cells (APCs). Important factors influencing the balance between donor- and patient-derived allo-immunological pressure include greater genetic disparity between patient and donor encoding more antigens that can induce alloreactivity[7,8]. Activation of professional APCs by tissue damage may increase the risk of GvHD after more toxic myeloablative (MA) compared with less toxic nonmyeloablative (NMA) conditioning regimens.[9] The risk of GvHD decreases over time as the professional patient APCs are gradually replaced by donor-derived APCs.[10] Finally, the recovery of regulatory T cells and *de novo* lymphopoiesis after transplantation may lead to a state of tolerance.[11,12]

Several strategies to modulate the allo-immunological pressure after alloSCT have been developed. Most patients receive prophylactic systemic immunosuppression with or without a form of T-cell depletion (TCD) to prevent rejection and GvHD.[13] The impact of TCD on patient- and/or donor-derived T cells depends on the method (*in vivo* versus *in vitro*) and timing (before or after alloSCT).[14-18] Excessive suppression of donor-derived allo-immunological pressure against the patient immune cells, HSCs, and tumor cells favors their persistence and eventual dominance, with the risk of graft failure and/or recurrence of the malignancy.[19-21] To improve engraftment or boost the GvL effect, unmodified donor lymphocyte infusion (DLI) can be administered after alloSCT.[22-25]

Our strategy of alemtuzumab-based TCD alloSCT followed by standard prophylactic DLI[26] aims to separate the establishment of donor hematopoiesis from the introduction of donor alloreactivity. The delayed introduction of donor alloreactivity allows the induction of a GvL effect without a high risk of GvHD necessitating systemic treatment. In this context, TCD permits analysis of BM repopulation in the absence of strong allo-immunological pressure. Obviation of the need for prophylactic pharmacologic immunosuppression facilitates analysis of natural immunological recovery. DLI is administered starting 3 months after alloSCT after the competitive repopulation of the BM and early T-cell expansion have taken place. This setting offers the unique opportunity to exclusively analyze the impact of donor alloreactivity introduced by DLI on persisting patient-derived HSCs and T cells. In a cohort of 281 patients, we examined lymphohematopoietic recovery and chimerism kinetics in the BM and circulating

immune cells in relation to conditioning and GvHD after TCD alloSCT and DLI.

METHODS

Study population

This observational study included all adult patients with acute myeloid leukemia (AML), acute lymphoblastic leukemia, or myelodysplastic syndrome in complete morphologic remission after intensive induction therapy who underwent a first BM or G-CSF-mobilized peripheral blood stem cell alloSCT using a standard conditioning and TCD protocol[27-29] at Leiden University Medical Center between 2005 and 2015. Exclusion criteria were use of a haploidentical donor or prescheduled posttransplant chemotherapy. The study was approved by the Medical Ethical Committee of Leiden University Medical Center (P03.114, P03.173, and P04.003). All patients provided signed informed consent for data collection and analysis. Data were analyzed as of August 2020.

Transplantation and DLI strategy

MA conditioning consisted of cyclophosphamide (60 mg/kg i.v. for 2 days) with 9 Gy total body irradiation or busulfan (4x0.8 mg/kg i.v. for 4 days). NMA conditioning consisted of fludarabine (50 mg/m^2 orally for 6 days) and busulfan (4x0.8 mg/kg i.v. for 2 days).

Standard *in vitro* TCD was performed by adding 20 mg of alemtuzumab (Sanofi Genzyme) to the graft prior to infusion.[28] Additional *in vivo* TCD depended on donor type and conditioning regimen. MA-conditioned patients with a 10/10 HLA-matched related donor (RD) did not receive any *in vivo* TCD; all other patients received 15 mg alemtuzumab i.v. on days -6 and -5 for MA conditioning or on days -4 and -3 for NMA conditioning. NMA-conditioned patients with an unrelated donor (UD) also received rabbit anti-thymocyte globulin (Sanofi Genzyme) on day -2 (2 mg/kg until April 2010 and 1 mg/kg thereafter). Only MA-conditioned patients with an UD or a 9/10 HLA-matched RD received ciclosporin as GvHD prophylaxis, which was tapered from 1 month and stopped within 3 months post-transplantation.

Preemptive DLI was administered for increasing or persisting mixed chimerism (MC) or the presence of minimal residual disease (MRD), whereas prophylactic DLI was administered regardless of chimerism or MRD status to all patients without clinically significant GvHD indicating allo-immunological pressure. Preemptive DLI was administered at escalating doses with ≥3 months between infusions until the development of GvHD or disappearance of MRD and/or MC. Prophylactic DLI was introduced in May 2010 and administered to all patients at 6 months post-transplantation in the absence of clinically significant GvHD or relapse (3M T cells/kg for patients with RD and 1.5M T cells/kg for patients with UD). Forty-two patients at high risk of early relapse (within 6 months post-alloSCT) also received prophylactic low-dose DLI (0.3M and 0.15M T cells/kg for RD and UD, respectively) at 3 months.[30] Indications for this early prophylactic DLI were very poor risk AML or high-risk acute lymphoblastic leukemia according to the HOVON criteria[31,32], MRD positivity at time

of transplantation, incomplete pretransplant treatment, therapy-related AML with unfavorable karyotype, AML with persisting underlying disease, and acute leukemia that relapsed early after the previous curative induction chemotherapy. Along with the unmodified DLI, patients could receive modified (purified or genetically modified) T-cell products as part of several clinical trials. Interferon could be administered to patients with an increasing number of blasts in the BM (but morphologically below 5%) or with MRD not responding to DLI.

Follow-up

BM cytology, lymphocyte counts, and BM chimerism analysis were performed at least every 3 months during the first 2 years post-transplantation. In a subset of patients, chimerism was also determined in granulocytes, monocytes, CD4+ T cells, CD8+ T cells, natural killer cells, and B cells on peripheral blood (Supplemental Methods). Because we were interested in the kinetics of chimerism following TCD alloSCT and unmodified DLI, we excluded all chimerism and cell count measurements obtained 1 week or longer after infusion of a modified T-cell product. Measurements obtained after relapse, interferon administration, chemotherapy, and second alloSCT were excluded as well. Measurements used in the comparison analyses were performed at 6 weeks (actual range 5-7 weeks), 3 months (2-4 months), and 6 months (5-7 months) after alloSCT. For the clinical outcomes, all patients, including those who received a modified T-cell product, were included to assess the overall outcome of our total strategy and to prevent selection bias, as recruitment for trials with modified T-cell products started a few weeks after alloSCT.

Definitions

Neutrophil recovery was defined as the first of 3 consecutive measurements with an absolute neutrophil count $>0.5\times10^9$/l, and non-engraftment was defined as detection of <5% donor BM chimerism without prior neutrophil recovery. The date of relapse was defined as the date of the first recurrence of ≥5% blasts on cytomorphologic BM examination or ≥1% blasts in peripheral blood after alloSCT (confirmed by BM biopsy if possible). Clinically significant GvHD was defined by therapeutic systemic immunosuppression (tIS) for GvHD for at least 14 days or until death or stopped as part of palliative care due to refractory GvHD, or continued use of GvHD prophylaxis beyond 3.5 months. tIS was started for acute GvHD grade II-IV according to the modified Glucksberg criteria, for extensive chronic GvHD according to the Seattle criteria, and for mild GvHD not responding to topical treatment.[33,34] We used tIS instead of the exact grading because our strategy aims to prevent GvHD necessitating systemic treatment. MC was defined as detection of patient DNA at or above the limit of detection, and full donor chimerism (FDC) was defined as undetectable patient DNA. DLI-induced cytopenia was defined as the development of severe neutropenia ($<0.5\times10^9$/l) after DLI not caused by relapse or infection.

Study objective and endpoints

The objective of this study was to determine the relative contribution of competitive repopulation and donor-derived anti-recipient allo-immunological pressure on the establishment of full-donor lymphohematopoietic chimerism after alloSCT. DLI-

induced increase in chimerism and development of GvHD were analyzed as complementary indicators for donor-derived anti-recipient allo-immunological pressure. The primary endpoint was the level of donor BM and T-cell chimerism at 6 weeks and 3 and 6 months after alloSCT. Secondary endpoints were primary engraftment, clinically significant GvHD as a surrogate for allo-immunological pressure, and BM chimerism kinetics during the first 2 years after DLI. Other secondary endpoints were overall survival, relapse-free survival, cumulative incidence of relapse, and non-relapse mortality during the first 5 years after alloSCT.

Analyses

An algorithm was developed to assess the chimerism response after the first unmodified DLI that patients received while having mixed BM chimerism. We defined this DLI as the 'starting DLI' for this analysis (Supplemental Methods).

The probabilities of overall survival and relapse-free survival from alloSCT with 95% confidence intervals (95%-CIs) were calculated using the Kaplan-Meier method. Follow-up from alloSCT was quantified using the reverse Kaplan-Meier method.[35] Cumulative incidences of neutrophil recovery as proxy for primary engraftment and clinically significant GvHD were calculated using competing risks models (Supplemental Methods).

To evaluate the effects of donor-derived allo-immunological pressure and DLI on BM repopulation and immunological recovery, donor chimerism in the BM and T cells was evaluated at 3 and 6 months after alloSCT and compared between groups based on whether patients had developed clinically significant GvHD, had received unmodified DLI without any clinically significant GvHD, or had neither. Because chimerism levels did not follow a normal distribution, groups were compared using the Mann-Whitney U test (2 groups) or Kruskall-Wallis test followed by, if applicable, the post hoc Dunn test with Holms adjustment for multiple comparisons (>2 groups). An (adjusted) p-value <0.05 was considered significant.

Software

All analyses were performed in R version 4.0.2 using the survival, cmprsk, prodlim, rstatix, ggplot2, ggpubr, gridExtra, and ggalluvial packages.

RESULTS

Population

A total of 281 patients were included in this study. The patients' baseline characteristics are summarized in Table 1. The median follow-up was 61 months (interquartile range [IQR] 43-85 months) after alloSCT. The clinical outcomes of our total strategy of TCD alloSCT followed by DLI are presented in Supplemental Results.

Successful primary engraftment after TCD alloSCT does not depend on MA conditioning or donor-derived allo-immunological pressure

The cumulative incidence of neutrophil recovery was 91% (95%-CI 88-94) at 4 weeks

after alloSCT and increased to 99% (95%-CI 97-100) at 2.5 months (Supplemental Figure 3). One patient, who underwent transplantation after MA conditioning, failed to engraft. Two patients died, at 2 and 12 days after alloSCT, before (non-)engraftment. Successful engraftment of all 103 evaluable NMA-conditioned patients demonstrates sufficient suppression of the patient immune cells by alemtuzumab, in combination with anti-thymocyte globulin in case of an UD, to prevent graft rejection.

To evaluate whether strong donor-derived alloimmune responses after alloSCT had a profound role in the primary engraftment in this cohort, we examined the development of clinically significant GvHD before any DLI after alloSCT in the 278 engrafted patients. At 3 months after alloSCT, the cumulative incidence of clinically significant GvHD was 13% (95%-CI 9-17) in the total cohort and only 2% (95%-CI 0-5) after NMA conditioning (Supplemental Figure 4). Together with the 99% probability of

	Total (N = 281)	MA, matched RD (N = 78)	MA, mismatched RD or UD (N = 99)	NMA, matched RD (N = 41)	NMA, UD (N = 63)
Age at alloSCT (years)					
median (range)	50 (18-73)	43 (18-60)	42 (19-59)	61 (28-72)	63 (40-73)
Disease					
AML	188 (67%)	47 (60%)	56 (57%)	33 (80%)	52 (83%)
ALL	76 (27%)	26 (33%)	39 (39%)	5 (12%)	6 (10%)
MDS	17 (6%)	5 (6%)	4 (4%)	3 (7%)	5 (8%)
Conditioning regimen					
Cyclo/TBI	171 (61%)	76 (97%)	95 (96%)	0 (0%)	0 (0%)
Cyclo/Bu	6 (2%)	2 (3%)	4 (4%)	0 (0%)	0 (0%)
Flu/Bu	103 (37%)	0 (0%)	0 (0%)	41 (100%)	62 (98%)
Flu/Bu/Cyclo*	1 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (2%)
Donor					
10/10 matched RD	119 (42%)	78 (100%)	0 (0%)	41 (100%)	0 (0%)
9/10 matched RD	2 (1%)	0 (0%)	2 (2%)	0 (0%)	0 (0%)
10/10 matched UD	120 (43%)	0 (0%)	63 (64%)	0 (0%)	57 (90%)
9/10 matched UD	39 (14%)	0 (0%)	33 (33%)	0 (0%)	6 (10%)
8/10 matched UD	1 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)
Graft source					
G-CSF-mobilized PBSC	266 (95%)	69 (88%)	94 (95%)	41 (100%)	62 (98%)
BM	15 (5%)	9 (12%)	5 (5%)	0 (0%)	1 (2%)
Period of alloSCT**					
first transplantation in cohort	2005-01-20	2005-01-20	2005-03-10	2008-08-14	2009-10-06
before May 2010	87 (31%)	35 (45%)	37 (37%)	9 (22%)	6 (10%)
since May 2010	194 (69%)	43 (55%)	62 (63%)	32 (78%)	57 (90%)

Table 1. Baseline characteristics of the total cohort and subgroups based on conditioning intensity and donor type. MA, myeloablative; NMA, nonmyeloablative; RD, related donor; UD, unrelated donor, AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; MDS, myelodysplastic syndrome; Cyclo, cyclophosphamide; TBI, total body irradiation; Bu, busulfan; Flu, fludarabine; G-CSF, granulocyte-colony stimulation factor; PBSC, peripheral blood stem cells; BM, bone marrow; alloSCT, allogeneic stem cell transplantation. *One patient received cyclophosphamide 750 mg/m^2 i.v. for 2 days in the conditioning regimen because a second consolidation course before transplantation was not given. **Prophylactic DLI has been included in the transplantation strategy since May 2010.

engraftment, these data show that primary engraftment after alemtuzumab-based TCD alloSCT was not impaired after MA or NMA conditioning and in the absence of clinically significant GvHD.

MC is more common in the T-cell compartment than in the BM

Because engraftment does not necessarily lead to persistent complete (100%) donor-derived hematopoiesis, we investigated chimerism kinetics in the 278 engrafted patients. Among the 223 patients alive without any prior cellular intervention, chemotherapy, interferon, or relapse and with evaluable BM chimerism at 3 months after alloSCT, 59% had FDC. Within the group with MC, the middle 50% (i.e., the IQR) had donor chimerism levels between 94% and 98%. To investigate whether the circulating immune cells also were of mixed origin after TCD alloSCT, we measured the level of donor chimerism in 6 immune cell types in a subset of patients, again excluding all samples after cellular intervention, chemotherapy, interferon, and relapse (Supplemental Figure 5; Supplemental Figure 6 and Supplemental Table 2 provide the cell counts in the total cohort). At 3 months after alloSCT, 73% to 78% of the patients showed FDC in the granulocytes, monocytes, natural killer cells, and B cells. The IQR of the donor chimerism values with MC in these cells ranged between 96% and 99%. In contrast, only 22% and 28% of the patients showed FDC in the CD4+ and CD8+ T-cell populations, respectively, and the IQR of donor chimerism within T cells with MC was 7% to 92%. Even in patients with complete donor-derived hematopoiesis, circulating T cells could be predominantly of patient origin at 3 months post-alloSCT (Figure 1).

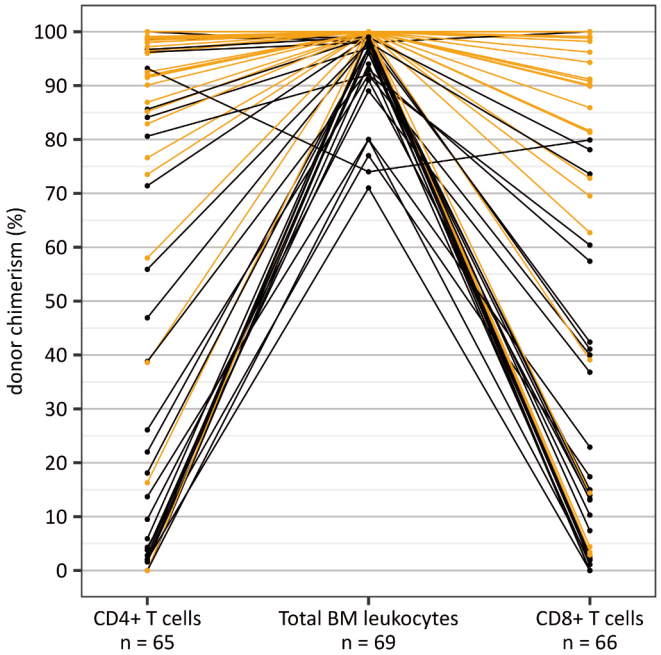


Figure 1. Donor chimerism in the BM and T cells at 3 months. Donor chimerism in the BM, CD4+ and CD8+ T cells at 3 months post-alloSCT without any prior cellular intervention, relapse, chemotherapy or interferon. Lines corresponding to patients with FDC in the BM are highlighted in orange.

Associations between conditioning intensity and clinically significant GvHD and BM and T-cell chimerism at 3 months after alloSCT

Although primary engraftment after TCD alloSCT was not affected by conditioning intensity or donor-derived allo-immunological pressure, these factors could influence the level of donor chimerism. To investigate the influence of conditioning intensity and allo-immunological pressure on the development of complete donor-derived hematopoiesis, we compared BM chimerism at 3 months after alloSCT between patient groups defined by conditioning intensity and development of clinically significant GvHD prior to measurement of chimerism (Figure 2A). In the absence of clinically significant GvHD, MA-conditioned patients had significantly higher donor BM chimerism (71% of the patients had FDC) compared to NMA-conditioned patients (32% FDC), showing that MA conditioning promoted the establishment of complete donor-derived hematopoiesis after alloSCT. To evaluate the effect of GvHD on the development of complete donor-derived hematopoiesis, we compared donor BM chimerism between MA-conditioned patients with and without clinically significant GvHD. Patients with clinically significant GvHD had higher donor BM chimerism at 3 months after alloSCT compared with those without (88% versus 71% FDC; adjusted p-value = 0.12).

To investigate the influence of conditioning intensity and clinically significant GvHD on T-cell chimerism, we compared the level of donor chimerism in CD4+ and CD8+ T cells at 3 months between the same groups for all patients with available T-cell chimerism (Figure 2B-C). In the absence of clinically significant GvHD, donor chimerism in CD4+ T cells and CD8+ T cells was significantly higher in MA-conditioned patients (33% had FDC in CD4+ T cells and 41% had FDC in CD8+ T cells) compared with NMA-conditioned patients (7% and 12%, respectively). In the MA-conditioned group, there was no significant difference in the level of donor chimerism between patients with and those without clinically significant GvHD: 43% versus 33% had FDC in CD4+ T cells (adjusted p-value 0.37) and 43% versus 41% had FDC in CD8+ T cells (adjusted p-value 0.74). Together, these data indicate that myeloablative conditioning led to higher donor T-cell chimerism after TCD alloSCT, but we did not find a significant effect of clinically significant GvHD on the level of CD4+ or CD8+ T-cell chimerism at 3 months after alloSCT. This may be explained by the immunosuppressive treatment that almost all patients with GvHD still were receiving at the time of chimerism measurement.

Donor chimerism in the BM and T cells increases after early DLI

Starting from 3 months, prophylactic and preemptive DLI was administered to induce an alloimmune response against patient-derived hematopoietic cells. To investigate the impact of this allo-immunological pressure by early DLI in the absence of GvHD, we compared donor BM chimerism between 3 and 6 months after alloSCT in patients who received unmodified DLI within 4 months after alloSCT but without any clinically significant GvHD up to 6 months, and in patients without any DLI or GvHD in this period. Of the 71 evaluable patients (51% NMA-conditioned) without any DLI or GvHD during this period, 66% showed FDC at 3 months and 61% did so at 6 months, illustrating that in absence of donor-derived allo-immunological pressure, mixed BM chimerism remained prevalent after TCD alloSCT. Thirty patients received unmodified DLI within 4 months after alloSCT, 30% after NMA conditioning. Notably, although only 38% of these patients showed FDC in the BM at 3 months after alloSCT, this

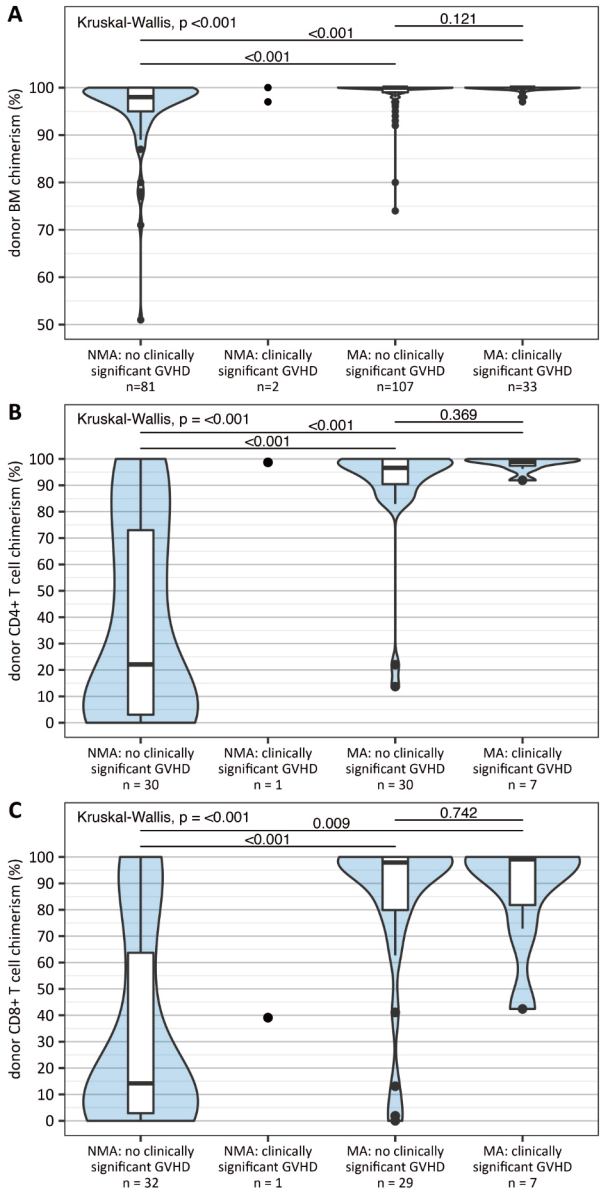


Figure 2. Donor chimerism in the BM and T cells at 3 months according to conditioning regimen intensity and the development of GvHD before the measurement. Donor chimerism in the BM (A), CD4+ T cells (B) and CD8+ T cells (C) at 3 months after TCD alloSCT without any prior cellular intervention, relapse, chemotherapy or interferon. T-cell chimerism was measured in a subset of patients. In 7 patients, either the CD4+ or the CD8+ fraction was missing. In all panels, the data are grouped based on conditioning intensity and development of clinically significant GvHD before the 3-month measurement. The boxplots are combined with violin plots showing the kernel probability density to visualize the distribution of the data. The lower and upper hinges of the boxplots correspond to the 25th and 75th percentiles, respectively. In each panel, the level of donor chimerism was compared among 3 groups as the 2 NMA-conditioned patients with GvHD were excluded from this test. The p-values for the pairwise comparisons are adjusted for multiple comparison.

percentage increased to 63% at 6 months after alloSCT, indicating that early unmodified DLI could increase donor BM chimerism without the concomitant development of GvHD.

To evaluate the impact of early DLI on donor T-cell chimerism, we investigated the kinetics of donor T-cell chimerism during the first 6 months after alloSCT in patients without any DLI or GvHD during this period and in patients who received DLI (Figure 3). Again, the 18 evaluable patients without any DLI or GvHD showed a stable pattern of MC, and almost all patients with an early DLI without any clinically significant GvHD (n=8) showed increasing levels of donor CD4+ and CD8+ T-cell chimerism.

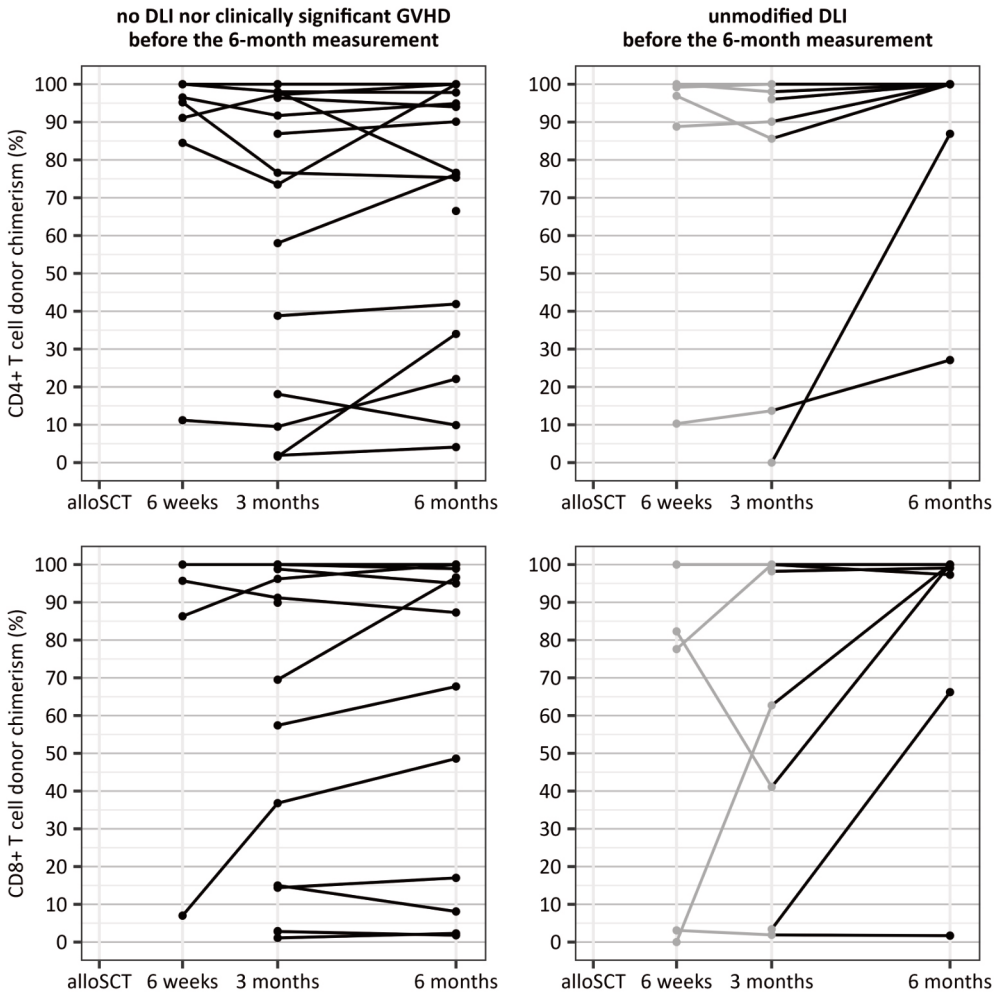


Figure 3. T-cell chimerism kinetics during the first 6 months after TCD alloSCT in patients without any clinically significant GvHD nor DLI before the 6-month measurement and in patients who received an unmodified DLI before the 6-month measurement. Patients who died, relapsed, or received chemotherapy, interferon, second alloSCT or a modified T-cell product before the 6-month measurement were excluded. In the second column, the chimerism measurements done before DLI are in grey, and the measurements done after DLI are in black, to visualize the impact of DLI on the level of donor T-cell chimerism.

To test whether these kinetics led to significant differences in the level of donor T-cell chimerism at 6 months after alloSCT and to compare the impact of DLI with the impact of clinically significant GvHD without DLI on chimerism, we compared the levels of 6-month donor chimerism in CD4+ and CD8+ T cells between these 2 patient groups, adding a third group comprising the 12 evaluable patients who developed clinically significant GvHD before the 6-month measurement, all without any prior DLI (Figure 4A-B). CD4+ T-cell donor chimerism was significantly higher in patients with DLI (75% FDC) or clinically significant GvHD (83%) compared to patients without any DLI or GvHD (24% FDC). CD8+ T-cell chimerism showed a similar trend, with 33% of the

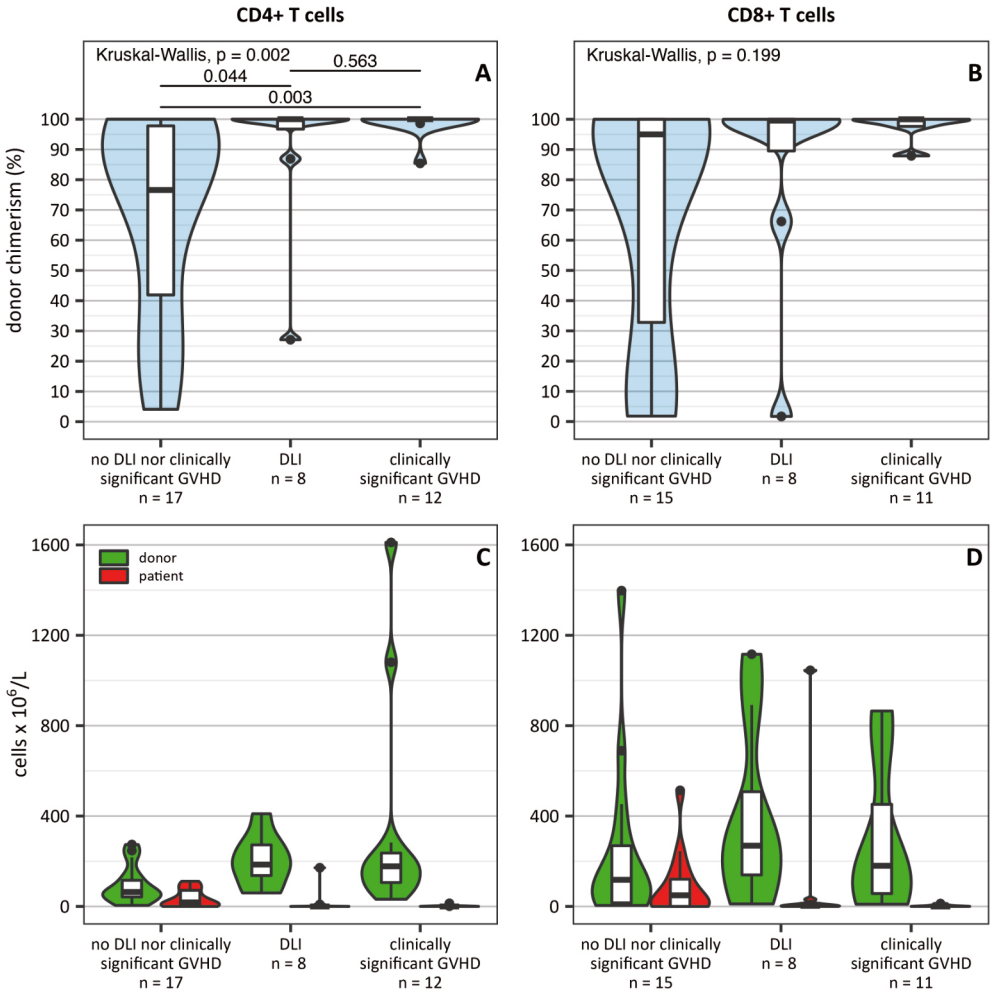


Figure 4. T-cell chimerism and patient/donor-specific counts at 6 months in patients with DLI, GvHD or neither. CD4+ and CD8+ T-cell chimerism (A and B) and patient/donor-specific counts (C and D) at 6 months after alloSCT in patients without prior DLI or clinically significant GvHD, patients with DLI before this measurement, and patients who had developed clinically significant GvHD after alloSCT without prior DLI. Patients who relapsed, or received chemotherapy, interferon, second alloSCT or a modified T-cell product before the 6-month measurement were excluded. The p-values for the pairwise comparisons in (A) are adjusted for multiple comparison.

patients without any DLI or GvHD having FDC, compared to 50% of those with DLI and 55% of those with clinically significant GvHD. Patients with DLI or clinically significant GvHD had both a lower number of circulating patient-derived T cells and a higher number of donor-derived T cells (Figure 4C-D). Together, these data show that 6-month donor CD4+ T-cell chimerism is significantly higher in patients with than in those without GvHD, and that early DLI can increase the level of donor T-cell chimerism in the absence of GvHD.

Strategy of dose-escalating DLIs can convert mixed hematopoiesis to durable complete donor-derived hematopoiesis without necessarily inducing clinically significant GvHD

To investigate the allo-immunological effects of our total DLI strategy, we developed an algorithm to quantify BM chimerism responses in the 65 patients with mixed hematopoiesis receiving unmodified DLI without any prior relapse (Methods). Clinical outcomes of all patients who received an unmodified DLI are presented in the Supplemental Results. The median level of donor BM chimerism in these patients at time of the starting DLI was 98% (IQR 94-99, Figure 5A). In 7 patients, the BM chimerism response could not be evaluated because of early death from severe GvHD after DLI (n=4) or early relapse (n=3) (Figure 5B, Supplemental Table 3). Within 2 months after starting DLI, 15 of the 65 patients (23%) converted to FDC and 9 (14%) showed a partial response with decreasing patient chimerism. Over time and with our dose-escalating DLI protocol, the numbers of patients with a response increased: 38 (58%) showed a response within 4 months and 46 (71%) did so within 7 months after starting DLI. At 25 months, 47 patients (72%) had converted to FDC, including 35 after 1 DLI, and 5 (8%) had shown a partial response. Six patients (9%) with available BM chimerism measurements after DLI did not show any response within this period, of whom 3 relapsed and 2 died within 25 months after the starting DLI. Only 1 patient completed the 25-month follow-up period without showing any chimerism response. Notably, this patient eventually converted to FDC in the BM at 29 months after the starting DLI, 6 months after the fourth DLI. After complete conversion, 4 of the 47 patients died and 4 relapsed within 25 months after the starting DLI. The other 39 (83%) patients with complete conversion were still alive and in complete remission at 25 months after DLI. Only 1 patient occasionally had some detectable patient DNA (Supplemental Figure 7).

To study whether GvHD is required for conversion to complete donor-derived hematopoiesis after DLI, we evaluated the development of clinically significant GvHD in the 47 patients with conversion from MC to FDC and found that 16 (34%) developed clinically significant GvHD within 25 months after starting DLI and 31 (66%) did not (Figure 5C). Together, these data show that our DLI strategy led to durable complete donor-derived hematopoiesis in the majority of the patients with mixed hematopoiesis receiving DLI after TCD alloSCT, without necessarily inducing clinically significant GvHD.

DISCUSSION

In this study, we leveraged a strategy of TCD alloSCT followed by standard prophylactic DLI to investigate how competitive repopulation and allo-immunological pressure influence the lymphohematopoietic recovery after alloSCT. The sequential introduction of donor hematopoiesis and alloreactivity enabled us to study these mechanisms

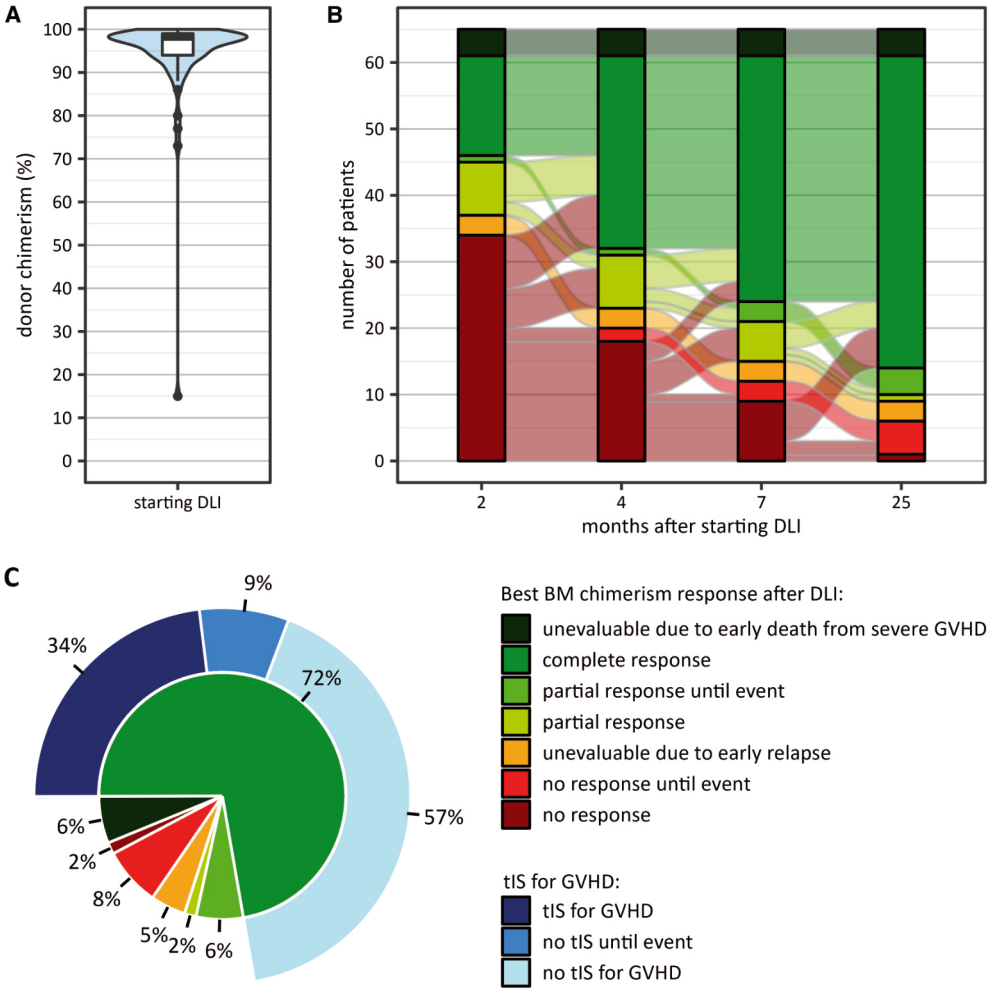


Figure 5. Alloimmune responses in the 65 patients with mixed hematopoiesis receiving unmodified DLI. Patients who relapsed or received chemotherapy, a second alloSCT, or a modified T-cell product before DLI or who continued GvHD prophylaxis after DLI were excluded. (A) The level of donor BM chimerism at time of DLI initiation. (B) The best BM chimerism response achieved by different time points after the first unmodified DLI. The events that terminated the evaluation period (death, relapse, chemotherapy, interferon or second alloSCT) are described in Supplemental Table 3. Note that recurrence of MC or terminating events occurring after a response are not shown in this plot. The current response, which considers these possibilities, is shown in Supplemental Figure 7. (C) Distribution of the best BM chimerism responses at 25 months after initiation of DLI (inner circle) and the use of tIS for GvHD by the converted patients during this period (outer ring). Five of these patients did not start tIS for GvHD but relapsed (n = 4) or died (n = 1) during this period.

separately, which is not possible in T-cell replete alloSCT or TCD alloSCT without standard DLI. Effective suppression of the patient-derived alloimmunity by the conditioning regimens sufficed for sustained engraftment without the need for myeloablative conditioning or evident donor-derived allo-immunological pressure. However, the development of complete donor-derived hematopoiesis depended on both competitive repopulation and allo-immunological pressure; the proportion of patients with FD BM chimerism at 3 months was lowest in the NMA-conditioned patients without any GvHD (32%), higher in the MA-conditioned patients without any GvHD (71%), and highest in the MA-conditioned patients who had developed GvHD (88%). In patients without GvHD, an alloimmune response against patient-derived hematopoietic cells could be efficiently induced by DLI even in the absence of concomitant GvHD. Following our total strategy of dose-escalating DLIs, 72% of the patients with mixed BM chimerism at time of DLI converted to complete donor-derived hematopoiesis. Only 34% of converting patients developed clinically significant GvHD after DLI, illustrating that the GvL effect can be separated from GvHD. For the establishment of a completely donor-derived T-cell compartment, some allo-immunological pressure seemed to be required.

Although the level of donor chimerism in CD4+ and CD8+ T cells at 3 months was higher after MA than NMA conditioning, only 33% and 41% of the MA-conditioned patients without any GvHD had FDC in these subsets, respectively. However, 83% of the patients who had developed GvHD, and 75% of the patients who had received an early DLI without developing any GvHD, had FDC in the CD4+ T cells at 6 months after alloSCT, compared to 24% of the patients without prior DLI or GvHD, showing that DLI also could convert mixed T-cell chimerism to FDC in the absence of GvHD. Together, these data indicate that the establishment of complete donor-derived hematopoiesis can be the result of competitive repopulation, but that donor-derived allo-immunological pressure is needed for the development of FD T-cell chimerism.

Because competitive repopulation can be sufficient to induce FD BM chimerism, the presence of FDC itself does not prove occurrence of an alloimmune response against patient hematopoietic cells or achievement of a meaningful GvL effect. This conclusion can explain why the value of FDC in predicting relapse remains controversial in different settings of alloSCT. For instance, Konuma et al[36] did not observe any association between FDC in the BM and relapse after MA single-unit umbilical cord blood transplantation. Owing to the MA conditioning and the relatively low allo-immunological pressure after cord blood transplantation[37], the achieved FDC might have been mainly the result of competitive repopulation. In contrast, Koreth et al[38] showed that having <90% donor chimerism in the BM or peripheral blood increased the risk of relapse after T-cell replete alloSCT following NMA conditioning. As in this case, the competitive repopulation probably played a more limited role, FDC was more likely a result from donor-derived allo-immunological pressure. The chimerism kinetics also can indicate whether alloreactivity played a role. Although FDC early after transplantation in the absence of GvHD may reflect the outcome of competitive repopulation, conversion from stable MC to FDC is most likely the result of an alloimmune response, leading to low relapse rates after chimerism conversion from MC to FDC, as observed in this study and as reported by others.[39-41] Therefore, not only the level of donor chimerism, but also the clinical setting and the chimerism kinetics, should be considered

when using chimerism to monitor the presence of donor-derived allo-immunological pressure and thereby the GvL effect in patients.

Because a conversion to FDC can be observed only in patients who have MC to start with, the ideal cell lineage for monitoring alloreactivity needs to show stable levels of MC after alloSCT in the majority of patients. Owing to the persistence of long-living patient-derived T cells, mixed T-cell chimerism is common after TCD alloSCT, and can exist in patients with a completely donor-derived hematopoiesis.[41-44] Therefore, changes in T-cell chimerism potentially could be used as marker for alloreactivity more often than BM chimerism. Applicability depends on the transplantation strategy. For instance, mixed T-cell chimerism is common after CD34+-selected alloSCT[46], whereas Carnevale-Schianca et al[45] observed 97% FDC at 28 days after MA alloSCT with posttransplant cyclophosphamide.

An important question is whether DLI can induce a sufficient GvL effect without needing to induce clinically significant GvHD as well. In concordance with our data, others have shown that conversion from MC to FDC can occur in the absence of GvHD after DLI, and that this conversion significantly decreases the risk of relapse.[24,41,47,48] The accumulating evidence that DLI can be effective in preventing relapse even without the induction of GvHD encourages further investigation into how the risk of GvHD after DLI can be decreased without losing the beneficial GvL effect. Several DLI modification strategies are being investigated that either remove cell subsets that are important for the development of GvHD (e.g., depletion of CD8+ T cells) or select only immune cells that target hematopoietic cells.[49] The toxicity of unmodified DLI can be reduced by administering prophylactic immunosuppression around DLI or by decreasing the initial DLI dose for patients with a higher risk of severe GvHD.[50]

In conclusion, we examined how the fundamental processes of BM repopulation and allo-immunological pressure shape the lymphohematopoietic recovery after TCD alloSCT and DLI. The suppression of the patient-derived allo-immunological pressure by the conditioning suffices for sustained engraftment without requiring intensive myeloablation or donor-derived allo-immunological pressure. We show that competitive repopulation can be sufficient to reach complete donor-derived hematopoiesis, but that some allo-immunological pressure is needed for the establishment of a completely donor-derived T-cell compartment, either by the development of GvHD or by administration of DLI. We illustrate that it is possible to separate GvL from GvHD, as conversion to durable complete donor-derived hematopoiesis following DLI did not require the induction of clinically significant GvHD.

ACKNOWLEDGEMENTS

The authors thank E.J.M. Bogers, S.M. Wallet-Malicka and H.A. Baelde for their data management and technical support.

REFERENCES

1. Quesenberry, P. J., Colvin, G. & Abedi, M. Perspective: fundamental and clinical concepts on stem cell homing and engraftment: a journey to niches and beyond. Exp Hematol 33, 9-19 (2005). doi: 10.1016/j.exphem.2004.10.012
2. Ozdemir, Z. N. & Civriz Bozdağ, S. Graft failure after allogeneic hematopoietic stem cell transplantation. Transfus Apher Sci 57, 163-167 (2018). doi: 10.1016/j.transci.2018.04.014
3. Gandy, K. L., Domen, J., Aguila, H. & Weissman, I. L. CD8+TCR+ and CD8+TCR- cells in whole bone marrow facilitate the engraftment of hematopoietic stem cells across allogeneic barriers. Immunity 11, 579-590 (1999). doi: 10.1016/s1074-7613(00)80133-8
4. van den Brink, M. R. & Burakoff, S. J. Cytolytic pathways in haematopoietic stem-cell transplantation. Nature reviews. Immunology 2, 273-281 (2002). doi: 10.1038/nri775
5. Kolb, H. J. Graft-versus-leukemia effects of transplantation and donor lymphocytes. Blood 112, 4371-4383 (2008). doi: 10.1182/blood-2008-03-077974
6. Weisdorf, D., Zhang, M. J., Arora, M. et al. Graft-versus-host disease induced graft-versus-leukemia effect: greater impact on relapse and disease-free survival after reduced intensity conditioning. Biol Blood Marrow Transplant 18, 1727-1733 (2012). doi: 10.1016/j.bbmt.2012.06.014
7. Flowers, M. E., Inamoto, Y., Carpenter, P. A. et al. Comparative analysis of risk factors for acute graft-versus-host disease and for chronic graft-versus-host disease according to National Institutes of Health consensus criteria. Blood 117, 3214-3219 (2011). doi: 10.1182/blood-2010-08-302109
8. Olsson, R., Remberger, M., Schaffer, M. et al. Graft failure in the modern era of allogeneic hematopoietic SCT. Bone Marrow Transplant 48, 537-543 (2013). doi: 10.1038/bmt.2012.239
9. Ferrara, J. L., Levine, J. E., Reddy, P. & Holler, E. Graft-versus-host disease. Lancet 373, 1550-1561 (2009). doi: 10.1016/s0140-6736(09)60237-3
10. Falkenburg, J. H. F. & Jedema, I. Graft versus tumor effects and why people relapse. Hematology. American Society of Hematology. Education Program 2017, 693-698 (2017). doi: 10.1182/asheducation-2017.1.693
11. Kamradt, T. & Mitchison, N. A. Tolerance and autoimmunity. N Engl J Med 344, 655-664 (2001). doi: 10.1056/nejm200103013440907
12. Roncarolo, M. G., Gregori, S., Lucarelli, B., Ciceri, F. & Bacchetta, R. Clinical tolerance in allogeneic hematopoietic stem cell transplantation. Immunological reviews 241, 145-163 (2011). doi: 10.1111/j.1600-065X.2011.01010.x
13. Penack, O., Marchetti, M., Ruutu, T. et al. Prophylaxis and management of graft versus host disease after stem-cell transplantation for haematological malignancies: updated consensus recommendations of the European Society for Blood and Marrow Transplantation. Lancet Haematol 7, e157-e167 (2020). doi: 10.1016/s2352-3026(19)30256-x
14. Marsh, R. A., Lane, A., Mehta, P. A. et al. Alemtuzumab levels impact acute GVHD, mixed chimerism, and lymphocyte recovery following alemtuzumab, fludarabine, and melphalan RIC HCT. Blood 127, 503-512 (2016). doi: 10.1182/blood-2015-07-659672
15. Lindemans, C. A., Chiesa, R., Amrolia, P. J. et al. Impact of thymoglobulin prior to pediatric unrelated umbilical cord blood transplantation on immune reconstitution and clinical outcome. Blood 123, 126-132 (2014). doi: 10.1182/blood-2013-05-502385
16. Bryant, A. R. & Perales, M. A. Advances in Ex Vivo T Cell Depletion - Where Do We Stand? Advances in cell and gene therapy 2 (2019). doi: 10.1002/acg2.29
17. Loeff, F. C., van Egmond, E. H. M., Moes, D. et al. Impact of alemtuzumab pharmacokinetics on T-cell dynamics, graft-versus-host disease and viral reactivation in patients receiving allogeneic stem cell transplantation with an alemtuzumab-based T-cell-depleted graft. Transplant immunology 57, 101209 (2019). doi: 10.1016/j.trim.2019.06.001
18. Williams, L., Cirrone, F., Cole, K. et al. Post-transplantation Cyclophosphamide: From HLA-Haploidentical to Matched-Related and Matched-Unrelated Donor Blood and Marrow Transplantation. Front Immunol 11, 636 (2020). doi: 10.3389/fimmu.2020.00636
19. Saad, A. & Lamb, L. S. Ex vivo T-cell depletion in allogeneic hematopoietic stem cell transplant: past, present and future. Bone Marrow Transplant 52, 1241-1248 (2017). doi: 10.1038/bmt.2017.22
20. Soiffer, R. J., Lerademacher, J., Ho, V. et al. Impact of immune modulation with anti-T-cell

antibodies on the outcome of reduced-intensity allogeneic hematopoietic stem cell transplantation for hematologic malignancies. Blood 117, 6963-6970 (2011). doi: 10.1182/blood-2011-01-332007
21. Soiffer, R. J., Kim, H. T., McGuirk, J. et al. Prospective, Randomized, Double-Blind, Phase III Clinical Trial of Anti-T-Lymphocyte Globulin to Assess Impact on Chronic Graft-Versus-Host Disease-Free Survival in Patients Undergoing HLA-Matched Unrelated Myeloablative Hematopoietic Cell Transplantation. J Clin Oncol 35, 4003-4011 (2017). doi: 10.1200/jco.2017.75.8177
22. Schmid, C., Labopin, M., Schaap, N. et al. Prophylactic donor lymphocyte infusion after allogeneic stem cell transplantation in acute leukaemia - a matched pair analysis by the Acute Leukaemia Working Party of EBMT. Br J Haematol 184, 782-787 (2019). doi: 10.1111/bjh.15691
23. Eefting, M., von dem Borne, P. A., de Wreede, L. C. et al. Intentional donor lymphocyte-induced limited acute graft-versus-host disease is essential for long-term survival of relapsed acute myeloid leukemia after allogeneic stem cell transplantation. Haematologica 99, 751-758 (2014). doi: 10.3324/haematol.2013.089565
24. Caldemeyer, L. E., Akard, L. P., Edwards, J. R. et al. Donor Lymphocyte Infusions Used to Treat Mixed-Chimeric and High-Risk Patient Populations in the Relapsed and Nonrelapsed Settings after Allogeneic Transplantation for Hematologic Malignancies Are Associated with High Five-Year Survival if Persistent Full Donor Chimerism Is Obtained or Maintained. Biol Blood Marrow Transplant 23, 1989-1997 (2017). doi: 10.1016/j.bbmt.2017.07.007
25. Krishnamurthy, P., Potter, V. T., Barber, L. D. et al. Outcome of donor lymphocyte infusion after T cell-depleted allogeneic hematopoietic stem cell transplantation for acute myelogenous leukemia and myelodysplastic syndromes. Biol Blood Marrow Transplant 19, 562-568 (2013). doi: 10.1016/j.bbmt.2012.12.013
26. Eefting, M., Halkes, C. J., de Wreede, L. C. et al. Myeloablative T cell-depleted alloSCT with early sequential prophylactic donor lymphocyte infusion is an efficient and safe post-remission treatment for adult ALL. Bone Marrow Transplant 49, 287-291 (2014). doi: 10.1038/bmt.2013.111
27. Barge, R. M., Starrenburg, C. W., Falkenburg, J. H. et al. Long-term follow-up of myeloablative allogeneic stem cell transplantation using Campath "in the bag" as T-cell depletion: the Leiden experience. Bone Marrow Transplant 37, 1129-1134 (2006). doi: 10.1038/sj.bmt.1705385
28. von dem Borne, P. A., Beaumont, F., Starrenburg, C. W. et al. Outcomes after myeloablative unrelated donor stem cell transplantation using both in vitro and in vivo T-cell depletion with alemtuzumab. Haematologica 91, 1559-1562 (2006).
29. von dem Borne, P. A., Starrenburg, C. W., Halkes, S. J. et al. Reduced-intensity conditioning allogeneic stem cell transplantation with donor T-cell depletion using alemtuzumab added to the graft ('Campath in the bag'). Current opinion in oncology 21 Suppl 1, S27-29 (2009). doi: 10.1097/01.cco.0000357472.76337.0e
30. Falkenburg, J., Schmid, C., Kolb, H., Locatelli, F. & Kuball, J. in The EBMT Handbook. Hematopoietic Stem Cell Transplantation and Cellular Therapies. (eds E. Carreras, C. Dufour, M. Mohty, & N. Kroger) 443-448 (Springer, 2019).
31. Rijneveld, A. W., van der Holt, B., de Weerdt, O. et al. Clofarabine added to intensive treatment in adult patients with newly diagnosed ALL: the HOVON-100 trial. Blood advances 6, 1115-1125 (2022). doi: 10.1182/bloodadvances.2021005624
32. Cruijsen, M., Hilberink, J. R., van der Velden, W. et al. Low relapse risk in poor risk AML after conditioning with 10-day decitabine, fludarabine and 2 Gray TBI prior to allogeneic hematopoietic cell transplantation. Bone Marrow Transplant 56, 1964-1970 (2021). doi: 10.1038/s41409-021-01272-3
33. Przepiorka, D., Weisdorf, D., Martin, P. et al. 1994 Consensus Conference on Acute GVHD Grading. Bone Marrow Transplant 15, 825-828 (1995).
34. Shulman, H. M., Sullivan, K. M., Weiden, P. L. et al. Chronic graft-versus-host syndrome in man. A long-term clinicopathologic study of 20 Seattle patients. The American journal of medicine 69, 204-217 (1980). doi: 10.1016/0002-9343(80)90380-0
35. Schemper, M. & Smith, T. L. A note on quantifying follow-up in studies of failure time. Control Clin Trials 17, 343-346 (1996). doi: 10.1016/0197-2456(96)00075-x
36. Konuma, T., Kato, S., Oiwa-Monna, M. et al. Early phase mixed chimerism in bone marrow does not affect long-term outcomes of myeloablative single-unit cord blood transplantation for adult

patients with hematological malignancies. Leukemia & lymphoma 57, 2848-2854 (2016). doi: 10.3109/10428194.2016.1171860
37. Takahashi, S., Ooi, J., Tomonari, A. et al. Comparative single-institute analysis of cord blood transplantation from unrelated donors with bone marrow or peripheral blood stem-cell transplants from related donors in adult patients with hematologic malignancies after myeloablative conditioning regimen. Blood 109, 1322-1330 (2007). doi: 10.1182/blood-2006-04-020172
38. Koreth, J., Kim, H. T., Nikiforow, S. et al. Donor chimerism early after reduced-intensity conditioning hematopoietic stem cell transplantation predicts relapse and survival. Biol Blood Marrow Transplant 20, 1516-1521 (2014). doi: 10.1016/j.bbmt.2014.05.025
39. Huisman, C., de Weger, R. A., de Vries, L., Tilanus, M. G. & Verdonck, L. F. Chimerism analysis within 6 months of allogeneic stem cell transplantation predicts relapse in acute myeloid leukemia. Bone Marrow Transplant 39, 285-291 (2007). doi: 10.1038/sj.bmt.1705582
40. Tang, X., Alatrash, G., Ning, J. et al. Increasing chimerism after allogeneic stem cell transplantation is associated with longer survival time. Biol Blood Marrow Transplant 20, 1139-1144 (2014). doi: 10.1016/j.bbmt.2014.04.003
41. Sheth, V., Potter, V., de Lavallade, H. et al. Mixed T cell lineage chimerism in acute leukemia/MDS using pre-emptive donor lymphocyte infusion strategy-Is it prognostic?-a single-center retrospective study. Blood cancer journal 11, 128 (2021). doi: 10.1038/s41408-021-00519-y
42. Storek, J., Geddes, M., Khan, F. et al. Reconstitution of the immune system after hematopoietic stem cell transplantation in humans. Semin Immunopathol 30, 425-437 (2008). doi: 10.1007/s00281-008-0132-5
43. Anandi, P., Tian, X., Ito, S. et al. Ex vivo T-cell-depleted allogeneic stem cell transplantation for hematologic malignancies: The search for an optimum transplant T-cell dose and T-cell add-back strategy. Cytotherapy 19, 735-743 (2017). doi: 10.1016/j.jcyt.2017.03.010
44. Mohty, M., Avinens, O., Faucher, C. et al. Predictive factors and impact of full donor T-cell chimerism after reduced intensity conditioning allogeneic stem cell transplantation. Haematologica 92, 1004-1006 (2007). doi: 10.3324/haematol.10971
45. Carnevale-Schianca, F., Caravelli, D., Gallo, S. et al. Post-Transplant Cyclophosphamide and Tacrolimus-Mycophenolate Mofetil Combination Governs GVHD and Immunosuppression Need, Reducing Late Toxicities in Allogeneic Peripheral Blood Hematopoietic Cell Transplantation from HLA-Matched Donors. Journal of clinical medicine 10 (2021). doi: 10.3390/jcm10061173
46. Fernández-Avilés, F., Urbano-Ispizua, A., Aymerich, M. et al. Serial quantification of lymphoid and myeloid mixed chimerism using multiplex PCR amplification of short tandem repeat-markers predicts graft rejection and relapse, respectively, after allogeneic transplantation of CD34+ selected cells from peripheral blood. Leukemia 17, 613-620 (2003). doi: 10.1038/sj.leu.2402854
47. Rujkijyanont, P., Morris, C., Kang, G. et al. Risk-adapted donor lymphocyte infusion based on chimerism and donor source in pediatric leukemia. Blood cancer journal 3, e137 (2013). doi: 10.1038/bcj.2013.39
48. Feliu, J., Potter, V., Grimaldi, F. et al. Full donor chimerism without graft-versus-host disease: the key factor for maximum benefit of pre-emptive donor lymphocyte infusions (pDLI). Bone Marrow Transplant (2019). doi: 10.1038/s41409-019-0695-x
49. Schmid, C., Kuball, J. & Bug, G. Defining the Role of Donor Lymphocyte Infusion in High-Risk Hematologic Malignancies. J Clin Oncol 39, 397-418 (2021). doi: 10.1200/jco.20.01719
50. Greiner, J., Gotz, M., Bunjes, D., Hofmann, S. & Wais, V. Immunological and Clinical Impact of Manipulated and Unmanipulated DLI after Allogeneic Stem Cell Transplantation of AML Patients. Journal of clinical medicine 9 (2019). doi: 10.3390/jcm9010039

SUPPLEMENTAL METHODS

Methodology for measuring chimerism and circulating immune cell counts

BM chimerism was determined on unfractionated BM samples by short-tandem-repeat (STR) PCR. For some patients, transplanted before 2007 with a sex-mismatched donor, unfractionated BM chimerism was determined by FISH analysis using Vysis CEP X/Y probes. The lower detection limit of the chimerism analyses was 1-2%, depending on the method and the selected markers. In a subset of patients, chimerism was also determined in granulocytes, monocytes, CD4+ T cells, CD8+ T cells, NK cells and B cells on peripheral blood by STR PCR. For this analysis, 50,000 cells per population were sorted by flow cytometry (Supplemental Table 1). Absolute numbers of circulating CD4+ T cells, CD8+ T cells, B cells and NK cells were measured routinely on anticoagulated fresh venous blood by flow cytometry with bead calibration (Trucount tubes, Becton Dickinson, Breda, The Netherlands) at a detection limit of $0.5x10^6$ cells/l.

Quantification of donor-derived alloimmune responses after DLI in patients with mixed chimerism

To evaluate whether DLIs can induce BM chimerism conversion from MC to FDC or improve the level of donor chimerism, we developed an algorithm to assess the best chimerism response after the first unmodified DLI that patients received while having MC in the BM. We defined this DLI as the 'starting DLI' for this analysis. Patients could enter the analysis only once: the analysis was not restarted if a patient received another DLI after recurrence of MC. The algorithm considered all BM chimerism measurements from 1 week until 25 months after the starting DLI, including measurements after successive DLIs and excluding measurements taken after relapse or administration of interferon, chemotherapy, or second alloSCT. Patients receiving DLI during continued GVHD prophylaxis or after a relapse, administration of chemotherapy, second alloSCT, or modified T cell product were excluded from this analysis. A complete donor-derived alloimmune response was defined as conversion to FDC. Partial donor-derived alloimmune response was defined as a relative decrease in patient chimerism of 50% or an absolute decrease of 20% when starting patient chimerism was at least 50%, 10% when starting patient chimerism was between 20% and 50%, or 5% when patient chimerism was <20%. These values were chosen to prevent that minor fluctuations in patient chimerism were defined as a response. Patients with evaluable BM chimerism measurements after DLI who failed to show a complete or partial response were considered to have no donor-derived alloimmune response. For partial and non-responders a distinction was made between patients who had completed the required follow up period and patients who had died, relapsed, or received chemotherapy, interferon, or second alloSCT within this period. For the evaluation of the durability of the chimerism responses we also considered loss of response, defined as the recurrence of patient chimerism after a complete response or an increase in patient chimerism (using the same cut-offs as described above) after a partial response. The number of DLIs before achieving the best response (until conversion to FDC for complete responders or until the start of decreasing patient chimerism for partial responders) or during the total evaluable follow-up (for non-responders) were recorded.

To evaluate whether clinically significant GVHD is required for chimerism conversion, we examined how many of the complete responders started tIS for GVHD within 25 months after the starting DLI.

Competing risks models

The cumulative incidence of neutrophil recovery as proxy for primary engraftment was calculated in a competing risks model starting at alloSCT and with non-engraftment and death as competing events. The cumulative incidence of clinically significant GVHD after TCD alloSCT was calculated with cellular intervention ((un)modified DLI, stem cell boost or second alloSCT), relapse, start of chemotherapy or interferon, and death as competing events. The cumulative incidences of clinically significant GVHD and of the development of acute GVHD grade II-IV or extensive chronic GVHD were calculated in separate competing risks models with relapse and death as competing events.

SUPPLEMENTAL RESULTS

Clinical outcomes of the total strategy

The 5-year overall and relapse-free survival were 49% (95%-CI 43-55) and 46% (95%-CI 40-52), respectively. The cumulative incidence of relapse was 24% (95%-CI 19-29) at this time, while the non-relapse mortality was 30% (95%-CI 24-35). The outcomes per conditioning and donor type are shown in Supplemental Figure 1. The 1-year cumulative incidence of clinically significant GVHD was 37% (95%-CI 32-43; see Supplemental Figure 2 for comparison with the overall grading of GVHD).

Development of GVHD after DLI and DLI-induced cytopenia

In total, 131 patients received an unmodified DLI after alloSCT without any prior relapse, chemotherapy or other cellular intervention or ongoing prophylactic immunosuppression. Of these patients, 65 had mixed BM chimerism at time of DLI, 59 FDC and for 7 the level of BM chimerism was unknown. 24 (37%) of the 65 patients with MC at time of DLI developed clinically significant GVHD compared to 9 (15%) of the 59 FDC patients. Of these 33 patients with GVHD after DLI, 14 died during GVHD (2 had FDC at time of DLI), while only one relapsed. Three patients, all with mixed BM chimerism (80-98% donor), showed DLI-induced cytopenia, all just before or at time of the start of GVHD.

SUPPLEMENTAL TABLES

Marker	Panel 1 Tube 1	Tube 2	Tube 3 (only used for sorting)	Panel 2 Tube 1
CD3	FITC	APC	-	APC
CD4	-	FITC	-	PB
CD8	-	PE	-	FITC
CD14	-	-	PE	APC-H7
CD16	PE	-	-	PE
CD19	APC	-	-	PE-Cy7
CD45	PerCP	PerCP	PerCP	PerCP
CD56	PE	-	-	PE

Supplemental Table 1. Fluorescence panels used for sorting and counting of the immune cells. APC, allophycocyanin; FITC, fluorescein isothiocyanate; PB, PacificBlue; PE, phycoerythrin; PerCP, peridinin-chlorophyll protein. All fluorochromes were from BD, Becton Dickinson, Breda, The Netherlands. The CD45intSSChiCD14- gate was used to identify and sort granulocytes.

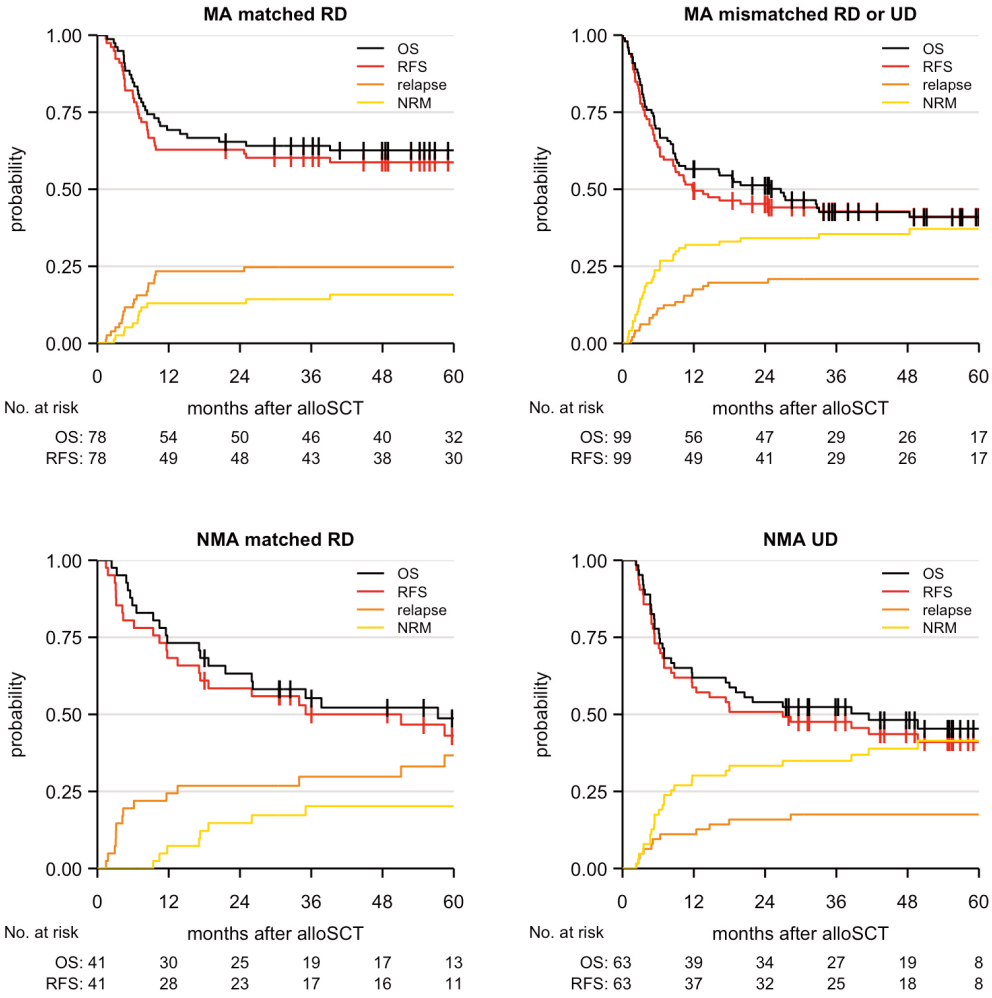
	Granulocytes	Monocytes	NK cells	B cells	CD4+ T cells	CD8+ T cells
Lower limit of reference range (cells/l)	1500	100	40	60	560	260
Time after alloSCT						
6 weeks	76%	90%	90%	24%	3%	23%
3 months	75%	95%	97%	61%	12%	32%
6 months	79%	97%	96%	79%	8%	43%

Supplemental Table 2. Recovery of immune cell subset counts after TCD alloSCT. Percentages of patients having immune cell counts of at least the lower limit of the reference range. Measurements after relapse or administration of chemotherapy, interferon, modified T cell product or second alloSCT were excluded.

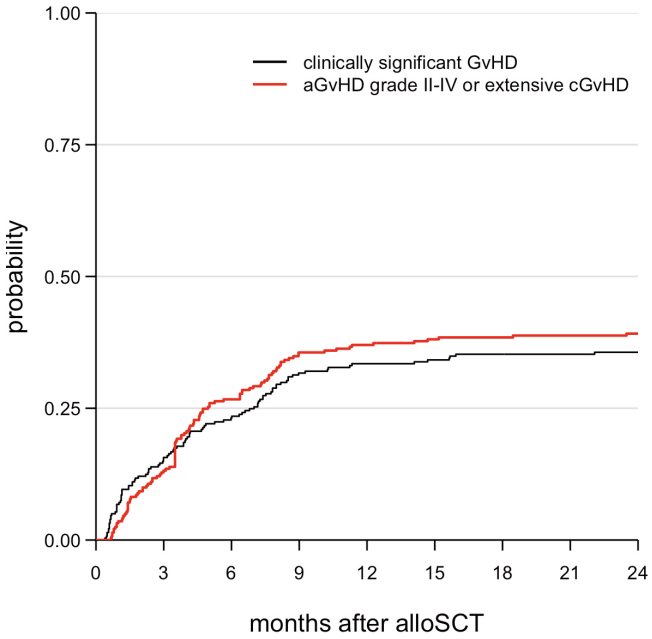
	Evaluable period from starting DLI			
Best response	**2 months**	**4 months**	**7 months**	**25 months**
Unevaluable due to early death from GvHD	4 (6%)	4 (6%)	4 (6%)	4 (6%)
Unevaluable due to early relapse	3 (5%)	3 (5%)	3 (5%)	3 (5%)
No response				
No event within period	34 (52%)	18 (28%)	9 (14%)	1 (2%)
Relapse, chemotherapy or interferon within period	0	2 (3%)	2 (3%)	3 (5%)
Death within period	0	0	1 (2%)	2 (3%)
Partial response				
No event within period	8 (12%)	8 (12%)	6 (9%)	1 (2%)
Relapse, chemotherapy or interferon within period	0	0	0	1 (2%)
Death within period	1 (2%)	1 (2%)	3 (5%)	3 (5%)
Complete response	15 (23%)	29 (45%)	37 (57%)	47 (72%)

Supplemental Table 3. Best BM chimerism response after DLI. Details regarding the best BM response and events that terminated the evaluation period.

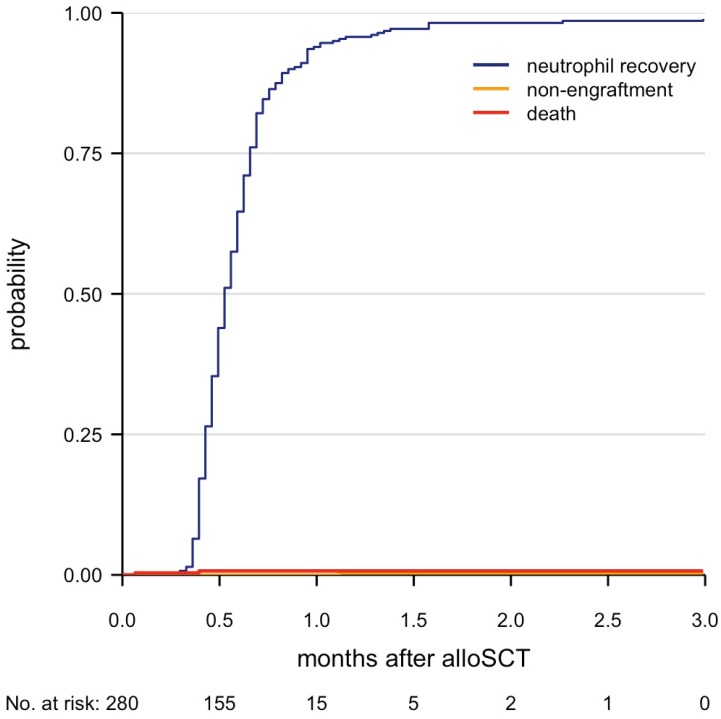
SUPPLEMENTAL FIGURES



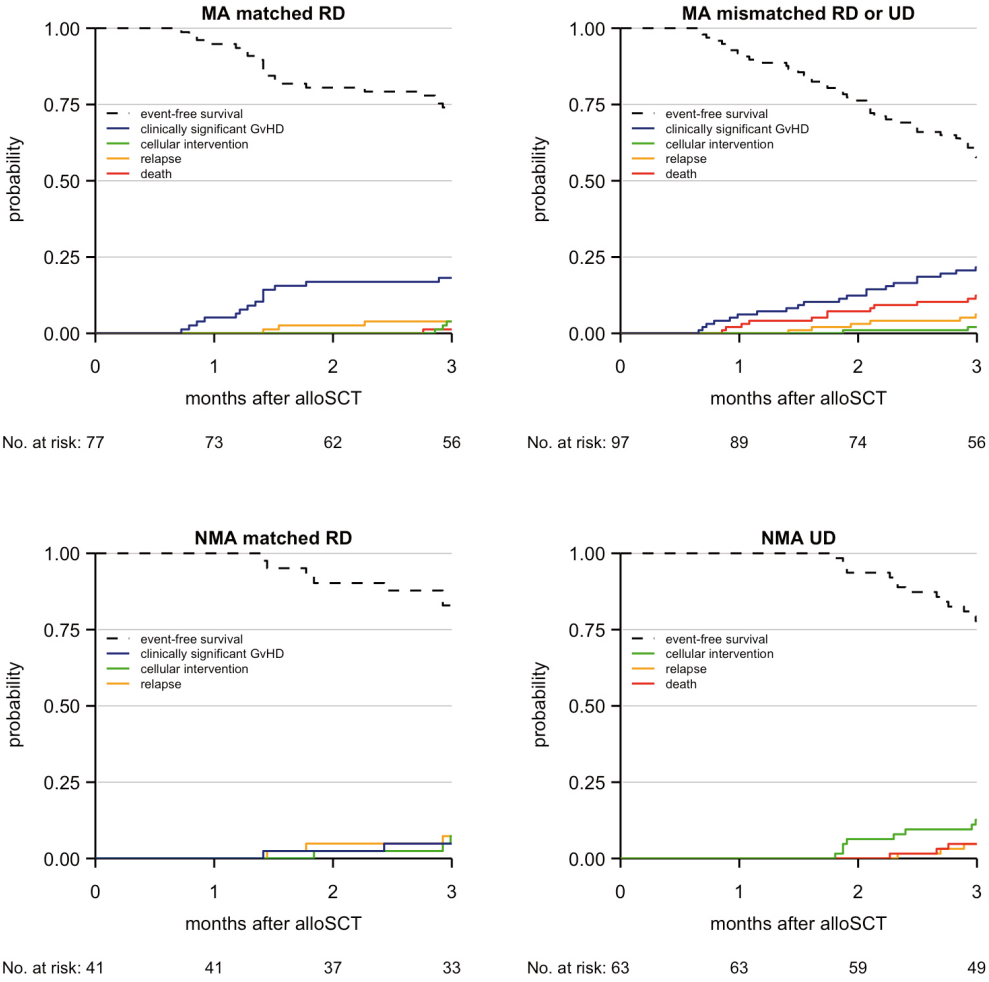
Supplemental Figure 1. Overall and relapse-free survival, relapse and non-relapse mortality. Kaplan-Meier curves for overall and relapse-free survival and cumulative incidence curves for relapse and non-relapse mortality per cohort. MA, myeloablative conditioned; NMA, nonmyeloablative conditioned; RD, related donor; UD, unrelated donor; OS, overall survival; RFS, relapse-free survival; NRM, non-relapse mortality. | indicates censoring times.



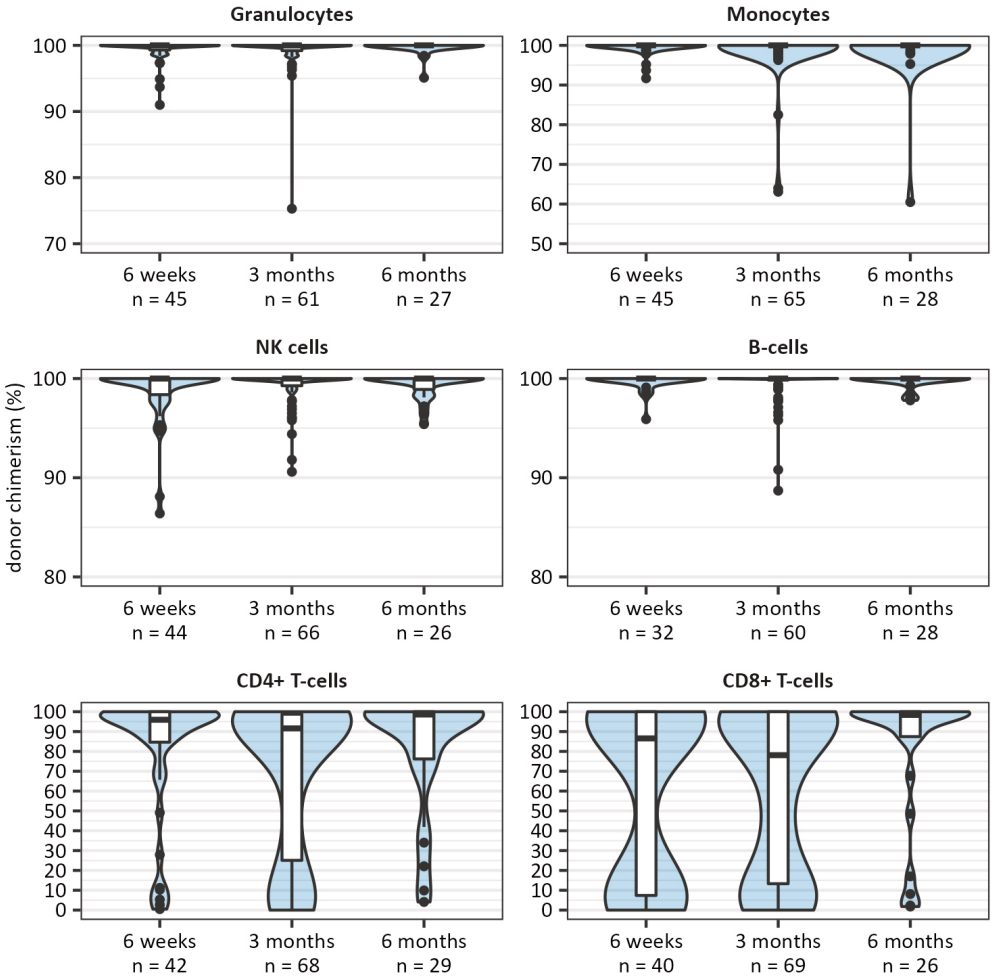
Supplemental Figure 2. Cumulative incidence of clinically significant GVHD. Cumulative incidence curves of clinically significant GVHD (GVHD requiring systemic treatment) and a combined curve of acute GVHD grade II-IV and extensive chronic GVHD. The cumulative incidences were calculated in separate competing risks models with relapse and death as competing events. The difference between the lines is caused by patients receiving tIS for lower grade GVHD not responding to topical treatment, patients with higher grade GVHD responding rapidly to topical treatment or requiring less than 14 days tIS, and patients with tIS for GVHD not proven by histology.



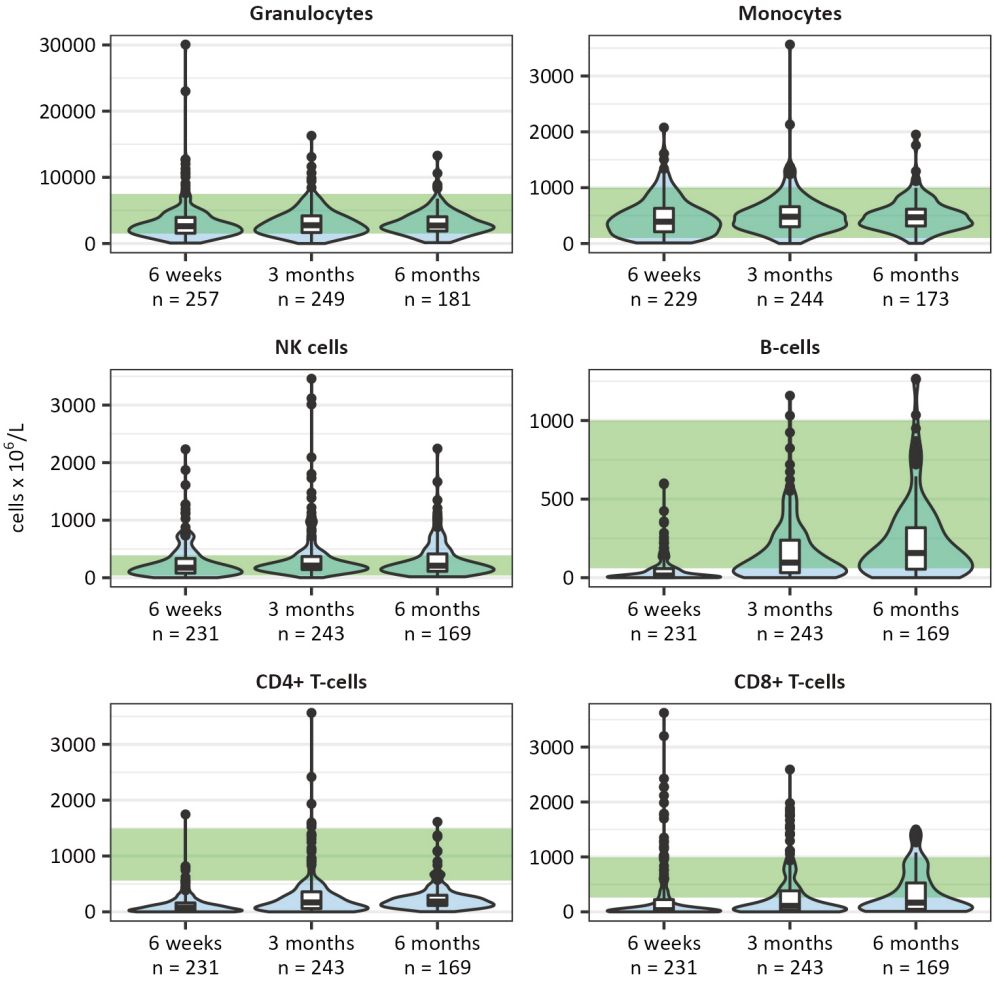
Supplemental Figure 3. Primary engraftment. Cumulative incidence curves of the competing events neutrophil recovery as a proxy for engraftment, non-engraftment and death. One patient never had neutrophils below 0.5×10^9/l and was excluded from this analysis.



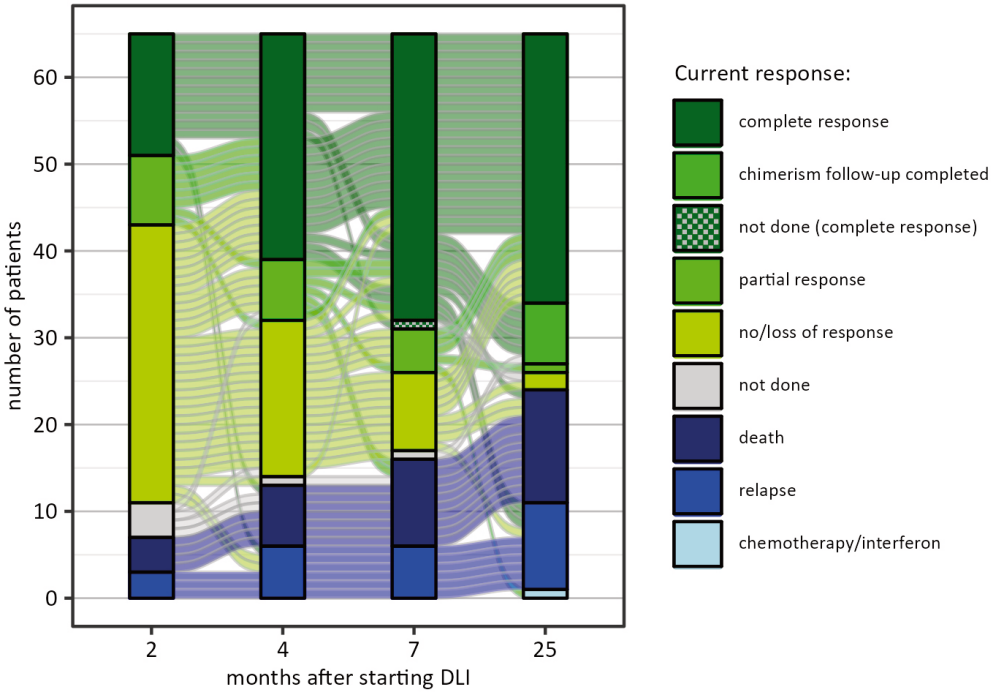
Supplemental Figure 4. Clinically significant GVHD after TCD alloSCT without DLI. Cumulative incidence curves of the competing events clinically significant GVHD, cellular intervention, relapse, start of chemotherapy or interferon, and death. Per plot only the curves corresponding to events observed in the subgroup are shown. The event-free survival was defined as the time from alloSCT until the occurrence of one of these events. MA, myeloablative; NMA, nonmyeloablative; RD, related donor; UD, unrelated donor; GVHD, graft-versus-host disease.



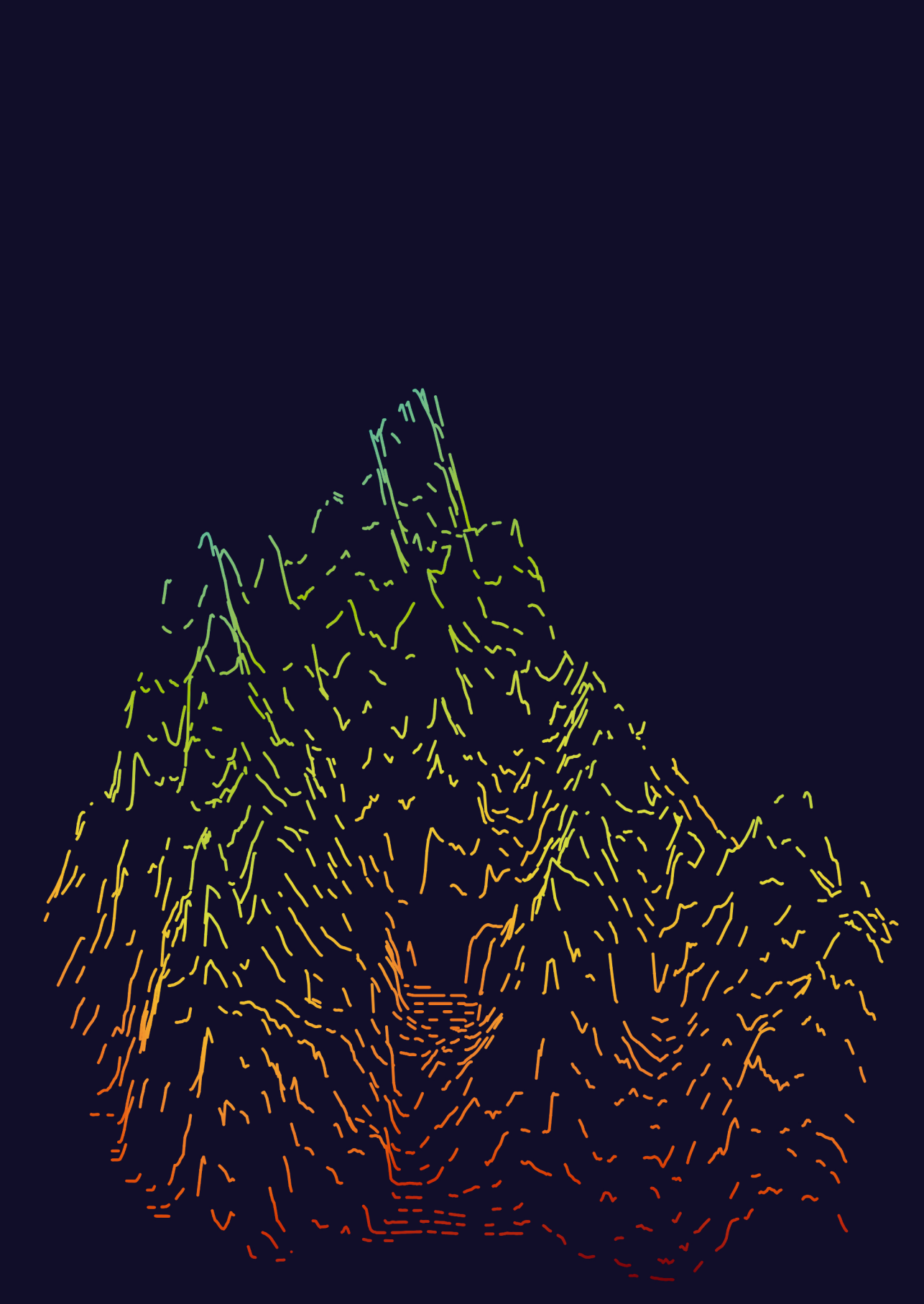
Supplemental Figure 5. Donor chimerism in immune cell subsets. Donor chimerism in immune cell subsets at 6 weeks, 3 months and 6 months after TCD alloSCT without any prior cellular intervention, relapse, chemotherapy or interferon. The lymphocyte counts of the total cohort, also including measurements after unmodified DLI, are shown in Supplemental Figure 6 and summarized in Supplemental Table 2.



Supplemental Figure 6. Recovery of absolute numbers of immune cell subsets after TCD alloSCT. Measurements after relapse or administration of chemotherapy, interferon, modified T cell product or second alloSCT were excluded. The green areas represent the reference ranges used in our laboratory.



Supplemental Figure 7. Current BM chimerism responses after unmodified DLI. Current BM chimerism responses in the 65 patients with mixed hematopoiesis receiving unmodified DLI. This plot provides insight in the durability of the achieved responses as patients with a response may lose their response or may for instance relapse after chimerism conversion. Per patient only the first terminating event (relapse, death, chemotherapy, interferon or second alloSCT) is considered: patients cannot move on to 'death' after a relapse. All patients in the green and grey areas were alive without any relapse, chemotherapy or interferon at the corresponding timepoint after their starting DLI. Per protocol, BM biopsies are stopped at 2 years after alloSCT if the patient has complete donor-derived hematopoiesis. These patients move to 'chimerism follow-up completed' in the plot. All other patients without any BM chimerism measurement within 2 months before the respective timepoint move to the 'not done' areas. Of these patients, those who showed a complete response in their previous measurement are shaded with green.



3

Joint models quantify associations between immune cell kinetics and allo-immunological events after allogeneic stem cell transplantation and subsequent donor lymphocyte infusion

Eva A.S. Koster*, Eduoard F. Bonneville*, Peter A. von dem Borne, Peter van Balen, Erik W.A. Marijt, Jennifer M.L. Tjon, Tjeerd J.F. Snijders, Daniëlle van Lammeren, Hendrik Veelken, Hein Putter, J.H. Frederik Falkenburg, Constantijn J.M. Halkes, Liesbeth C. de Wreede

* shared first authorship

Front Immunol. 2023 Aug 1;14:1208814.

ABSTRACT

Alloreactive donor-derived T cells play a pivotal role in alloimmune responses after allogeneic hematopoietic stem cell transplantation (alloSCT); both in the relapse-preventing Graft-versus-Leukemia (GvL) effect and the potentially lethal complication Graft-versus-Host-Disease (GvHD). The balance between GvL and GvHD can be shifted by removing T cells via T-cell depletion (TCD) to reduce the risk of GvHD, and by introducing additional donor T cells (donor lymphocyte infusions [DLI]) to boost the GvL effect. However, the association between T-cell kinetics and the occurrence of allo-immunological events has not been clearly demonstrated yet. Therefore, we investigated the complex associations between the T-cell kinetics and alloimmune responses in a cohort of 166 acute leukemia patients receiving alemtuzumab-based TCD alloSCT. Of these patients, 62 with an anticipated high risk of relapse were scheduled to receive a prophylactic DLI at 3 months after transplant. In this setting, we applied joint modelling which allowed us to better capture the complex interplay between DLI, T-cell kinetics, GvHD and relapse than traditional statistical methods. We demonstrate that DLI can induce detectable T-cell expansion, leading to an increase in total, CD4+ and CD8+ T-cell counts starting at 3 months after alloSCT. CD4+ T cells showed the strongest association with the development of alloimmune responses: higher CD4 counts increased the risk of GvHD (hazard ratio 2.44, 95% confidence interval 1.45-4.12) and decreased the risk of relapse (hazard ratio 0.65, 95% confidence interval 0.45-0.92). Similar models showed that natural killer cells recovered rapidly after alloSCT and were associated with a lower risk of relapse (HR 0.62, 95%-CI 0.41-0.93). The results of this study advocate the use of joint models to further study immune cell kinetics in different settings.

INTRODUCTION

The curative potential of allogeneic stem cell transplantation (alloSCT) in the treatment of hematological malignancies depends on the introduction of donor-derived alloreactive T cells.[1] These T cells recognize nonself antigens on patient-derived cells and can, once activated, expand and eliminate those cells. Targeting antigens on lymphohematopoietic cells including the malignant cells leads to the desired Graft-versus-Leukemia (GvL) effect and prevents relapse. However, when other tissues of the patient are targeted, Graft-versus-Host-Disease (GvHD) may develop.[2] Natural killer (NK) cells may discriminate between healthy and non-healthy (e.g., virus-infected or malignant) cells by acting on signals from inhibitory and activating receptors that bind to the target cell. In the setting of alloSCT, early NK cell recovery can protect against relapse and viral infections.[3,4] However, NK cells do not appear to be important effector cells in GvHD.[5]

To reduce the risk of severe GvHD, donor T-cell depletion (TCD) can be applied, although this will decrease the GvL effect.[6] In order to restore the GvL effect to prevent relapse, TCD alloSCT can be combined with the administration of donor lymphocyte infusions (DLIs) after transplant.[2,7,8] DLI as part of a preemptive strategy is administered to patients with detectable minimal residual disease (MRD) or with residual patient hematopoiesis: mixed chimerism (MC). DLI as part of a prophylactic strategy is given to all patients in whom no GvHD has developed as sign of alloreactivity. The alloreactive potential of DLI decreases over time after alloSCT: both the efficacy (GvL effect) and toxicity (GvHD) are highest early after alloSCT.[9,10] Therefore, administration preferably starts a few months after alloSCT to allow for sufficient GvL without severe GvHD.[11]

Since T cells are pivotal for alloimmune responses, several groups have investigated T-cell kinetics after alloSCT and their impact on the development of GvHD or relapse. However, as shown in the recent review by Yanir et al.[12], the reported results are inconsistent, and their interpretation is complicated by several factors. First, T cells can be patient- or donor-derived, while only donor-derived T cells are responsible for GvHD and GvL. Second, the T-cell changes following alloSCT are the combined result of *de novo* T-cell generation from infused hematopoietic stem cells starting at least 6 months after alloSCT, homeostatic proliferation of T cells present in the patient or graft, T-cell expansion during infections and expansion of alloreactive T cells responsible for GvL and GvHD. Especially cytomegalovirus (CMV) reactivations are common during the first 3 months after alloSCT and strongly affect the kinetics of both T cells and NK cells after alloSCT.[13-15] This may distort the association between the kinetics of the main T-cell subsets and specific alloimmune responses, i.e., the presence of GvHD and the absence of relapse as a result of the GvL effect. Third, factors that could influence both the T-cell kinetics and the risks of GvHD and relapse, such as the conditioning regimen, donor type and the use and method of TCD, should be properly accounted for. Finally, ignoring clinical events or interventions during follow-up can also be problematic: over time, the patients that have not yet experienced an event like relapse, death or the development of GvHD, become less representative of the population at the beginning of follow-up. As death by definition prevents further T-cell measurements and the possibility of experiencing subsequent GvHD and relapse, bias is created by considering the patients who died as having non-informatively dropped out (i.e. that their measurements *could*

have been measured if kept under follow-up). Likewise, DLI and the use of posttransplant prophylactic immunosuppression are known to affect the risks of relapse and GvHD, but may also affect the T-cell kinetics.[16-23] To fully understand the complex interplay between all these factors, sophisticated statistical methods are required that properly model the T-cell kinetics themselves, along with their association with GvHD or relapse. Joint modelling captures the T-cell trajectories and the clinical events simultaneously, accounting for informative dropout, as well as the measurement error and heterogeneity in individual trajectories.[24]

In this study, we performed joint modelling to investigate the complex associations between the immune cell kinetics and alloreactivity in a cohort of 166 patients receiving an alloSCT for acute leukemia or myelodysplastic syndrome (MDS). All patients received an alemtuzumab-based TCD alloSCT after nonmyeloablative conditioning without any posttransplant prophylactic immunosuppression. Patients with an anticipated high risk of relapse were scheduled to receive an early low-dose DLI prophylactically at 3 months after alloSCT, while prophylactic DLI administration for the other patients started at 6 months. In this unique setting we investigated the impact of the early low-dose DLI on the T-cell and NK cell kinetics during the first 6 months after transplant and the association between these kinetics and the development of clinical events.

METHODS

Study population

This retrospective study included all adult patients with acute myeloid leukemia, acute lymphoblastic leukemia or MDS in complete morphologic remission after intensive induction therapy who received their first alloSCT from a 9 or 10 out of 10 HLA-matched donor using nonmyeloablative conditioning and alemtuzumab-based TCD[25] between March 2008 and December 2019 at Leiden University Medical Center (LUMC, Leiden, The Netherlands). Two patients who were transplanted while receiving systemic immunosuppression for a non-transplant indication (polymyalgia rheumatica and cryptogenic organizing pneumonia) were excluded because of the potential impact of the ongoing systemic immunosuppression on the immune cell recovery after alloSCT. All patients signed informed consent for data collection and analysis. Data were analyzed as of July 2021.

Transplantation and DLI strategy

As conditioning regimen patients received either fludarabine (6 days 50 mg/m^2 orally or 30 mg/m^2 intravenously) and busulfan (2 days 4x0.8 mg/kg intravenously), or the FLAMSA regimen: fludarabine (5 days 30 mg/m^2 intravenously), cytarabine (4 days 2000 mg/m^2 intravenously), amsacrine (4 days 100 mg/m^2 intravenously) and busulfan (4 days 4x0.8 mg/kg intravenously). In both regimens, TCD was performed by adding 20 mg alemtuzumab (Sanofi Genzyme, Naarden, The Netherlands) to the graft before infusion and by administering 15 mg alemtuzumab intravenously on days -4 and -3. Patients with an unrelated donor (UD) received rabbit-derived anti-thymocyte globulin (ATG; Sanofi Genzyme) additionally on day -2 (until April 2010 2mg/kg and thereafter 1mg/kg). None of the patients received posttransplant GvHD prophylaxis.

The dose of unmodified preemptive and prophylactic DLIs was based on donor type and timing after alloSCT. Standard DLIs given at 6 months after alloSCT contained $3x10^6$ or $1.5x10^6$ T cells/kg for patients with a related donor (RD) or an UD, respectively. Early low-dose DLIs given at 3 months after alloSCT contained $0.3x10^6$ or $0.15x10^6$ T cells/kg for patients with a RD or an UD, respectively. Since May 2010, all patients without any relapse and without GvHD requiring systemic immunosuppressive treatment at 6 months after alloSCT prophylactically (i.e., irrespective of chimerism or posttransplant MRD status) were planned to receive the standard DLI. Patients who were considered to have a high risk of relapse based on the disease characteristics or MRD status at time of alloSCT or who received the FLAMSA regimen were also scheduled to receive the early low-dose DLI prophylactically at 3 months after alloSCT. All patients, including those transplanted before May 2010, could receive preemptive DLIs in case of MC or MRD positivity, starting from 3 months after alloSCT. Additionally, as part of several clinical trials, patients could receive modified T-cell products prophylactically or virus-specific T-cell infusions to treat severe viral infections.

Monitoring of CMV and absolute numbers of circulating immune cells

CMV serostatus was assessed in all patients and donors before alloSCT. After transplant CMV was monitored routinely by PCR on peripheral blood samples in all patients. Absolute numbers of circulating total (CD3+), CD4+CD8- and CD4-CD8+ T cells, B cells and NK cells were measured routinely at predefined timepoints on anticoagulated fresh venous blood by flow cytometry with bead calibration (Trucount tubes, BD Biosciences). Samples were measured either on a FACSCalibur using anti-CD3-APC, anti-CD4-FITC, anti-CD8-PE, and anti-CD45-PerCP or with anti-CD3-FITC, anti-CD16-PE, anti-CD19-APC, anti-CD45-PerCP, and anti-CD56-PE, or on a FACSCanto using anti-CD3-APC, anti-CD4-PB, anti-CD8-FITC, anti-CD16-PE, anti-CD19-PE Cy7, anti-CD45-PerCP, and anti-CD56-PE (all from BD). The lower detection limit was $0.5x10^6$ cells/l.

Definitions of events

Relapse was defined as the recurrence of at least 5% blasts on cytomorphologic bone marrow examination or at least 1% blasts in peripheral blood (if possible, confirmed by BM biopsy). We defined clinically significant GvHD as the start of therapeutic systemic immunosuppression for GvHD.[26] We defined 'other failure' as the occurrence of an adverse event with a potential impact on the immune cell kinetics: death, graft failure, start of systemic immunosuppression for a non-GvHD indication, and virus-specific T-cell infusion for a severe viral infection (whichever occurred first). Graft failure was defined as the occurrence of >95% patient BM chimerism in all lineages tested or refractory granulopenia (granulocyte count $<0.5x10^9$/l) in the absence of relapse or ongoing myelotoxic medication.

For this study we analyzed the T-cell and NK cell kinetics and events during the first 6 months after alloSCT, during which the early immunological recovery and most CMV reactivations take place. Furthermore, during this period the impact of the early low-dose DLI can be assessed, as the standard DLI is given to all eligible patients around 6 months after alloSCT. As part of the analyses assessing the net impact of the early low-dose DLI on the T-cell and NK cell kinetics and clinical events, patients receiving a

standard DLI or modified T-cell product as part of a clinical trial were censored at 7 days after this infusion. We considered this to be non-informative censoring, since these interventions were prophylactic and not driven by the clinical course of the patient. For the T-cell kinetics we considered the circulating cell counts of the total (CD3+) T-cell population and the two major T-cell subpopulations: the CD4+CD8- and the CD4-CD8+ T cells.

Statistical analyses

Probabilities of overall survival (OS) and relapse-free survival (RFS) after alloSCT with associated 95% confidence intervals (95%-CI) were calculated by the Kaplan-Meier method. The cumulative incidences of clinically significant GvHD and relapse from time of alloSCT were estimated by means of the Aalen-Johansen method, treating other failure (as described in the previous section) as a third competing risk.

To study the complex interplay between the immune cell kinetics, DLI and clinically relevant endpoints (GvHD and relapse), two joint models were developed; model I starting at time of alloSCT and model II at time of the early low-dose DLI.

Shared-parameter joint models consist of two components: a longitudinal submodel, and a time-to-event submodel.[24] The former often takes the form of a mixed-effects regression model, and the latter is generally assumed to follow a proportional hazards structure, similar to a Cox model (for one or possibly multiple endpoints such as GvHD or relapse). The mixed-effects model allows to model cell count trajectories over time, while appropriately accounting for both the heterogeneity in subject-specific trajectories (using random effects) and measurement error. These two submodels are linked together via an association structure. Practically speaking, this allows the hazard of a particular event to depend on characteristics of an individual's specific trajectory, such as the 'true' underlying (i.e. in absence of measurement error) value over time. In turn, this enables the estimation of an association between a longitudinal marker (e.g. CD3 counts) and the risk of a clinical event (e.g. GvHD). In the presence of an association, the estimated trajectories themselves will be corrected for bias related to the measurements being terminated by the occurrence of endpoints (generally known as 'informative dropout').

Below follows a concise description of the joint models developed for the present application. Detailed explanation of the statistical models and the underlying rationale can be found in the Statistical Supplement. For all models, absolute cell counts were analyzed on the log scale after setting measurements under the detection limit to 0.5. This only occurred at earliest timepoints where because of the lymphodepletion by the conditioning regimen and TCD, the counts are expected to be around zero.

Model I (starting from alloSCT)

To investigate the effect of early low-dose DLI on the kinetics of the T-cell and NK cell counts after TCD alloSCT, we performed an intention-to-treat (ITT) analysis with a baseline group distinguishing between those scheduled for early low-dose DLI because of a high anticipated risk of relapse (henceforth 'high risk' group) and those who were not ('non-high risk' group). We chose this approach instead of a per-protocol analysis since we could not properly define a control group of patients who did not receive early

DLI but could have been candidates as we did not know for each patient who was not scheduled for early DLI whether he/she would have been able to receive it.

Figure 1A shows a schematic overview of joint model I. The model was run separately for each T-cell subset, respectively using CD3, CD4 or CD8 counts, and the total NK

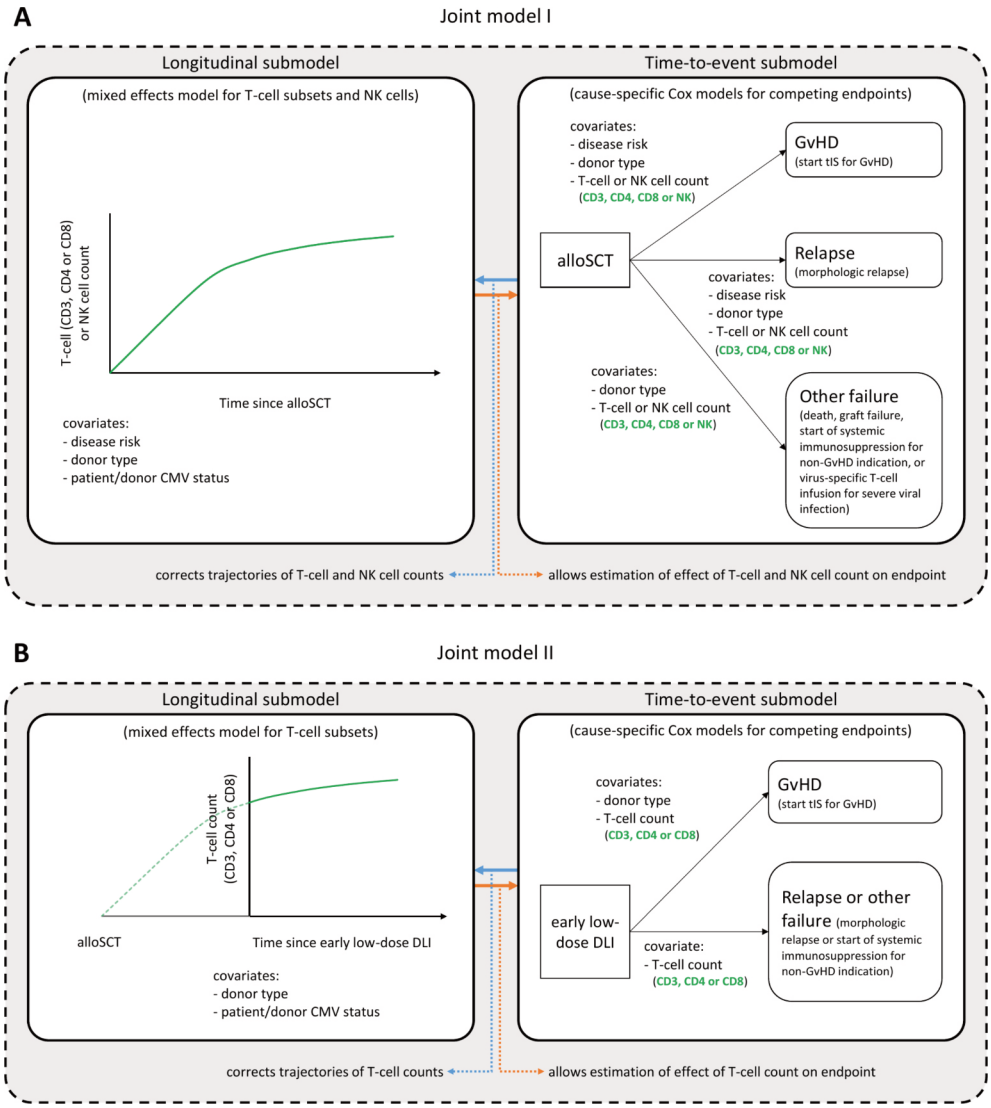


Figure 1. Structure of the joint models. Graphical description of the two joint models. Joint model I (A) starts at time of alloSCT, joint model II (B) at time of the early low-dose DLI. Each model consists of a longitudinal and a time-to-event submodel and was run in turn for each T-cell subset, considering either the CD3+, CD4+ or CD8+ T-cell counts, and the NK cell counts. These are the outcome of the longitudinal submodel and a time-dependent covariate in the time-to-event submodel. All other variables in each submodel are baseline covariates. Per endpoint of the time-to-event submodels, the clinical events that occurred during the relevant time period (first 6 months after alloSCT or first 3 months after the early low-dose DLI) are described. The NK cells were only analyzed in model I. See the Statistical Supplement for a detailed description of the model structures.

counts. All patients started at time of alloSCT and were followed-up until 6 months after alloSCT or until the occurrence of an earlier endpoint (GvHD, relapse or other failure), whichever occurred first. The longitudinal submodel was a linear mixed-effects model, which used restricted cubic splines to flexibly model the log counts over time. The baseline covariates included in this submodel were disease risk (non-high risk or high risk), donor type (RD or UD with ATG-containing conditioning regimen) and patient/donor CMV status (both seronegative [CMV -/-] or not). The patient/donor CMV status was included as simple fixed effect, and both disease risk and donor type were included as part of a three-way interaction with time. This was in order to both properly accommodate the expected slower lymphocyte recovery in patients treated with ATG, and to evaluate a difference in trajectories between the disease risk groups. The time-to-event submodel comprised three cause-specific proportional hazards models, with GvHD, relapse and other failure as competing events. As predictors, they each contained the time-dependent current value (i.e. the underlying 'true' value at a given timepoint, as estimated by the longitudinal submodel) of the log immune cell count, as well as the baseline factors donor type and disease risk. The latter was omitted as a covariate from the model for 'other failure' due to the limited number of events.

To investigate whether the current slope (i.e. rate of increase or decrease of counts at a given moment) of the T-cell counts was associated with the development of GvHD, we also extended the models by adding the current slope of the log counts in addition to the current value to the time-to-event submodel (so-called 'time-dependent slopes' parametrization).

Model II (starting from early low-dose DLI)

To further investigate the T-cell kinetics after the early low-dose DLI, we constructed a joint model including only the patients who actually received the early low-dose DLI without any prior event of interest (Figure 1B). Since NK cells recover rapidly after alloSCT[27] (expected before the administration of early low-dose DLI in this study), they were not considered for model II. The time-scale was taken from DLI instead of from alloSCT, and follow-up was restricted to 3 months after this DLI, until administration of a second DLI, or until the occurrence of a terminating event, whichever occurred first. The disease risk factor was omitted since all included patients belonged to the high risk group. Since only 7 patients had a non-GvHD event within 3 months after the early low-dose DLI (Supplemental Figure 1), relapse and other failure were combined into one composite endpoint to compete with GvHD and the donor type factor was omitted for this composite endpoint.

Software

All analyses were performed in R version 4.2.1 using the packages JM[28] (version 1.5-2), survival[29] (version 3.4.0) and nlme[30] (3.1-157). Full code needed to reproduce the results of the present work is available at https://github.com/survival-lumc/ImmuneReconstJM, and structured using the targets[31] (version 0.14.0) package.

RESULTS

Population

166 patients were included in this study. Baseline characteristics are presented in Table 1. All surviving patients had at least 12 months follow-up since alloSCT. OS and RFS at

	Total cohort (N = 166)	Intention for early low-dose DLI (N = 62)	No intention for early low-dose DLI (N = 104)
Age at alloSCT (years)			
median (range)	63 (28-78)	64 (31-78)	63 (28-73)
Disease			
AML	133 (80%)	46 (74%)	87 (84%)
ALL	17 (10%)	10 (16%)	7 (7%)
MDS	16 (10%)	6 (10%)	10 (10%)
Nonmyeloablative conditioning			
Flu/Bu	150 (90%)*	46 (74%)	104 (100%)*
Flu/Bu/Ara-C/Amsa (FLAMSA)	16 (10%)	16 (26%)	0
Donor			
RD, 10/10 HLA matched	57 (34%)	20 (32%)	37 (36%)
UD, 10/10 HLA matched	101 (61%)	39 (63%)	62 (60%)
UD, 9/10 HLA matched	8 (5%)	3 (5%)	5 (5%)
Graft source			
G-CSF mobilized PBSC	165 (99%)	62 (100%)	103 (99%)
BM	1 (1%)	0	1 (1%)
CMV serostatus patient/donor			
+/+	79 (48%)	32 (52%)	47 (45%)
+/-	25 (15%)	8 (13%)	17 (16%)
-/+	11 (7%)	4 (6%)	7 (7%)
-/-	51 (31%)	18 (29%)	33 (32%)
Main reason for intention for early low-dose DLI			
FLAMSA regimen	-	16 (26%)	-
MRD+ at time of alloSCT	-	14 (23%)	-
AML/MDS: EVI1 overexpression	-	9 (15%)	-
AML: monosomal karyotype	-	8 (13%)	-
AML: ASXL mutation, only one remission induction course, or persisting underlying disease	-	4 (6%)	-
ALL: t(9;22)	-	4 (6%)	-
ALL: hypodiploidy, no CR1, or t(4;11)	-	4 (6%)	-
Therapy-related AML	-	2 (3%)	-
AML: progression before alloSCT	-	1 (2%)	-

Table 1. Baseline characteristics. Intention for early low-dose DLI is based on the anticipated high risk of relapse after alloSCT. DLI, donor lymphocyte infusion; alloSCT, allogeneic stem cell transplantation; AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; MDS, myelodysplastic syndrome; Flu, fludarabine; Bu, busulfan; Ara-C, cytarabine; Amsa, amsacrine; RD, related donor; UD, unrelated donor; G-CSF, granulocyte-colony stimulation factor; PBSC, peripheral blood stem cells; BM, bone marrow. *One patient had not received a second consolidation course before transplant and received 2 days cyclophosphamide 750 mg/m^2 intravenously additionally to the conditioning regimen.

6 months after alloSCT were 77% (95%-CI 71-83) and 70% (95%-CI 64-77), respectively. A total of 62 patients were considered to have a high risk of relapse and were scheduled for an early low-dose DLI, of whom 42 actually received it after a median interval of 3.1 months (range: 2.7-4.4) without any prior event of interest (Supplemental Figure 1). Twenty patients did not receive an early low-dose DLI: 10 because of early relapse, 9 because of early other failures (death [n=1], graft failure [n=2], start of systemic immunosuppression for a non-GvHD indication [n=4], or administration of a virus-specific T-cell infusion [n=2]), and 1 patient did not receive the early low-dose DLI because of mild skin GvHD requiring topical treatment. All 19 events occurred within 4 months after alloSCT. The patient with mild skin GvHD remained event-free for at least 51 months after alloSCT. None of the 104 non-high risk patients received an early low-dose DLI. At 6 months after alloSCT, the cumulative incidence of clinically significant GvHD was 26% (95%-CI 15-37) and 5% (95%-CI 0-9) for the high risk patients scheduled for early low-dose DLI and the non-high risk patients, respectively (Supplemental Figure 2). All clinically significant GvHD in the high risk patients occurred after administration of the early low-dose DLI (but before standard DLI) of which 88% occurred in patients receiving DLI from an UD after an ATG-containing conditioning regimen.

T-cell trajectories after alloSCT and DLI

DLI-related increase of T-cell counts after 3 months after alloSCT observed in patients with an unrelated donor

To investigate whether administration of the early low-dose DLI increased the numbers of circulating T cells during the first 6 months after alloSCT, we performed an ITT analysis using model I (see Methods) to compare the 62 high risk patients who were scheduled for early low-dose DLI with the 104 non-high risk patients who were not. All patients had at least 2 T-cell measurements with a median of 6 measurements per patient (interquartile range: 5-8). Although patients showed very different T- cell kinetics over time (Supplemental Figure 3), the model was flexible enough to capture the different shapes of patient-specific trajectories (Figure 2). Patients who were CMV seropositive or who had a CMV seropositive donor had significantly higher CD3 and CD8 counts during the first 6 months after TCD alloSCT compared to CMV seronegative patients with a CMV seronegative donor, corresponding to a significant increase on the log scale of 0.49 (95%-CI 0.31-0.67) and 0.45 (95%-CI 0.08-0.80) for CD3+ and CD8+ T cells, respectively. For instance, the model-based CD3 count at 6 months for a non-high risk patient with a RD was $425x10^6$/l if CMV -/- compared to $694x10^6$/l for any other CMV serostatus combination. The model-based CD8 count at this time was $222x10^6$/l compared to $347x10^6$/l, respectively, suggesting expansion of CMV-specific T cells. A same trend was observed for the CD4 counts (increase of 0.11 on the log scale, 95%-CI 0-0.23). As shown in Figure 3, patients with an UD had lower T-cell counts during the first 3 months after TCD alloSCT than patients with a RD, illustrating the enduring effect of the additional ATG that was given to all patients with an UD. We observed no significant difference in the cell count trajectories between the disease risk groups for patients with a RD. In contrast, in patients with an UD the CD4 trajectories started to diverge at 3 months after alloSCT, resulting in higher cell counts in the high risk patients intended to receive an early low-dose DLI at 3 months. The CD3 and CD8 counts

showed similar trends. Taken together, these data show that a strategy of early low-dose DLI can lead to T-cell expansion.

CD3, CD4 and CD8 counts increase after early low-dose DLI

To investigate whether the T-cell counts increased after the early low-dose DLI as the ITT-analysis suggested, we used model II including only the 42 patients who actually received this DLI without any prior event and modelled the kinetics during the first 3 months after DLI. One of the 42 patients did not have any T-cell measurement during this period and was excluded. Baseline characteristics of the 41 included patients are described in Supplemental Table 1. These patients had at least one T-cell measurement during the 3-month period after early low-dose DLI with a median of 4 measurements (interquartile range: 2-5). Again, a flexible model was constructed to capture the different shapes of the T-cell kinetics of the included patients (Supplemental Figure 4 and Supplemental Figure 5). The model-based trajectories of the total, CD4+ and CD8+ T-cell counts (Figure 4) showed increasing T-cell counts after DLI, with similar effects of the patient/donor CMV serostatus and donor type on the T-cell counts as in the earlier models.

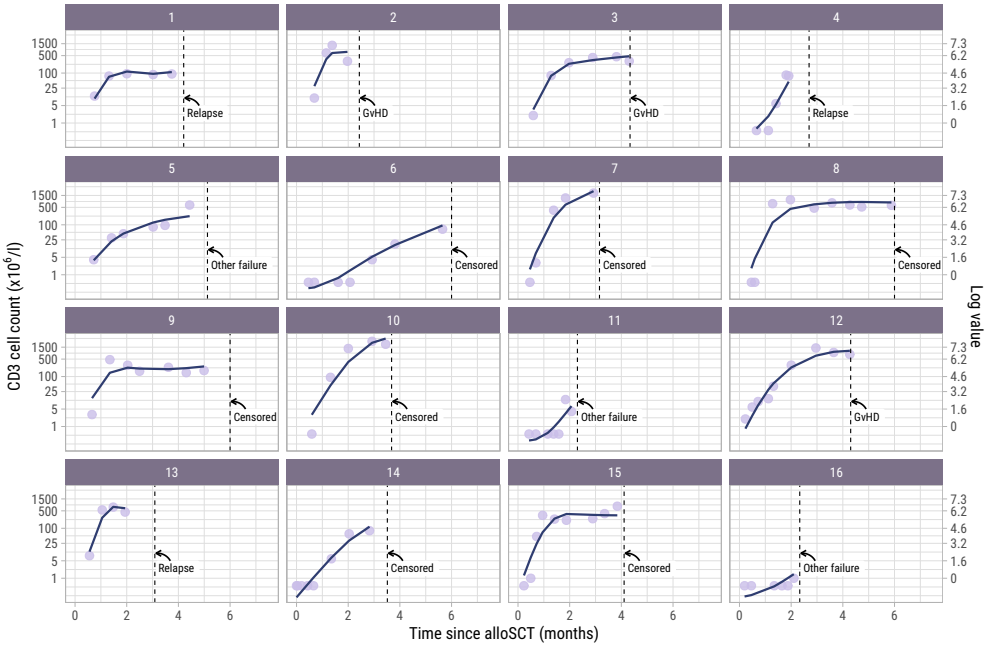


Figure 2. Observed versus estimated CD3 counts from alloSCT. Observed (dots) and estimated subject-specific trajectories (solid line) of a random subset of 16 patients in the dataset. The estimated trajectories are based on the longitudinal submodel of model I. Dotted lines show the time of terminating event or administrative censoring because of administration of a modified T-cell product or standard DLI. The secondary axis shows the cell counts on the log scale, which is the scale used for modelling. For example, a cell count of 1 on the primary axis corresponds to log(1) = 0 on the secondary axis.

Associations between T-cell kinetics and alloimmune responses after alloSCT and DLI

Higher CD3 and CD4 counts are associated with a higher risk of GvHD

To study the association between the T-cell kinetics and the development of GvHD or relapse after TCD alloSCT and DLI, we added disease risk and donor type as time-fixed covariates alongside the time-dependent T-cell counts in the cause-specific submodels (with GvHD, relapse and other failure as competing events) of model I. As shown in Figure 5, donor type showed no significant association with the risk of GvHD, although

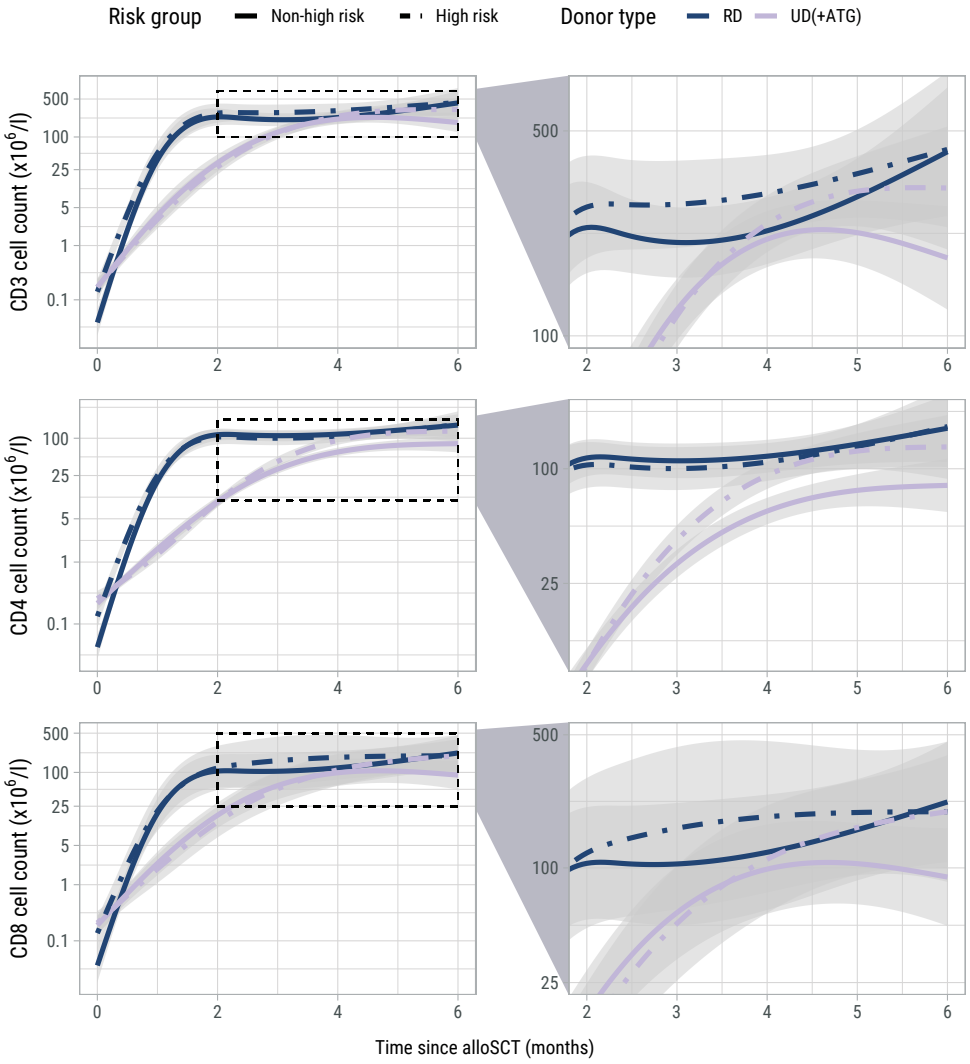


Figure 3. Model-based T-cell count trajectories after alloSCT. Predicted average trajectories of the total, CD4+ and CD8+ T-cell counts during the first 6 months after alloSCT, based on the longitudinal submodel of model I. For all predicted trajectories, the patient/donor CMV status was set to -/-. 95% confidence intervals are shown in grey. The right column zooms in on a specific part of the total trajectory.

in the CD4 model a trend for higher risk in patients with an UD despite the ATG in the conditioning regimen was observed (hazard ratio [HR] 2.7, 95%-CI 1.0-7.4). High risk patients, who were scheduled for early low-dose DLI, had a considerably higher risk of GvHD compared to non-high risk patients with HRs ranging between 6.3 (CD8 model, 95%-CI 2.1-18.8) and 7.3 (CD4 model, 95%-CI 2.4-22.2), indicating an alloimmune effect of the early low-dose DLI in this setting. The current values of the log CD4 and CD3 counts significantly increased the risk of GvHD (HR 2.4 (95%-CI 1.4-4.1) and HR 1.5 (95%-CI 1.0-2.3) for CD4+ T cells and CD3+ T cells, respectively), while CD8+ T cells showed a similar trend (HR 1.3, 95%-CI 0.9-1.8). These HRs represent the relative increase in GvHD risk for an increase of one in the log counts, assuming same disease risk and donor type. These results indicate that the absolute total numbers of circulating CD4+ and CD3+ T cells after alloSCT and DLI are informative for the development of GvHD.

We hypothesized that not only the current value but also the slope of the T-cell counts would be associated with the development of an alloimmune response. To investigate this, we extended the time-to-event submodel of model I by additionally including the current slope of the T-cell counts as a covariate for all endpoints. However, we observed no association between the slope of any of the T-cell subsets and the development of GvHD (p-values 0.59-0.87). We therefore retained the simpler version of model I with only the current value.

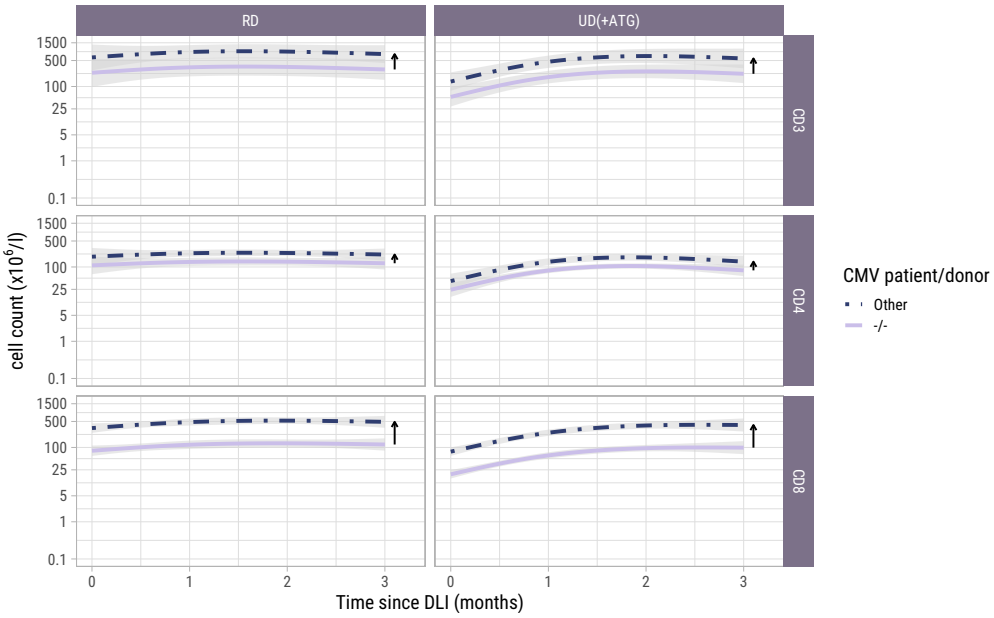


Figure 4. Model-based T-cell count trajectories after early low-dose DLI. Predicted average trajectories of the total, CD4+ and CD8+ T-cell counts during the first 3 months after early low-dose DLI. These are based on the longitudinal submodel of model II. 95% confidence intervals are shown in grey. The distance between the two lines in each panel (and further visualized by the adjacent arrows) corresponds to the CMV patient/donor effect on the trajectories. Namely, higher cell counts are predicted for patient/donor pairs where at least one is CMV seropositive, relative to a pair where both are CMV seronegative.

Protective effect of CD4+ T cells against relapse and other failure

To investigate whether higher T-cell counts were associated with a lower risk of relapse, we examined the risk factors for relapse in the time-to-event submodel of model I. Despite the ATG, patients with an UD had a significantly lower risk of relapse than patients with a RD (HRs ranging between 0.2 (95%-CI 0.1-0.5) and 0.3 (95%-CI 0.1-0.8), Figure 5). A trend was observed for higher relapse risk in the high risk patients (HR 2.1 in all models, 95%-CI for CD4+ T cells: 0.9-5.0, respectively), suggesting that the addition of early low-dose DLI to the strategy did not completely compensate for the higher relapse risk. While CD3+ and CD8+ T cells showed no significant association with relapse, higher CD4 counts decreased the risk of relapse significantly (HR 0.6, 95%-CI 0.5-0.9).

Of the 36 patients who experienced other failures, 6 died, 8 developed graft failure, 18 required systemic immunosuppression for a non-GvHD indication (of whom 9 received rituximab for EBV) and 4 received a virus-specific T-cell infusion for a severe viral infection. Only in the CD8 model a trend was observed for a higher risk of other failure in patients with an UD receiving an ATG-containing conditioning regimen (HR 2.6, 95%-CI 1.0-6.9). Higher CD4+ T-cell counts significantly lowered the hazard of the composite endpoint other failure (HR 0.7, 95%-CI 0.6-1.0).

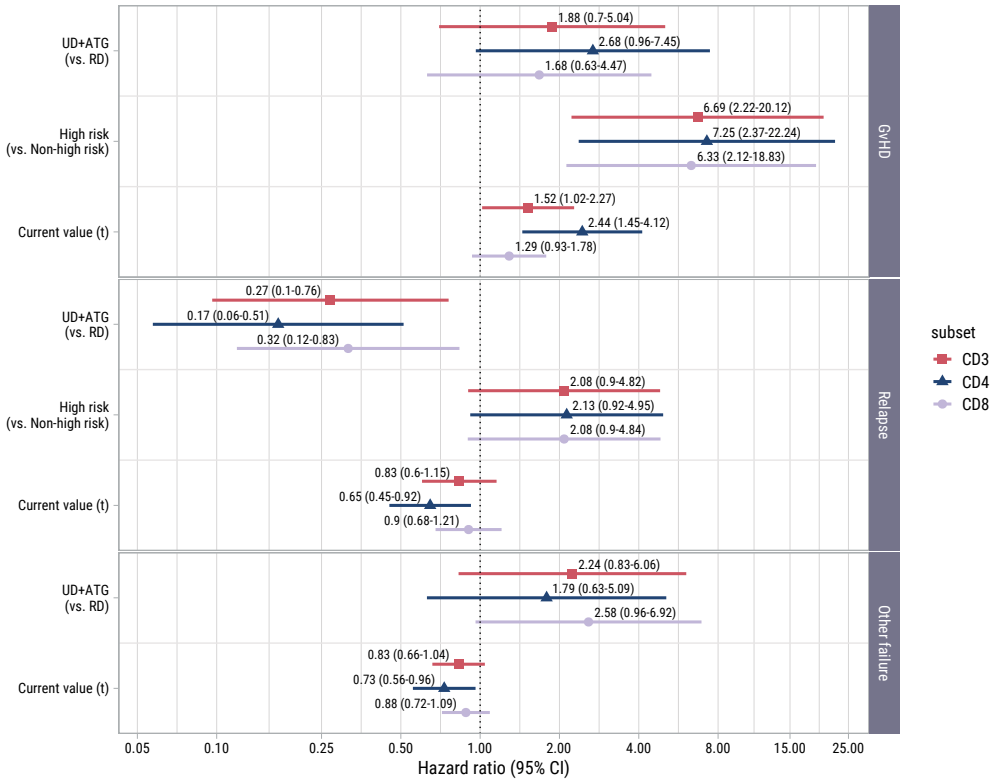


Figure 5. Forest plot for ITT analysis. Hazard ratios with associated 95% confidence intervals for donor type, disease risk and current value of the log of total, CD4+ or CD8+ T-cell counts on the events of interest. These are based on the time-to-event submodel of model I (see Figure 1A).

T-cell counts after early low-dose DLI retain their association with the development of GvHD

To investigate whether the T-cell kinetics were also associated with the development of alloimmune responses in the postDLI setting, we used the time-to-event submodel of model II starting from early low-dose DLI with GvHD and non-GvHD events as competing events. We observed no significant association between the current values and the very heterogenous composite endpoint of relapse and other failure (Figure 6). However, patients with an UD had a considerably higher risk of GvHD with HRs ranging between 7.0 (CD8+ T cells, 95%-CI 1.5-32.1) and 22.5 (CD4+ T cells, 95%-CI 3.7-138.9) compared to patients with a RD. For all T-cell subsets, higher current values increased the risk of GvHD with HRs ranging between 1.6 (CD8+ T cells, 95%-CI 1.0-2.6) and 6.7 (CD4+ T cells, 95%-CI 2.1-21.5). These data show that in the subset of patients receiving early low-dose DLI, total CD3+, CD4+ and CD8+ T-cell counts after DLI are associated with the development of GvHD.

NK cell kinetics and associations with alloimmune responses after alloSCT

To investigate the NK cell kinetics and their association with GvHD and relapse, we returned to model I starting at alloSCT. As shown in Supplemental Figure 6, the NK cell counts recovered rapidly, reaching the normal levels of 40-390x10^6 NK cells/l for almost all patients within 2 months, before the time of administration of the early low-dose DLI. As shown in Figure 7, CMV seropositive patients or patients with a CMV seropositive donor had significantly higher NK counts than CMV -/- patients, as was seen for the T-cell subsets. In contrast to T-cell kinetics, patients with an UD and ATG did not have a slower recovery of NK counts compared to patients with a RD and no ATG. Furthermore, there was no association between the risk group and NK counts, indicating that there was no impact of DLI on the NK cell kinetics. Higher current NK counts were associated with a higher risk of GvHD (HR 1.95 per unit log count increase,

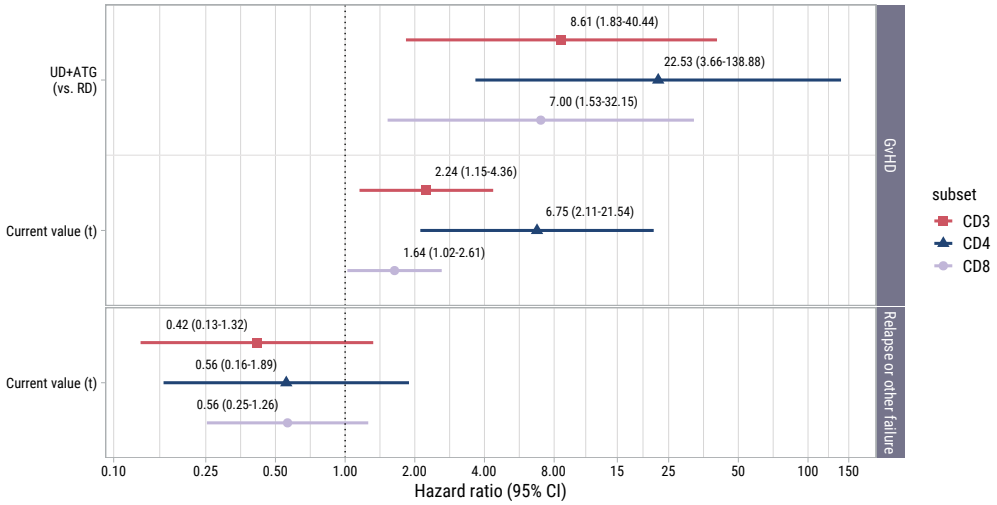


Figure 6. Forest plot for postDLI models. Hazard ratios with associated 95% confidence intervals for donor type and current value of the log of total, CD4+ or CD8+ T-cell counts on the events of interest. These are based on the time-to-event submodel of model II (see Figure 1B).

95%-CI 1.10-3.47) and a lower risk of relapse (HR 0.62, 95%-CI 0.41-0.93) but had no significant association with the risk of other failure. We hypothesized that the observed association between the NK count and GvHD may not be due to a direct effect of the NK cells, but instead reflected the high correlation between the NK and CD4 count trajectories, the latter being expected to be the main driver of GvHD. We therefore ran a cause-specific Cox model for GvHD, which included disease risk and donor type as time-fixed covariates, and both CD4 and NK counts as time-dependent covariates. In this model, CD4 counts were significantly associated with the development of GvHD (HR 2.08, 95%-CI 1.16-3.74) while the HR for the NK cell counts was 1.07 (p-value 0.83), supporting that the CD4+ T cells were the important drivers for the development of GvHD.

DISCUSSION

In this study we investigated the interplay between immune cell kinetics and alloimmune responses after both TCD alloSCT and subsequent DLI using joint modelling. In the ITT analysis we observed significantly more GvHD in the high risk patients intended to receive an early low-dose DLI and an increase in T-cell counts starting at 3 months after alloSCT in high risk patients with an UD receiving an ATG-containing conditioning regimen. The ITT allocation was solely based on the disease characteristics of the patients. Since all patients were in complete remission at time of alloSCT, the TCD strategy was similar between the disease risk groups, and all GvHD in the high risk group only occurred after DLI, the only plausible explanation for both the higher risk of GvHD and the associated T-cell expansion is the administration of the early low-dose DLI. We

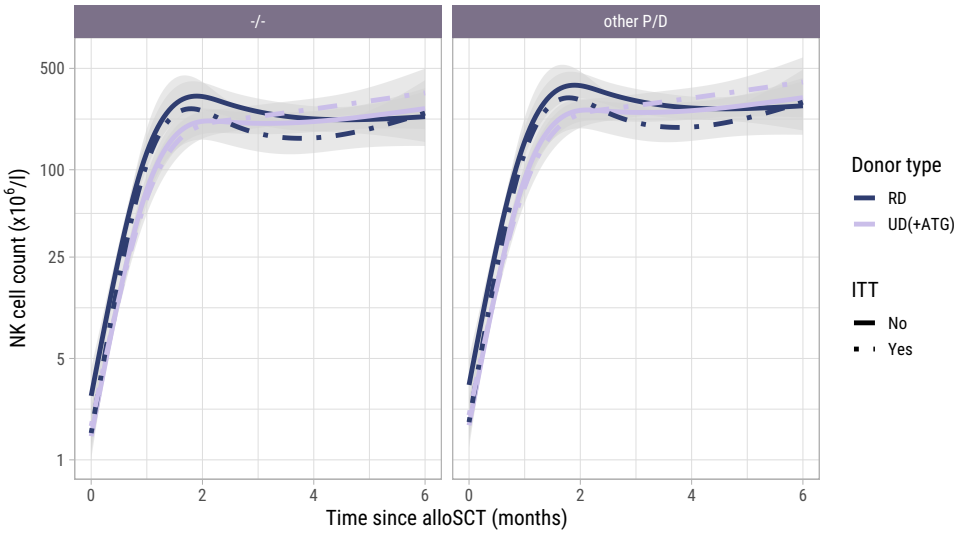


Figure 7. Model-based NK cell count trajectories after alloSCT. Predicted average trajectories of the NK cell counts during the first 6 months after alloSCT, based on the longitudinal submodel of model I. The left panel shows the predicted trajectories for CMV seronegative patients with a CMV seronegative donor, the right panel the predicted trajectories for patients with any other patient/donor CMV serostatus combination. 95% confidence intervals are shown in grey.

also observed significant associations between the CD4 counts and alloimmune responses after TCD alloSCT and DLI: an increase in CD4+ T cells was associated with a higher risk of GvHD and at the same time a lower risk of relapse suggesting establishment of a GvL effect. Interestingly, we only observed DLI-induced T-cell expansion in patients transplanted using an UD. This likely reflects an alloimmune response as GvHD was mainly seen in patients with an UD after receiving a DLI, and the T-cell counts after DLI were associated with the development of GvHD. The alloreactive T-cell expansion may have been more easily detectable in patients with an UD compared to RD because of the deeper lymphopenia at time of DLI due to the long-lasting immunosuppressive effect of ATG that patients with an UD received.[13] In addition, the high prevalence of HLA-DP mismatches, targeted by CD4+ T cells, in patients with an UD[32-34] could contribute to the strong association between CD4+ T cells and the development of GvHD. In contrast to T cells, NK cells recovered early after alloSCT and were not significantly influenced by donor type and TCD, consistent with previous studies[13,35,36], nor by DLI. As previously reported[3,37], higher NK counts were associated with a lower risk of relapse. The joint model also suggested that higher NK counts were associated with a higher risk of GvHD. However, in an exploratory cause-specific Cox model, this association between NK cells and GvHD disappeared after adjusting for the CD4 counts, indicating that the CD4+ T cells were the important drivers for GvHD.

Our results suggest a DLI-induced T-cell expansion measurable in total numbers of the major T-cell subsets where others did not observe a significant effect of DLI on the T-cell kinetics.[18-21] This may be due to several factors. Our comparatively larger cohort size (other studies usually included less than 25 patients) allowed for detection of more subtle differences. Furthermore, the strategy of administering early prophylactic DLI to a subset of patients based on their relapse risk provided an intervention and control group who were treated according to the same transplantation strategy. Lastly, conclusions drawn can be influenced by the choice of the statistical method. For example, matched pair analysis as used by Guillaume et al.[19] and Schultze-Florey at al[21] only allowed them to compare the cells counts between two timepoints. The repeated measures analysis used by Nikiforow et al.[20] and the mixed model used by Bullucini et al.[18] allowed to compare the trajectories over time but could not account for informative dropout. Because we used joint modelling, we could flexibly model the T-cell trajectories over a longer period of time and properly account for informative dropout and random variation. To our knowledge, thus far only a single study used joint modelling to study T-cell kinetics after alloSCT.[38] We now have used this technique to investigate the immunological effects of DLI.

There are several limitations to our study. The total CD3, CD4 and CD8 counts are crude measures for potentially alloreactive T cells, as only donor-derived T cells can induce GvHD and GvL and the counts are not informative about the subpopulations, activation status or kinetics of specific T-cell clones. Thus, if we had measured the chimerism status and clonality, we might have expected to find stronger associations between the T-cell kinetics and the clinical events. Moreover, our ITT approach attenuated the observed effects of DLI on the T-cell kinetics and clinical endpoints as not all high risk patients received the early low-dose DLI and most patients who did receive this DLI did not receive it at exactly the same time after transplant. Therefore, we constructed model II starting from early low-dose DLI to see whether similar associations

were observed. Joint modelling requires substantial numbers of both clinical events and longitudinal measurements to estimate associations with sufficient accuracy. Despite our comparatively larger sample size, the modest numbers of clinical events limited both the accurate estimation of association parameters (between T-cell counts and the endpoints), as well as the inclusion of additional risk factors for each endpoint. This was especially noticeable in our models focusing on the subset of the patients actually receiving an early low-dose DLI. Due to the limited number of events, we used suboptimal composite endpoints such as 'other failure' and 'relapse and other failure', which hampered estimation of the association between the T-cell kinetics and these endpoints.

Further studies are necessary to assess the clinical implications of the findings from the present work. Aside from validation of our findings, larger studies must be performed to investigate the predictive utility of the T-cell and NK cell counts. While these counts are crude measures, they are often measured standardly and therefore attractive biomarkers for predicting alloimmune responses in patients receiving alloSCT and/or DLI. Further investigation of the immune cell kinetics in other alloSCT settings is needed to see whether similar associations between the T-cell and NK cell kinetics and alloimmune responses can be observed when using joint modelling. For instance, the recent machine learning analysis by McCurdy et al. also suggested important roles of CD4+ T cells in the development of acute GvHD and of NK cells in the development of relapse after alloSCT with posttransplant cyclophosphamide.[37] For DLI, we would suggest to perform a prospective study where the T-cell counts are measured at time of DLI and every week after DLI during the first 6 weeks. Most GvHD develops within this period and by measuring more often, dynamic prediction tools (i.e. updated personalized probabilities of GvHD given measurement history) could be developed.[39] In order to develop such tools however, one would ideally need to model the T-cell subsets and NK cells *jointly* as part of a multivariate joint model, which will account for the correlation between each subset, but may be complicated to fit and will require larger sample sizes. In our study, we were not able to present such a multivariate joint model because of both sample size and software limitations. Nevertheless, results from the exploratory time-dependent cause-specific Cox model for GvHD with both the CD4 and NK counts hint at the importance of modelling immune subsets jointly. Generally speaking, further characterization of the circulating T-cell subsets, differentiation and metabolic fitness could provide valuable additional insight in future studies on T-cell kinetics.[40,41]

In summary, joint modelling allowed us to capture the associations between DLI, T-cell and NK cell counts, GvHD and relapse in a very complex clinical setting, even with modest numbers of patients and events. NK cells recover early after alloSCT and may have a protective effect against relapse. We demonstrate that DLI can induce detectable T-cell expansion and observe that the CD4+ T cells show the strongest association with the development of alloimmune responses. Higher CD4 counts increase the risk of GvHD and decrease the risk of relapse.

ACKNOWLEDGMENTS

The authors would like to thank E.J.M. Bogers, S.M. Wallet-Malicka and H.A. Baelde for the data management and technical support.

REFERENCES

1. Horowitz, M. M., Gale, R. P., Sondel, P. M. et al. Graft-versus-leukemia reactions after bone marrow transplantation. Blood 75, 555-562 (1990).
2. Falkenburg, J. H. F. & Jedema, I. Graft versus tumor effects and why people relapse. Hematology. American Society of Hematology. Education Program 2017, 693-698 (2017). doi: 10.1182/asheducation-2017.1.693
3. Dunbar, E. M., Buzzeo, M. P., Levine, J. B. et al. The relationship between circulating natural killer cells after reduced intensity conditioning hematopoietic stem cell transplantation and relapse-free survival and graft-versus-host disease. Haematologica 93, 1852-1858 (2008). doi: 10.3324/haematol.13033
4. Minculescu, L., Marquart, H. V., Friis, L. S. et al. Early Natural Killer Cell Reconstitution Predicts Overall Survival in T Cell-Replete Allogeneic Hematopoietic Stem Cell Transplantation. Biol Blood Marrow Transplant 22, 2187-2193 (2016). doi: 10.1016/j.bbmt.2016.09.006
5. Simonetta, F., Alvarez, M. & Negrin, R. S. Natural Killer Cells in Graft-versus-Host-Disease after Allogeneic Hematopoietic Cell Transplantation. Front Immunol 8, 465 (2017). doi: 10.3389/fimmu.2017.00465
6. Busca, A. & Aversa, F. In-vivo or ex-vivo T cell depletion or both to prevent graft-versus-host disease after hematopoietic stem cell transplantation. Expert Opin Biol Ther 17, 1401-1415 (2017). doi: 10.1080/14712598.2017.1369949
7. Eefting, M., Halkes, C. J., de Wreede, L. C. et al. Myeloablative T cell-depleted alloSCT with early sequential prophylactic donor lymphocyte infusion is an efficient and safe post-remission treatment for adult ALL. Bone Marrow Transplant 49, 287-291 (2014). doi: 10.1038/bmt.2013.111
8. Eefting, M., de Wreede, L. C., Halkes, C. J. et al. Multi-state analysis illustrates treatment success after stem cell transplantation for acute myeloid leukemia followed by donor lymphocyte infusion. Haematologica 101, 506-514 (2016). doi: 10.3324/haematol.2015.136846
9. Yun, H. D. & Waller, E. K. Finding the sweet spot for donor lymphocyte infusions. Biol Blood Marrow Transplant 19, 507-508 (2013). doi: 10.1016/j.bbmt.2013.02.005
10. Krishnamurthy, P., Potter, V. T., Barber, L. D. et al. Outcome of donor lymphocyte infusion after T cell-depleted allogeneic hematopoietic stem cell transplantation for acute myelogenous leukemia and myelodysplastic syndromes. Biol Blood Marrow Transplant 19, 562-568 (2013). doi: 10.1016/j.bbmt.2012.12.013
11. Falkenburg, J., Schmid, C., Kolb, H., Locatelli, F. & Kuball, J. in The EBMT Handbook. Hematopoietic Stem Cell Transplantation and Cellular Therapies. (eds E. Carreras, C. Dufour, M. Mohty, & N. Kroger) 443-448 (Springer, 2019).
12. Yanir, A., Schulz, A., Lawitschka, A., Nierkens, S. & Eyrich, M. Immune Reconstitution After Allogeneic Haematopoietic Cell Transplantation: From Observational Studies to Targeted Interventions. Front Pediatr 9, 786017 (2021). doi: 10.3389/fped.2021.786017
13. Bosch, M., Dhadda, M., Hoegh-Petersen, M. et al. Immune reconstitution after anti-thymocyte globulin-conditioned hematopoietic cell transplantation. Cytotherapy 14, 1258-1275 (2012). doi: 10.3109/14653249.2012.715243
14. Stern, L., McGuire, H. M., Avdic, S. et al. Immunoprofiling reveals cell subsets associated with the trajectory of cytomegalovirus reactivation post stem cell transplantation. Nat Commun 13, 2603 (2022). doi: 10.1038/s41467-022-29943-9
15. Hassan, N., Eldershaw, S., Stephens, C. et al. CMV reactivation initiates long-term expansion and differentiation of the NK cell repertoire. Front Immunol 13, 935949 (2022). doi: 10.3389/fimmu.2022.935949
16. Lewalle, P., Triffet, A., Delforge, A. et al. Donor lymphocyte infusions in adult haploidentical transplant: a dose finding study. Bone Marrow Transplantation 31, 39-44 (2003). doi: 10.1038/sj.bmt.1703779
17. Schmaelter, A.-K., Waidhauser, J., Kaiser, D. et al. Alterations of Peripheral Blood T Cell Subsets following Donor Lymphocyte Infusion in Patients after Allogeneic Stem Cell Transplantation. Hemato 2, 692-702 (2021).
18. Bellucci, R., Alyea, E. P., Weller, E. et al. Immunologic effects of prophylactic donor lymphocyte infusion after allogeneic marrow transplantation for multiple myeloma. Blood 99, 4610-4617

(2002). doi: 10.1182/blood.v99.12.4610
19. Guillaume, T., Gaugler, B., Chevallier, P. et al. Escalated lymphodepletion followed by donor lymphocyte infusion can induce a graft-versus-host response without overwhelming toxicity. Bone Marrow Transplantation 47, 1112-1117 (2012). doi: 10.1038/bmt.2011.231
20. Nikiforow, S., Kim, H. T., Daley, H. et al. A phase I study of CD25/regulatory T-cell-depleted donor lymphocyte infusion for relapse after allogeneic stem cell transplantation. Haematologica 101, 1251-1259 (2016). doi: 10.3324/haematol.2015.141176
21. Schultze-Florey, C. R., Kuhlmann, L., Raha, S. et al. Clonal expansion of CD8+ T cells reflects graft-versus-leukemia activity and precedes durable remission following DLI. Blood advances (2021). doi: 10.1182/bloodadvances.2020004073
22. Toor, A. A., Sabo, R. T., Roberts, C. H. et al. Dynamical System Modeling of Immune Reconstitution after Allogeneic Stem Cell Transplantation Identifies Patients at Risk for Adverse Outcomes. Biol Blood Marrow Transplant 21, 1237-1245 (2015). doi: 10.1016/j.bbmt.2015.03.011
23. Gooptu, M., Kim, H. T., Howard, A. et al. Effect of Sirolimus on Immune Reconstitution Following Myeloablative Allogeneic Stem Cell Transplantation: An Ancillary Analysis of a Randomized Controlled Trial Comparing Tacrolimus/Sirolimus and Tacrolimus/Methotrexate (Blood and Marrow Transplant Clinical Trials Network/BMT CTN 0402). Biol Blood Marrow Transplant 25, 2143-2151 (2019). doi: 10.1016/j.bbmt.2019.06.029
24. Rizopoulos, D. (Boca Raton: Chapman and Hall/CRC, 2012).
25. von dem Borne, P. A., Starrenburg, C. W., Halkes, S. J. et al. Reduced-intensity conditioning allogeneic stem cell transplantation with donor T-cell depletion using alemtuzumab added to the graft ('Campath in the bag'). Current opinion in oncology 21 Suppl 1, S27-29 (2009). doi: 10.1097/01.cco.0000357472.76337.0e
26. Koster, E. A. S., von dem Borne, P. A., van Balen, P. et al. Competitive Repopulation and Allo-Immunologic Pressure Determine Chimerism Kinetics after T Cell-Depleted Allogeneic Stem Cell Transplantation and Donor Lymphocyte Infusion. Transplant Cell Ther 29, 268.e261-268.e210 (2023). doi: 10.1016/j.jtct.2022.12.022
27. Elfeky, R., Lazareva, A., Qasim, W. & Veys, P. Immune reconstitution following hematopoietic stem cell transplantation using different stem cell sources. Expert review of clinical immunology 15, 735-751 (2019). doi: 10.1080/1744666x.2019.1612746
28. Rizopoulos, D. JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. Journal of Statistical Software 35, 1 - 33 (2010). doi: 10.18637/jss.v035.i09
29. Therneau, T. M. A Package for Survival Analysis in R, <https://CRAN.R-project.org/package=survival> (2023).
30. Pinheiro, J., Bates, D. & R Core Team. nlme: Linear and Nonlinear Mixed Effects Models, <https://CRAN.R-project.org/package=nlme> (2023).
31. Landau, W. M. The targets R package: a dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. Journal of Open Source Software 6, 2959 (2021). doi: 10.21105/joss.02959
32. Mariano, L., Zhang, B. M., Osoegawa, K., Lowsky, R. & Fernandez-Vina, M. Assessment by Extended-Coverage Next-Generation Sequencing Typing of DPA1 and DPB1 Mismatches in Siblings Matching at HLA-A, -B, -C, -DRB1, and -DQ Loci. Biol Blood Marrow Transplant 25, 2507-2509 (2019). doi: 10.1016/j.bbmt.2019.07.033
33. Shaw, B. E., Gooley, T. A., Malkki, M. et al. The importance of HLA-DPB1 in unrelated donor hematopoietic cell transplantation. Blood 110, 4560-4566 (2007). doi: 10.1182/blood-2007-06-095265
34. Fleischhauer, K., Shaw, B. E., Gooley, T. et al. Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study. The Lancet. Oncology 13, 366-374 (2012). doi: 10.1016/s1470-2045(12)70004-9
35. Penack, O., Fischer, L., Stroux, A. et al. Serotherapy with thymoglobulin and alemtuzumab differentially influences frequency and function of natural killer cells after allogeneic stem cell transplantation. Bone Marrow Transplant 41, 377-383 (2008). doi: 10.1038/sj.bmt.1705911
36. Ito, A., Kitano, S., Tajima, K. et al. Impact of low-dose anti-thymocyte globulin on immune reconstitution after allogeneic hematopoietic cell transplantation. International journal of hematology 111, 120-130 (2020). doi: 10.1007/s12185-019-02756-1

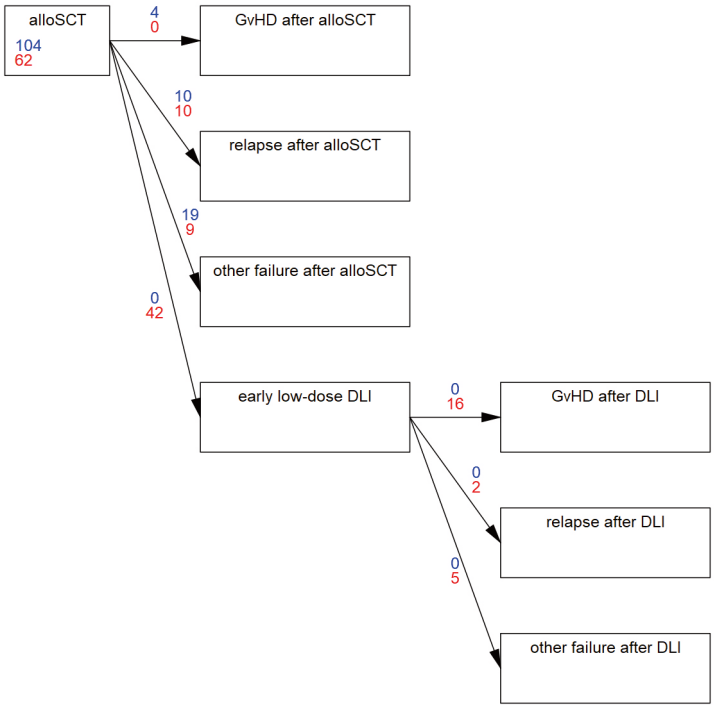
37. McCurdy, S. R., Radojcic, V., Tsai, H. L. et al. Signatures of GVHD and relapse after posttransplant cyclophosphamide revealed by immune profiling and machine learning. Blood 139, 608-623 (2022). doi: 10.1182/blood.2021013054
38. Salzmann-Manrique, E., Bremm, M., Huenecke, S. et al. Joint Modeling of Immune Reconstitution Post Haploidentical Stem Cell Transplantation in Pediatric Patients With Acute Leukemia Comparing CD34(+)-Selected to CD3/CD19-Depleted Grafts in a Retrospective Multicenter Study. Front Immunol 9, 1841 (2018). doi: 10.3389/fimmu.2018.01841
39. Andrinopoulou, E. R., Rizopoulos, D., Geleijnse, M. L. et al. Dynamic prediction of outcome for patients with severe aortic stenosis: application of joint models for longitudinal and time-to-event data. BMC Cardiovasc Disord 15, 28 (2015). doi: 10.1186/s12872-015-0035-z
40. Dekker, L., de Koning, C., Lindemans, C. & Nierkens, S. Reconstitution of T Cell Subsets Following Allogeneic Hematopoietic Cell Transplantation. Cancers (Basel) 12 (2020). doi: 10.3390/cancers12071974
41. Uhl, F. M., Chen, S., O'Sullivan, D. et al. Metabolic reprogramming of donor T cells enhances graft-versus-leukemia effects in mice and humans. Science translational medicine 12 (2020). doi: 10.1126/scitranslmed.abb8969

SUPPLEMENTAL TABLES

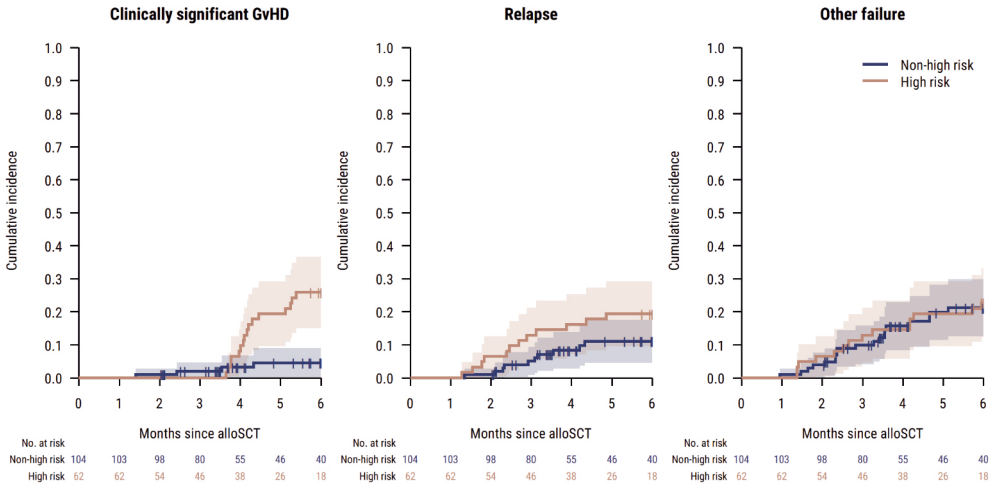
	Early low-dose DLI (N = 41)
Age at alloSCT (years)	
median (range)	65 (31-74)
Disease	
AML	29 (71%)
ALL	8 (20%)
MDS	4 (10%)
Nonmyeloablative conditioning	
Flu/Bu	30 (73%)
Flu/Bu/Ara-C/Amsa (FLAMSA)	11 (27%)
Donor	
RD, 10/10 HLA matched	12 (29%)
UD, 10/10 HLA matched	27 (66%)
UD, 9/10 HLA matched	2 (5%)
Graft source	
G-CSF mobilized PBSC	42 (100%)
CMV serostatus patient/donor	
+/+	17 (41%)
+/-	5 (12%)
-/+	3 (7%)
-/-	16 (39%)
Reason for early low-dose DLI	
Conditioning using the FLAMSA regimen	11 (27%)
MRD+ at time of alloSCT	10 (24%)
ALL: t(9;22)	3 (7%)
ALL: t(4;11), hypodiploidy, or not in CR1	3 (7%)
AML: monosomal karyotype	5 (12%)
AML/MDS: EV1 overexpression	6 (15%)
AML: ASXL mutation, only 1 intensive remission induction course, or persisting CMML	3 (7%)

Supplemental Table 1. Baseline characteristics of the 41 evaluable patients with early low-dose DLI. DLI, donor lymphocyte infusion; alloSCT, allogeneic stem cell transplantation; AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; MDS, myelodysplastic syndrome; Flu, fludarabine; Bu, busulfan; Ara-C, cytarabine; Amsa, amsacrine; RD, related donor; UD, unrelated donor; G-CSF, granulocyte-colony stimulation factor; PBSC, peripheral blood stem cells; BM, bone marrow; MRD, minimal residual disease; CR1, first complete morphological remission; CMML, chronic myelomonocytic leukemia

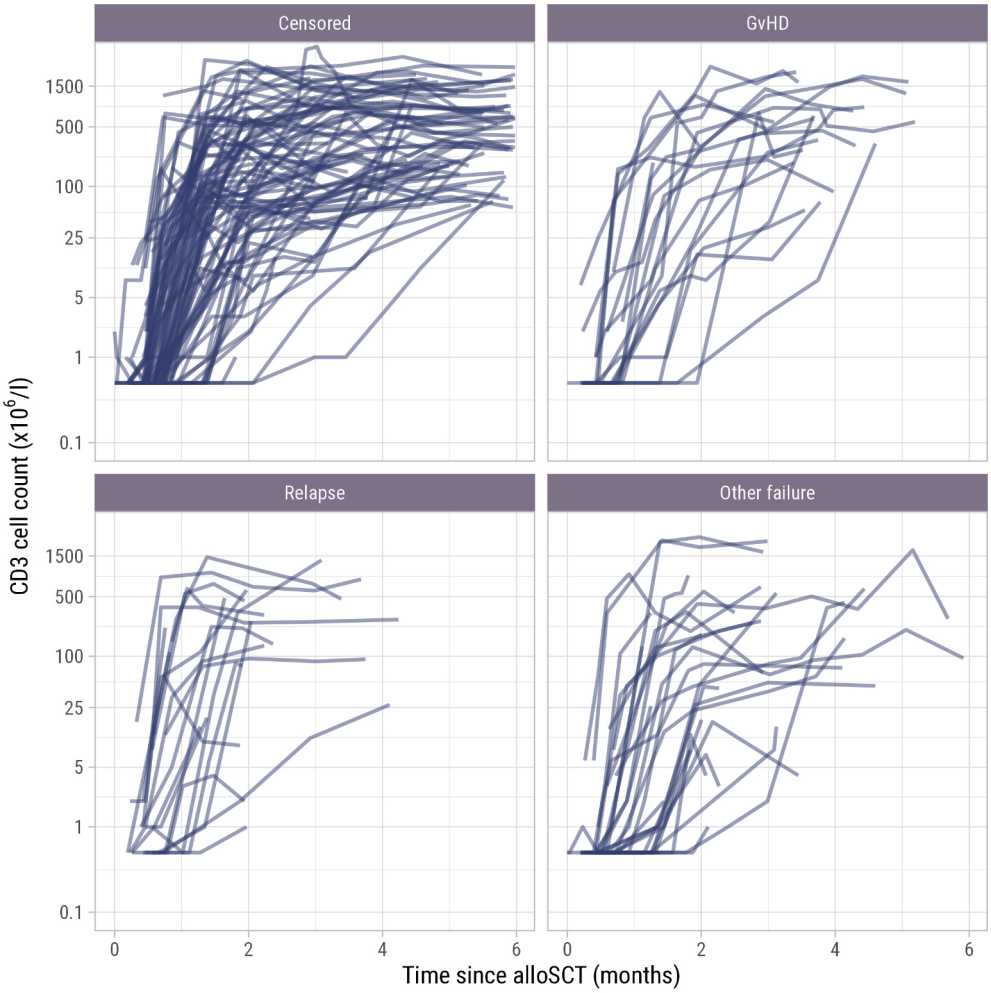
SUPPLEMENTAL FIGURES



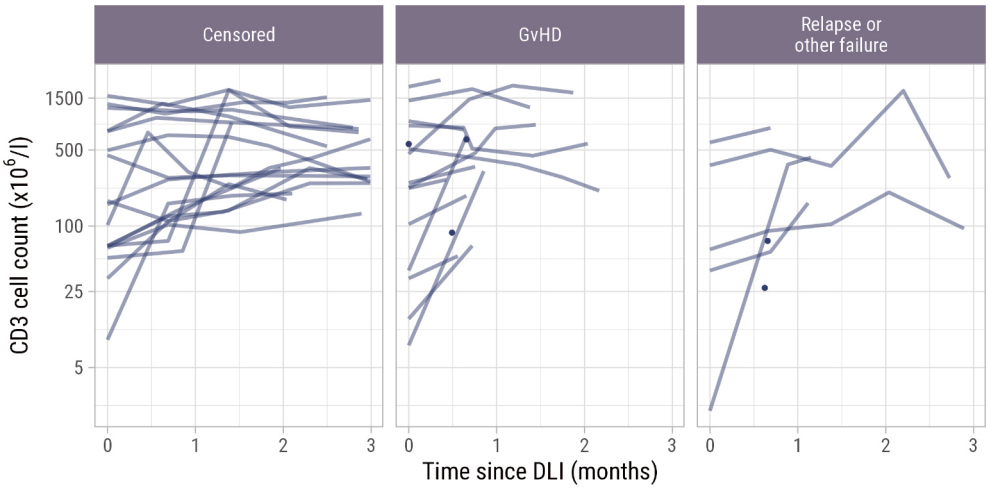
Supplemental Figure 1. Flow diagram of events during the first 6 months after alloSCT. Flow diagram of the events of interest after alloSCT and early low-dose DLI. The numbers in the top left box show the total numbers of included high risk patients scheduled for early low-dose DLI (red) and non-high risk patients (blue). The numbers next to the arrows show the numbers of the patients who had the respective event during the first 6 months after alloSCT without any prior administration of a modified T-cell product or standard DLI (blue: non-high risk, red: high risk). For instance, all high-risk patients received an early low-dose DLI or developed clinically significant GvHD, relapse or other failure before this DLI could be administered, except one patient who only had mild GvHD and did not need any systemic immunosuppression: therefore, the red numbers along the leftmost set of arrows add up to 61 while 62 started in the left box.



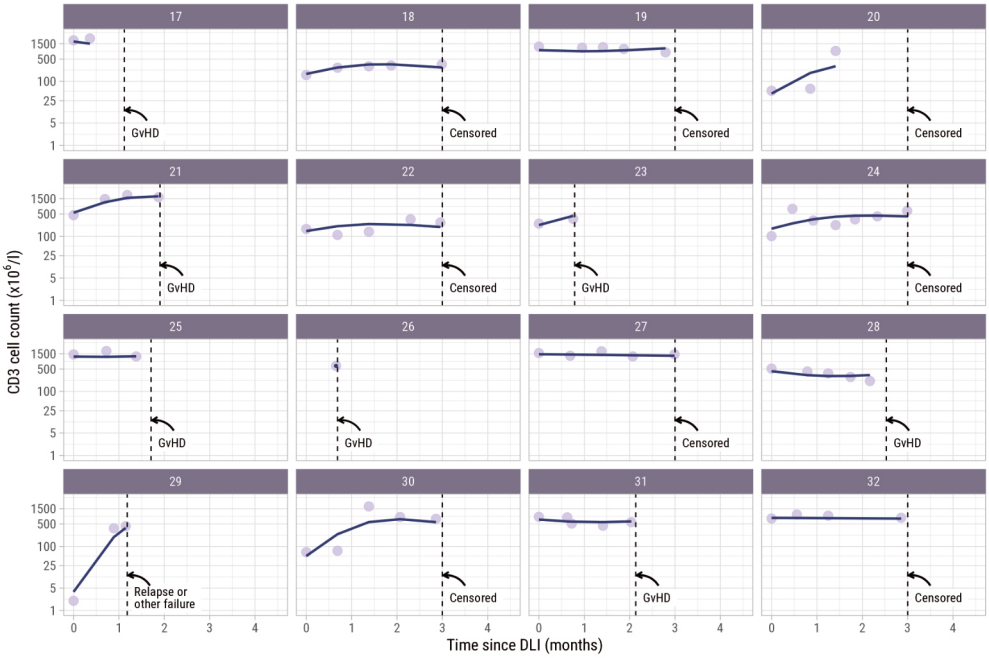
Supplemental Figure 2. Cumulative incidence of GvHD, relapse and other failure per disease risk group. Cumulative incidence of the competing events GvHD, relapse and other failure with associated 95% confidence intervals stratified by disease risk. Patients with a high anticipated risk of relapse were scheduled to receive an early low-dose DLI at 3 months after alloSCT. Contrary to Supplemental Figure 1, early low-dose DLI was not treated as an event in this figure. Patients who received a modified T-cell product or standard DLI were censored at 7 days after this DLI, indicated by |.



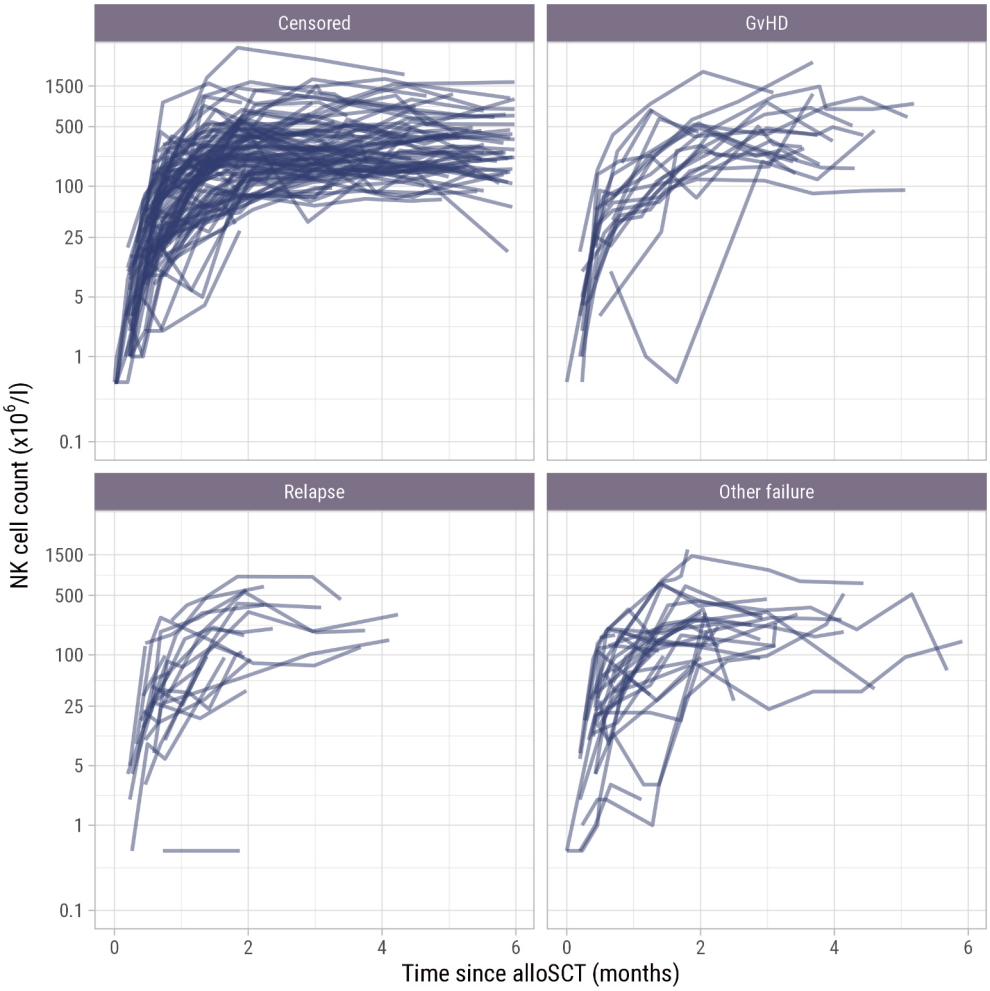
Supplemental Figure 3. Trajectories of total T-cell counts from alloSCT per terminating event. All observed trajectories for the CD3 counts during the first 6 months after alloSCT per terminating event. Patients were censored at 6 months after alloSCT, or 7 days after administration of a standard DLI or modified T-cell product, whichever occurred first. There was no loss to follow-up.



Supplemental Figure 4. Trajectories total T-cell counts after early low-dose DLI per terminating event. All observed trajectories for the CD3 counts during the first 3 months after early low-dose DLI per terminating event. The single points correspond to patients with only a single measurement between their DLI and terminating event. Patients were censored at 6 months after alloSCT, or 7 days after administration of a standard DLI or modified T-cell product, whichever occurred first. There was no loss to follow-up.



Supplemental Figure 5. Observed versus estimated CD3 counts after early low-dose DLI. Observed (dots) and estimated subject-specific trajectories (solid lines) of a random subset of 16 patients in the dataset. The estimated trajectories are based on the longitudinal submodel of model II. Dotted lines show the time of terminating event or administrative censoring because of administration of a modified T-cell product or standard DLI at 6 months after alloSCT.



Supplemental Figure 6. Trajectories of NK cell counts from alloSCT per terminating event. All observed trajectories for the NK counts during the first 6 months after alloSCT per terminating event. Patients were censored at 6 months after alloSCT, or 7 days after administration of a standard DLI or modified T-cell product, whichever occurred first. There was no loss to follow-up.

STATISTICAL SUPPLEMENT

The present Supplemental material is a 'Statistical Supplement' to the main article, providing mathematical summaries of the models used.

Joint model I

Joint models only consider measurements taken prior to the occurrence of the clinical events of interest. Occasionally, the measurement time and event time coincide: for example, T-cell counts may be recorded on the same day as the start of therapeutic systemic immunosuppression for Graft-versus-Host-Disease (GvHD). In order to retain the information of the measurements taken at event times, we set the time of these measurements to one day earlier, which assumes that the measurement at the event time was representative of the T-cell counts the day before the event. However, we excluded measurements at time of relapse, since the presence of blasts in the peripheral blood could lead to incorrect counts of the normal T cells. We also excluded measurements at time of autologous recovery, as donor-derived T cells were no longer present, and therefore also no potentially alloreactive T cells capable of inducing GvHD or Graft-versus-leukemia (GvL) effect.

Model formulation

The longitudinal submodel assumes that the true underlying (log) immune cell counts (either CD3, CD4, CD8, or NK) for the i^{th} patient are given by

$$\begin{aligned} m_i(t) = \beta_0 + \sum_{q=1}^{3}(\beta_q + b_{iq})B_q(t) + \sum_{q=1}^{3}\beta_{q+3}\{B_q(t) \times \text{Risk}_i\} + \sum_{q=1}^{3}\beta_{q+6}\{B_q(t) \times \text{Donor}_i\} \\ + \sum_{q=1}^{3}\beta_{q+9}\{B_q(t) \times \text{Risk}_i \times \text{Donor}_i\} + \beta_{13}\text{CMV}_i + \beta_{14}\text{Risk}_i + \beta_{15}\text{Donor}_i \\ + \beta_{16}\{\text{Risk}_i \times \text{Donor}_i\}, \end{aligned}$$

with random effects vector $b_i \sim \mathcal{N}(0, D)$. The observations for the i^{th} patient at timepoints t_{ij} $(j = 1, \ldots, n_i)$ are given by

$$y_{ij} = m_i(t_{ij}) + \epsilon_{ij},$$

where $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ are independent random error terms.

Risk_i, Donor_i and CMV_i respectively represent the dummy variables for baseline disease risk (the intention-to-treat variable, high-risk compared to non-high risk), donor type (unrelated compared to related donor) and patient/donor Cytomegalovirus (CMV) serostatus at baseline (any one of patient or donor positive, compared to patient and donor both negative).

Time since allogeneic stem cell transplantation (alloSCT) was modelled flexibly assuming restricted (natural) cubic splines with two internal knots placed at the 33.3% and 66.7% percentiles of the measurement times. This is represented above by $B_q(t)$, corresponding to the q^{th} basis function of the spline. The fixed effects part of the model posits a three-way interaction between time, donor type and baseline disease risk, as well as a main effect of patient/donor CMV status. The three-way interaction was constructed to a) capture the slower expected average trajectory of patients with an

unrelated donor, due to the use of anti-thymocyte globulin (ATG) in this group; and b) to test for a difference in average trajectories between baseline disease risk groups.

In terms of random effects, this models assumes random slopes b_{iq} (one for each basis function), and a fixed intercept. This fixed intercept was justified given that this cohort underwent T-cell depleted (TCD) alloSCT, and all patients were therefore expected to start follow-up with immune cell counts close to zero. The random slopes were assumed to be normally distributed with mean zero, with unstructured covariance matrix D.

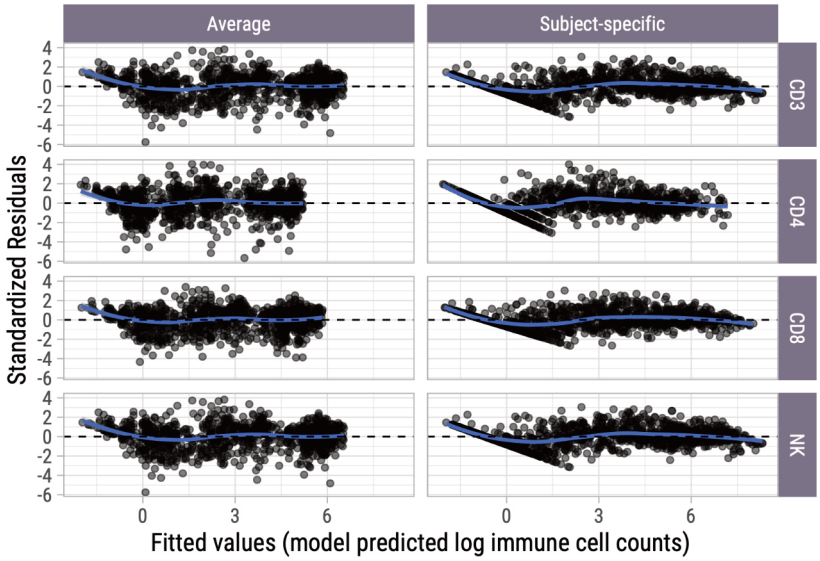
The time-to-event submodel was composed of multiple cause-specific proportional hazards models as

$$h_{1i}(t) = h_{10}(t)\exp\left\{\gamma_{11}\text{Donor}_i + \gamma_{12}\text{Risk}_i + \alpha_1 m_i(t)\right\},$$
$$h_{2i}(t) = h_{20}(t)\exp\left\{\gamma_{21}\text{Donor}_i + \gamma_{22}\text{Risk}_i + \alpha_2 m_i(t)\right\},$$
$$h_{3i}(t) = h_{30}(t)\exp\left\{\gamma_{31}\text{Donor}_i + \alpha_3 m_i(t)\right\},$$

where the $h_{ki}(t)$ for $k \in \{1, 2, 3\}$ respectively represent the cause-specific hazards of GvHD, relapse, and other failures. The cause-specific baseline hazards $h_{k0}(t)$ were approximated on the log scale using cubic B-splines with three internal knots. The above corresponds to the 'current value' parametrization of the joint model, where the $\exp(\alpha_k)$ would represent the hazard ratio (for cause k) when comparing two patients (with same covariates) whose 'true' (model-based) underlying log immune cell values at a particular timepoint $m_i(t)$ differ by one. The γ_{kp} coefficients are interpreted analogously to main effects in standard cause-specific Cox proportional hazards models.

In addition to the current value parametrization, we also ran the models assuming a time-dependent slopes association structure as $\alpha_{k1}m_i(t) + \alpha_{k2}\{dmi(t)/dt\}$.

Goodness of fit



On the previous page we present standardized residuals plots, which summarize how well the model fits the data overall (i.e. across all observations) - both for the average and subject-specific trajectories. The fitted (i.e. log immune cell counts predicted by the model) values are plotted against the standardized distance between the observed measurement and the predicted value. The blue line is a smoothed average of the standardized residuals as a function of the fitted values, and should ideally be horizontal at 0.

Joint model II

Model formulation

For model II, the time scale was no longer from alloSCT, but instead from time of early low-dose donor lymphocyte infusion (DLI). Therefore, this model was only run among the subset that *did* in fact receive an early low-dose DLI before the occurrence of other competing events. Furthermore, some patients did not have a T-cell measurement on the day of DLI but only a few days prior. For these patients, we used the measurement closest to DLI taken within the last week before DLI as the measurement at time of DLI (time 0).

The longitudinal submodel was again a linear mixed-effects model, where the true underlying log T-cell counts are given by

$$m_i(t) = (\beta_0 + b_{i0}) + \sum_{q=1}^{2}(\beta_q + b_{iq})B_q(t) + \sum_{q=1}^{2}\beta_{q+2}\{B_q(t) \times \text{Donor}_i\} + \beta_5 \text{CMV}_i,$$

with random effects vector $b_i \sim \mathcal{N}(0, D)$. Observations for i^{th} patient are again given by

$$y_{ij} = m_i(t_{ij}) + \epsilon_{ij},$$

where $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ are independent random error terms.

Time was again modelled with restricted cubic splines, but in contrast to model I, we used a single internal knot. The focus on a shorter timespan resulted in a reduced sample size, and fewer measurements per person. For consistency with model I, this average trajectory was allowed to differ across donor types (two-way interaction). In this model, disease risk at baseline was redundant as we ran the model among those having actually received an early low-dose DLI. A fixed effect for patient/donor CMV serostatus was also added to the model. This model comprised both random intercepts b_{i0} and random slopes b_{iq}, assumed to follow normal distributions with mean zero and unstructured covariance matrix.

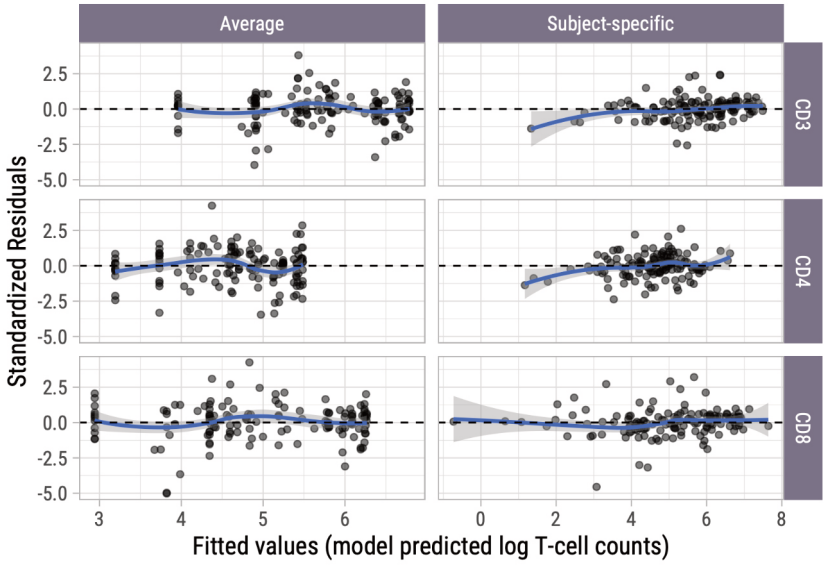
Due to a limited number of events, relapse and other failures were merged into a composite endpoint. The time-to-event submodel was therefore specified as

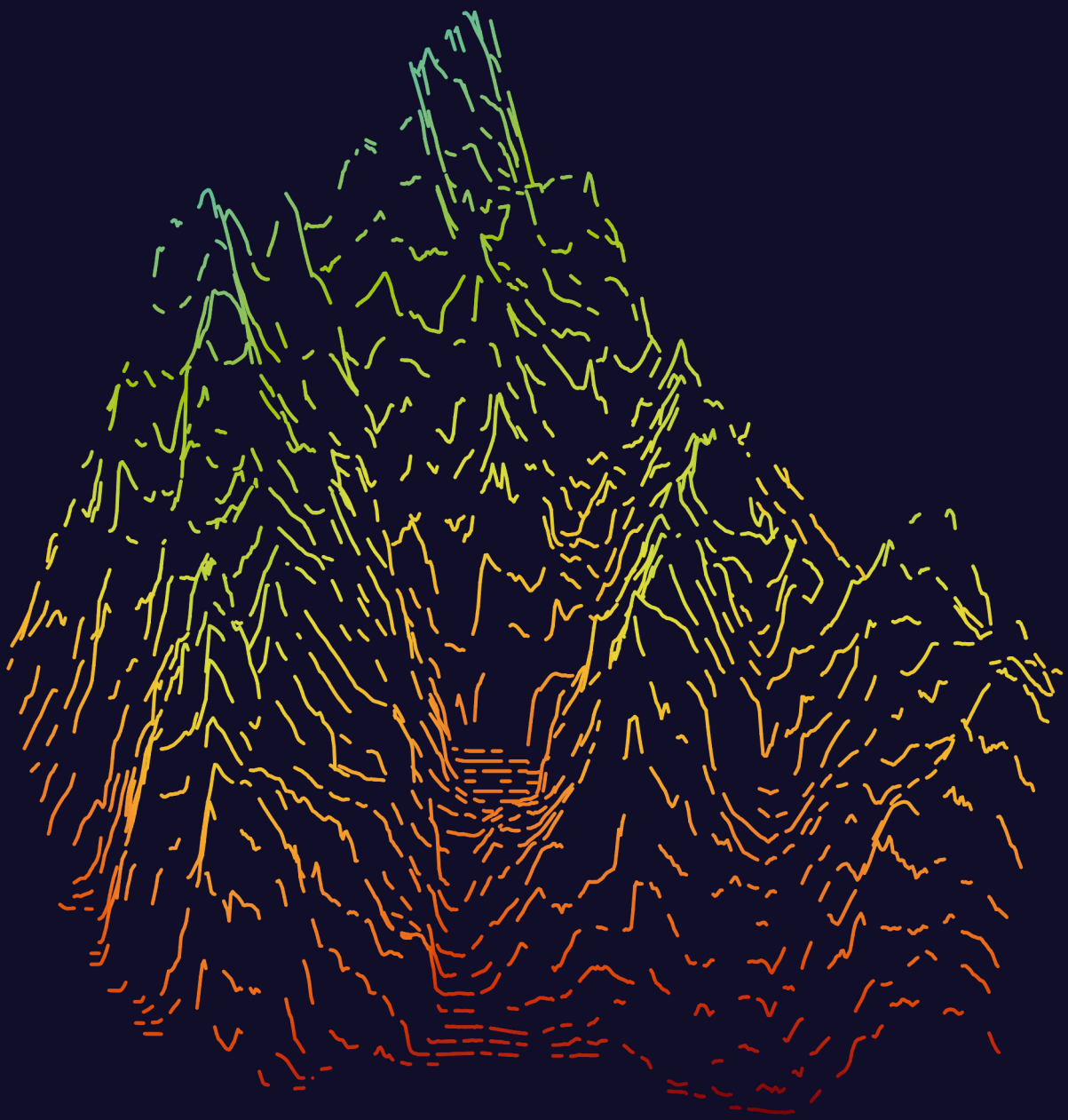
$$h_{1i}(t) = h_{10}(t) \exp\left\{\gamma_{11}\text{Donor}_i + \alpha_1 m_i(t)\right\},$$
$$h_{2i}(t) = h_{20}(t) \exp\left\{\alpha_2 m_i(t)\right\},$$

where the $h_{ki}(t)$ for $k \in \{1, 2\}$ respectively represent the cause-specific hazards of GvHD and the composite of relapse and other failures for subject i. The cause-specific

baseline hazards $h_{k0}(t)$ were approximated on the log scale using cubic B-splines with two internal knots. In this joint model, only the current value parametrization was explored.

Goodness of fit





4

Risk factors for graft-versus-host-disease after donor lymphocyte infusion following T-cell depleted allogeneic stem cell transplantation

Eva A.S. Koster, Peter A. von dem Borne, Peter van Balen, Erik W.A. Marijt, Jennifer M.L. Tjon, Tjeerd J.F. Snijders, Daniëlle van Lammeren, Hendrik Veelken, J.H. Frederik Falkenburg, Constantijn J.M. Halkes, Liesbeth C. de Wreede

Front Immunol. 2024 Mar 13;15:1335341.

ABSTRACT

Unmodified donor lymphocyte infusions (DLI) after allogeneic stem cell transplantation (alloSCT) can boost the beneficial Graft-versus-Leukemia (GvL) effect but may also induce severe Graft-versus-Host-Disease (GvHD). To improve the balance between GvL and GvHD, it is crucial to identify factors that influence the alloreactivity of DLI. We investigated the effects of the presence of patient-derived antigen-presenting cells at time of DLI as estimated by the bone marrow (BM) chimerism status, lymphopenia as measured by the absolute lymphocyte count (ALC) at time of DLI, and the presence of a viral infection (*de novo* or reactivation) close to DLI on the risk of GvHD after DLI. The cohort consisted of patients with acute leukemia or myelodysplastic syndrome who prophylactically or pre-emptively received DLI as standard care after alemtuzumab-based alloSCT. In patients at high risk for relapse, DLI was administered at 3 months after alloSCT (n=88) with a dose of $0.3x10^6$ or $0.15x10^6$ T cells/kg in case of a related or unrelated donor, respectively. All other patients (n=76) received $3x10^6$ or $1.5x10^6$ T cells/kg, respectively, at 6 months after alloSCT. For both DLIs, patients with reduced-intensity conditioning and an unrelated donor had the highest risk of GvHD. For DLI given at three months, viral infection within 1 week before and 2 weeks after DLI was an additional significant risk factor (hazard ratio (HR) 3.66 compared to no viral infection) for GvHD. At six months after alloSCT, viral infections were rare and not associated with GvHD. In contrast, mixed BM chimerism (HR 3.63 for ≥5% mixed chimerism compared to full donor) was an important risk factor for GvHD after DLI given at six months after alloSCT. ALC of $<1000x10^6$/l showed a trend for association with GvHD after this DLI (HR 2.05 compared to $\geq 1000x10^6$/l, 95% confidence interval 0.94-4.45). Furthermore, the data suggested that the presence of a viral infection close to the DLI at three months or ≥5% mixed chimerism at time of the DLI at six months correlated with the severity of GvHD, thereby increasing their negative impact on the current GvHD-relapse-free survival. These data demonstrate that the risk factors for GvHD after DLI depend on the setting of the DLI.

INTRODUCTION

The Graft-versus-Leukemia (GvL) effect of allogeneic hematopoietic stem cell transplantation (alloSCT) results from elimination of persisting malignant hematopoietic cells by donor-derived alloreactive T cells.[1] The GvL effect can provide enduring relapse-free survival but can be accompanied by Graft-versus-Host-Disease (GvHD) when non-hematopoietic cells are targeted.[2] T-cell depletion (TCD) reduces the risk of severe GvHD, but increases the relapse risk by reduction of the GvL effect.[3,4] To boost the GvL effect, unmodified donor lymphocyte infusions (DLI) can be administered after alloSCT.[5] A third of the patients develops clinically relevant GvHD after DLI.[6] Although GvHD is a complication, it does not necessarily mean treatment failure: if GvHD resolves, the patient is unlikely to experience an eventual relapse due to the established concomitant GvL effect.[7,8] The long-term health status of patients with resolved GvHD is comparable to those who did not develop GvHD.[9] Thus, GvHD is a temporary undesired state in contrast to relapse or death as definitive failures. This is captured by the endpoint current GvHD-relapse-free survival (cGRFS) which incorporates recovery from GvHD.[10] Estimation of cGRFS requires advanced statistical methods that can take the end date of GvHD into account, such as multi-state models.[10-12]

Different DLI strategies can be applied to achieve an optimal balance between GvL and GvHD.[13] A reactive strategy is to give only therapeutic DLI to relapsed patients who need a strong alloimmune response to survive. A preemptive strategy administers DLI to patients based on biomarkers that may herald relapse such as mixed chimerism (MC) or minimal residual disease (MRD). In a prophylactic strategy, DLIs are given to all patients without any GvHD independent on additional biomarkers. Several factors known to influence the alloreactivity of DLI are usually taken into account to determine the DLI dose.[14] First, DLIs with higher T-cell doses induce more GvHD and GvL.[15] Second, patients with an unrelated donor (UD) or HLA-mismatched donor have more allo-antigens that can provoke an alloimmune response and often receive a lower dose than patients with an HLA-matched related donor (RD). Third, the DLI dose is also dependent on the timing after alloSCT, since the alloreactive potential of DLI decreases over time due to changes in the host environment.[16,17] Early after transplantation, professional antigen-presenting cells (APCs) required to activate naïve T cells are still patient-derived and therefore highly capable of activating donor-derived alloreactive T cells. Tissue damage by the conditioning regimen and infections, which occur relatively frequently during the first months after alloSCT, leads to a pro-inflammatory environment that promotes activation of alloreactive T cells.[18,19] Moreover, the conditioning-induced lymphopenia stimulates the outgrowth of (alloreactive) T cells by homeostatic proliferation and promotes activation of these T cells.[20,21] Over time after alloSCT, tissue damage is repaired, patient-derived professional APCs are replaced by donor-derived APCs, lymphopenia disappears, infections become rare, and higher T-cell doses are needed to induce a sufficient GvL effect after DLI.

Despite dose adjustments based on timing and donor type, the effect of a single DLI is highly variable between patients, ranging from patients not responding at all to patients succumbing to severe GvHD. To avoid excessive toxicity in the prophylactic or preemptive setting, it is crucial to better understand which factors influence the efficacy and toxicity of DLI. Since development of clinically relevant GvHD represents the

clearest indicator for induction of alloreactivity after DLI, we aimed to identify risk factors for GvHD after prophylactic or preemptive DLI following alemtuzumab-based TCD alloSCT. Focusing on conditions that promote T-cell activation, we investigated the effects of the presence of patient-derived APCs in the bone marrow (BM) as measured by the BM chimerism level at time of DLI, the presence of lymphopenia as measured by the absolute lymphocyte count (ALC) at time of DLI, and the occurrence of viral infections (i.e., *de novo* infections or reactivations) close to DLI. We also investigated the impact of potential risk factors on the course of GvHD: GvHD only requiring short-term therapeutic systemic immunosuppression (tIS), GvHD requiring long-term tIS, or lethal GvHD. To assess their clinical relevance, we transformed these effects into cGRFS probabilities.

METHODS

Study population

This retrospective study included all adult patients with acute myeloid leukemia, acute lymphoblastic leukemia or myelodysplastic syndrome in complete morphologic remission who received an alloSCT from a 10/10 HLA-matched donor using a standard conditioning and TCD protocol[22-24] at Leiden University Medical Center (LUMC, Leiden, The Netherlands) between 2005 and 2019. Patients scheduled to receive azacitidine or daratumumab (in 1 patient with CD38 positive acute lymphoblastic leukemia) as pharmacological maintenance therapy after alloSCT were excluded. All patients signed informed consent for data collection and analysis. Data were analyzed as of July 2021.

Transplantation and DLI protocol

The protocols for the myeloablative and reduced-intensity conditioning regimens (MAC and RIC, respectively), TCD and GvHD prophylaxis are described in the Supplemental Methods. The dose of unmodified preemptive and prophylactic DLI was based on donor type and timing after alloSCT. DLI at 3 months after alloSCT contained low doses of $0.3x10^6$ and $0.15x10^6$ T cells/kg in case of RD and UD, respectively. DLI at 6 months after alloSCT contained $3x10^6$ and $1.5x10^6$ T cells/kg, respectively. All patients could receive preemptive DLI in case of MC or MRD positivity, starting from 3 months after alloSCT. Subsequent preemptive DLI could be given in escalating doses with at least 3 months between DLI. Since May 2010, patients who were considered to have a high risk of relapse or who received the FLAMSA regimen received prophylactic low-dose DLI at 3 months. In addition, all eligible patients without any relapse or GvHD requiring systemic treatment received prophylactic DLI at 6 months after alloSCT regardless of chimerism or MRD status. Furthermore, selected patients could receive modified T-cell products within several clinical trials.

Definitions of clinical events and DLI cohorts

Relapse was defined as recurrence of at least 5% blasts on cytomorphologic BM examination, at least 1% blasts in the peripheral blood or the presence of extramedullary disease. Graft failure was defined as the occurrence of >95% mixed BM chimerism in all

lineages tested or refractory granulopenia (granulocyte count $<0.5x10^9$/l) in the absence of relapse and ongoing myelotoxic medication. To have a clear definition of clinically relevant GvHD with exact starting and stopping dates, essential for statistical modeling, we considered administration of tIS for acute or chronic GvHD instead of the grading of GvHD. For the analyses, we only considered tIS which was given for at least 14 days or until death, or which was stopped within 1 week before death from GvHD. In the latter case, the last week before death was added to the tIS episode. If a patient stopped tIS but had to restart tIS again within 2 months due to the recurrence of GvHD, both tIS episodes were combined into one episode. cGRFS was defined as the probability of being alive without relapse and currently not using any tIS for GvHD.

To investigate the clinical outcomes after DLI, two subcohorts were defined. The low-dose 3-month DLI cohort included all patients who were scheduled to receive a prophylactic or preemptive low-dose DLI at 3 months after alloSCT and received it within 6 months after alloSCT without any prior relapse, tIS for GvHD or cellular intervention besides infusion of virus-specific T cells. The 6-month DLI cohort consisted of all patients who were scheduled to receive a prophylactic or preemptive 6-month DLI as first DLI and received it within 9 months after alloSCT without any prior relapse, tIS for GvHD or cellular intervention besides infusion of virus-specific T cells. Both subcohorts were thus independent.

BM chimerism, ALC and viral infections

The methods for measuring BM chimerism, ALC and viral infections are described in the Supplemental Methods. The BM chimerism level was used as a measurement of the presence of patient-derived APCs in the BM at time of DLI. Three chimerism categories were defined: full donor chimerism (FDC; no detectable patient material), low MC (detectable patient material but <5%), and high MC (≥5% patient material).

Lymphopenia was defined as ALC $<1000x10^6$/l, the lower limit of normal in our laboratory. For patients receiving the 3-month DLI, three ALC categories were defined: ALC $<500x10^6$/l, ALC between 500 and $999x10^6$/l and ALC $\geq1000x10^6$/l. For patients who received the 6-month DLI as first DLI, only two categories were used, <1000 and $\geq1000x10^6$/l, since most patients had ALC $\geq500x10^6$/l at that time.

All viral infections (*de novo* or reactivation) confirmed by PCR that occurred within 1 week before and 8 weeks after DLI without any prior relapse, second DLI or tIS were considered.

Statistical analyses

Follow-up after alloSCT was quantified using the reversed Kaplan-Meier method.[25] The cumulative incidence of tIS after the first DLI (DLI1) was estimated in a competing risks model starting at time of DLI1 with start of tIS as the event of interest and relapse, death and second DLI (DLI2) as competing events. The cumulative incidence of death during treatment for GvHD from start of tIS was estimated in a competing risks model starting at time of start tIS after DLI1 with death as the event of interest and relapse, stop tIS and DLI2 as competing events.

To investigate risk factors for requiring tIS for GvHD and death during tIS and to

estimate cGRFS after DLI, several Markov time-inhomogeneous multi-state models were constructed. See the Supplemental Methods for a brief explanation of the methodology of multi-state modelling. The structure of the main multi-state model is shown in Figure 1. The model used DLI1 as the starting state and time and considered the following events: death, relapse, start and stop of tIS for GvHD, and DLI2. Separate states were used for events after DLI1 and for events after second DLI (e.g., 'relapse after DLI1' and 'relapse after DLI2'). The probability of cGRFS over time was calculated as the sum of the probabilities of being in one of the relevant states in the multi-state model (i.e., 'DLI1', 'stop tIS after DLI1', 'DLI2' and 'stop tIS after DLI2'). The probabilities of death after start of tIS, being alive with clinically GvHD, relapse-free survival (RFS) and overall survival (OS) were calculated analogously. The outcomes after the low-dose 3-month DLI and the 6-month DLI were analyzed using two separate versions of this

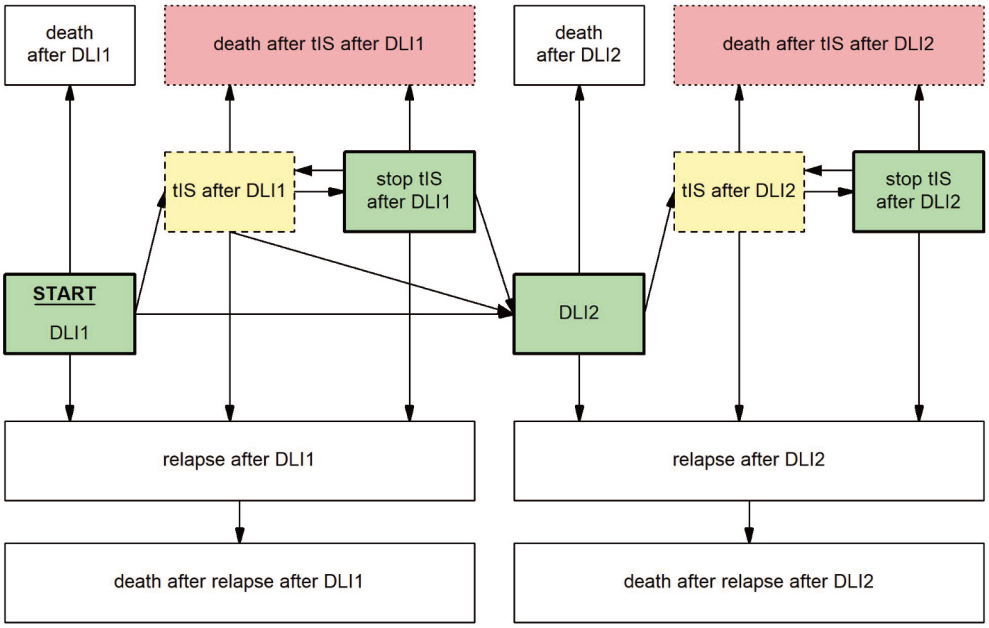


Figure 1. Multi-state model to evaluate the development and outcome of GvHD and other clinical events after DLI. Boxes represent states, arrows represent transitions. Starting state and time was DLI1. From here, patients could move to the state 'relapse after DLI1' at time of relapse, 'death after DLI1' at time of death, 'tIS after DLI1' at time of the start of tIS for GvHD and 'DLI2' at time of the administration of a second DLI, whichever occurred first. From the state 'relapse after DLI1' patients could only enter the state 'death after relapse after DLI1'. From the state 'tIS after DLI1' patients could move to 'stop tIS after DLI1' at time of stop of all tIS, 'relapse after DLI1' at time of relapse, 'death after tIS after DLI1' at time of death or 'DLI2' at time of the administration of a second DLI, whichever occurred first. From the state 'stop tIS after DLI1' patients could return to 'tIS after DLI1' when patients had to restart tIS for recurrent GvHD, 'relapse after DLI1' at time of relapse, 'death after tIS after DLI1' at time of death or 'DLI2' at time of the administration of a second DLI, whichever occurred first. After DLI2, similar states were constructed, except that any further DLIs were ignored. The cGRFS is the sum of the probabilities of all green (thick border) states, the probability of being alive with GvHD the sum of all yellow (dashed border) states, the probability of death after start of tIS for GvHD the sum of all red (dotted border) states, the RFS the sum of all green (thick border) and yellow (dashed border) states, and the OS the sum of all non-death states. For these summarizing measures, no distinction was made between states after the first DLI or after multiple DLIs.

model, omitting all transitions and states that were not used by the included patients (Supplemental Figures 1 and 2).

The effects of BM chimerism, ALC and viral infections on the risk of clinically relevant GvHD after DLI were estimated using separate multivariable Cox proportional hazards regression models for the transition from 'DLI1' to 'tIS after DLI1': 3 models were fitted for the low-dose 3-month DLI and two for the 6-month DLI (only chimerism and ALC). Since donor type and conditioning/TCD regimen have been recognized as important factors for GvHD after DLI[16], conditioning/donor type (MAC UD, RIC RD and RIC UD vs MAC RD) was included in all models, while BM chimerism (low MC and high MC vs FDC), ALC ($<500x10^6$/l and 500-$999x10^6$/l vs $\geq 1000x10^6$/l for the 3-month DLI or $<1000x10^6$/l vs $\geq 1000x10^6$/l for the 6-month DLI), or viral infection were added as the only other covariate per model. Viral infection was time-varying: patients could start as having no viral infection or as having an early viral infection if they had a viral infection during the last week before DLI. After DLI, the variable could change to 'early viral infection' at time of the first viral infection if this occurred within 2 weeks after DLI or to 'late-onset viral infection' at time of the first viral infection occurring beyond 2 weeks after DLI.

To identify risk factors for death during treatment for GvHD, univariable Cox proportional hazards regression models were fitted for the transition from 'tIS after DLI1' to 'death after tIS after DLI1' with either patient age at time of alloSCT or with the presence of early viral infection (3-month DLI) or high MC (6-month DLI). Two-sided p-values <0.05 were considered statistically significant for all Cox models. All models were based on complete cases only: patients with missing values for the included covariates were excluded.

To illustrate the impact of early viral infections on the outcome after the low-dose 3-month DLI, an extended version of the multi-state model was constructed with two starting states: 'DLI1 without early viral infection' for patients without any viral infection during the last week before DLI and 'DLI1 with early viral infection' for patients with a viral infection during this period (Supplemental Figure 3). To evaluate the impact of the identified transition-specific risk factors on the probability of cGRFS, the probability of being alive with GvHD, and the probability of death after start of tIS after the 6-month DLI, the Cox models for the two transitions were integrated as components in a multi-state model. This model was used to predict the outcomes after the 6-month DLI for reference patients with different baseline characteristics.

Software

All analyses were performed in R version 4.3.1 using the packages survival[26], prodlim[27], cmprsk[28], mstate[29], ggplot2[30], and ComplexUpset[31].

RESULTS

Cohort

388 patients were included in this study (Supplemental Table 1). Median follow-up after alloSCT was 76 months (interquartile range 32-110). 88 patients received the low-dose

3-month DLI prophylactically or pre-emptively at a median of 3.2 months after alloSCT (range 2.7-5.2) and 76 the 6-month DLI as first DLI at a median of 6.3 months after alloSCT (range 4.8-8.9; Table 1). 79 (20%) patients could not receive any DLI because of early relapse (n=44), death (n=23), or graft failure (n=12; Supplemental Figure 4). 66 (17%) other patients developed clinically relevant GvHD after alloSCT and therefore were not eligible for DLI. 42 patients received a modified T-cell product as part of a clinical study, and 9 received a DLI not according to our standard prophylactic/preemptive DLI protocol (different cell dose (n=6), DLI for a viral infection (n=2) or DLI in combination with interferon (n=1)). The remaining 28 patients did not receive any DLI within the first 9 months after alloSCT because of alloSCT before May 2010 (n=12), (temporary) donor unavailability (n=3) or physician's decision (n=13).

Similar incidences of GvHD after low-dose 3-month DLI and 6-month DLI

The 3-month cumulative incidence of clinically relevant GvHD was 28% (95%-CI 20-40) after the low-dose 3-month DLI and 30% (95%-CI 22-43) after the 6-month DLI.

	Low-dose 3-month DLI (N = 88)	**6-month DLI (N = 76)**
Age at alloSCT (years)		
median (range)	58 (18-74)	57 (19-76)
Disease		
AML	59 (67%)	56 (74%)
ALL	23 (26%)	9 (12%)
MDS	6 (7%)	11 (14%)
Conditioning		
MAC: Cyclo/TBI	35 (40%)	33 (43%)
MAC: Cyclo/Bu	1 (1%)	1 (1%)
RIC: Flu/Bu*	38 (43%)	42 (55%)
RIC: Flu/Bu/Ara-C/Amsa	14 (16%)	0
Donor		
RD	39 (44%)	30 (39%)
UD	49 (56%)	46 (61%)
Graft source		
G-CSF mobilized PBSC	84 (95%)	70 (92%)
BM	4 (5%)	6 (8%)
CMV serostatus patient/donor		
+/+	43 (49%)	33 (43%)
+/-	13 (15%)	12 (16%)
-/+	6 (7%)	4 (5%)
-/-	26 (30%)	27 (36%)
EBV serostatus patient/donor		
+/+	78 (89%)	59 (78%)
+/-	6 (7%)	7 (9%)
+/unknown	0	4 (5%)
-/+	3 (3%)	6 (8%)
-/-	1 (1%)	0

Table continues on next page.

The probability of death during tIS after one DLI was 15% (95%-CI 9-24) and 16% (95%-CI 9-27) at 12 months after the 3- and 6-month DLI, respectively (Supplemental Figures 5 and 6). Figures 2 and 3 show how the state probabilities add up to the overall survival, relapse-free survival, and cGRFS probabilities. For example, the cGRFS decreased during the first months after DLI but later increased as patients with GvHD could stop their tIS after the GvHD was resolved. Notably, none of the patients with GvHD after DLI relapsed, demonstrating the concomitant GvL effect. 1- and 5-year cGRFS probabilities were 55% (95%-CI 45-66) and 48% (95%-CI 38-61) after 3-month DLI and 57% (95%-CI 46-69) and 67% (95%-CI 57-79) after 6-month DLI, respectively. Together, these data show that the tenfold dose difference effectively equalized the

Main indication of first DLI		
ALL: t(9;22)	11 (12%)	-
ALL: hypodiploidy, complex karyotype, or t(4;11)	3 (3%)	-
ALL: high white blood cell count at diagnosis	4 (5%)	-
ALL: no CR1	2 (2%)	-
AML: monosomal karyotype	10 (11%)	-
AML: complex karyotype	1 (1%)	
AML/MDS: EV1 overexpression	15 (17%)	-
AML: ASXL mutation	2 (2%)	
AML: FLT3 mutation	1 (1%)	-
AML/MDS: FLAMSA regimen	14 (16%)	
AML: progression during remission-induction	1 (1%)	-
AML/MDS: no intensive treatment or no consolidation	4 (5%)	-
AML/MDS: persisting CMML	1 (1%)	-
MRD+ at time of alloSCT	11 (12%)	-
Preemptive for MC	8 (9%)	34 (45%)
Standard prophylactic DLI	-	42 (55%)
BM chimerism at time of first DLI		
FDC	28 (33%)	25 (34%)
Low MC (1-4% mixed chimerism)	32 (38%)	30 (41%)
High MC (≥5% mixed chimerism)	24 (29%)	19 (26%)
Unknown	4	2
ALC at time of first DLI (x10^6/l)		
≥1000	41 (47%)	45 (61%)
500-999	29 (33%)	20 (27%)
<500	17 (20%)	9 (12%)
Unknown	1	2

Table 1. Baseline characteristics of the patients who received either a low-dose 3-month DLI or 6-month DLI as first DLI. Characteristics are given at time of alloSCT unless otherwise indicated. DLI, donor lymphocyte infusion; alloSCT, allogeneic stem cell transplantation; AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; MDS, myelodysplastic syndrome; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; Cyclo, cyclophosphamide; TBI, total body irradiation; Bu, busulfan; Flu, fludarabine; Ara-C, cytarabine; Amsa, amsacrine; RD, related donor; UD, unrelated donor; G-CSF, granulocyte-colony stimulation factor; PBSC, peripheral blood stem cells; BM, bone marrow; CMV, cytomegalovirus; EBV, Epstein-Barr virus; CR, complete morphological remission; CMML, chronic myelomonocytic leukemia; MRD, minimal residual disease; MC, mixed chimerism; FDC, full donor chimerism; ALC, absolute lymphocyte count. *One patient had not received a second consolidation course before transplant and received 2 days cyclophosphamide 750 mg/m^2 intravenously additionally to the conditioning regimen.

GvHD risk between low-dose 3-month DLI and 6-month DLI. Because 16% of patients died within 1 year after DLI during treatment for GvHD (Figures 2 and 3), we investigated risk factors for the development of clinically relevant GvHD and the occurrence of death during tIS.

Viral infections close to low-dose 3-month DLI increase the risk of GvHD after this DLI

First, we analyzed the low-dose 3-month DLI. To investigate whether the presence of patient-derived APCs in the BM increased the risk of GvHD after this DLI, we examined the chimerism model (Figure 4A). RIC patients with an UD had a hazard ratio (HR) of 3.2 (95%-CI 1.1-9.1) for developing GvHD compared to MAC RD patients. However, there was no significant effect of chimerism (p-values 0.9 and 0.8 for low and high MC compared to FDC, respectively) on the risk of clinically relevant GvHD after this DLI. To investigate whether lymphopenia increased the risk of GvHD after the 3-month DLI, we examined the ALC model (Figure 4B). Again, RIC UD was a significant

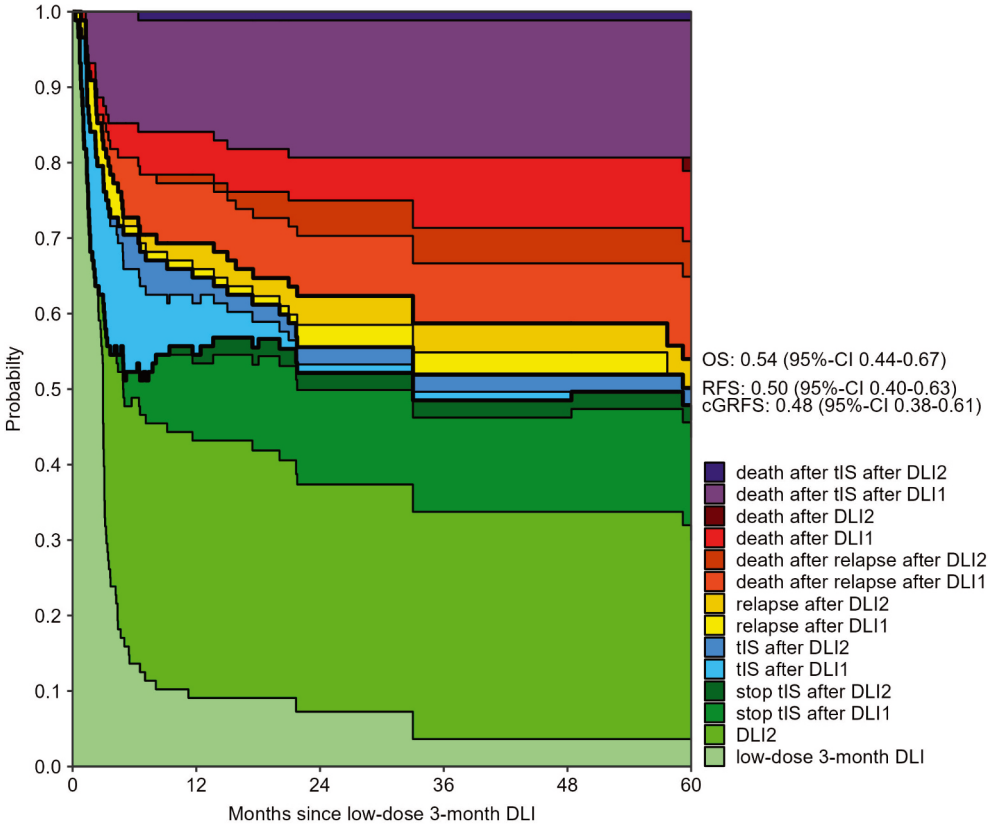


Figure 2. Outcomes after low-dose 3-month DLI. Stacked transition probabilities from state DLI1 (low-dose 3-month DLI) estimated in the non-parametric model in Supplemental Figure 1. The difference between two adjacent curves represents the probability of being in the corresponding state. 39 patients reached the second DLI as planned. Bold lines show the overall survival (OS), relapse-free survival (RFS) and current GvHD-relapse-free survival (cGRFS), of which the 5-year probabilities with 95%-CI are stated next to the figure.

risk factor while ALC showed no significant effect on GvHD after DLI (p-values 0.9 and 0.6 for ALC 500-999x10^6/l and <500x10^6/l compared to ≥1000x10^6/l, respectively). We then investigated the correlation between viral infections close to the 3-month DLI and the development of GvHD after DLI. 34 of the 88 patients with a 3-month DLI had a viral infection within the last week before and first 8 weeks after DLI: 28 had an early viral infection (25 before or at time of DLI and 3 within 2 weeks after DLI) and 6 a late-onset viral infection (>2 weeks after DLI). Most common pathogens were cytomegalovirus (CMV; n=15), adenovirus (n=7) and Epstein-Barr virus (EBV; n=5; Supplemental Figure 7A). The model with viral infection revealed that patients with an early viral infection had a HR of 3.7 (95%-CI 1.7-7.9) for developing clinically relevant GvHD compared to those without any viral infection (Figure 4C). Patients with a late-onset viral infection did not have a higher risk of GvHD (p-value 0.7).

Since the ALC at time of the low-dose 3-month DLI was higher in patients with a viral infection (Supplemental Figure 8), viral infections may have confounded the correlation between ALC and GvHD. Therefore, to explore whether ALC is a risk factor for GvHD

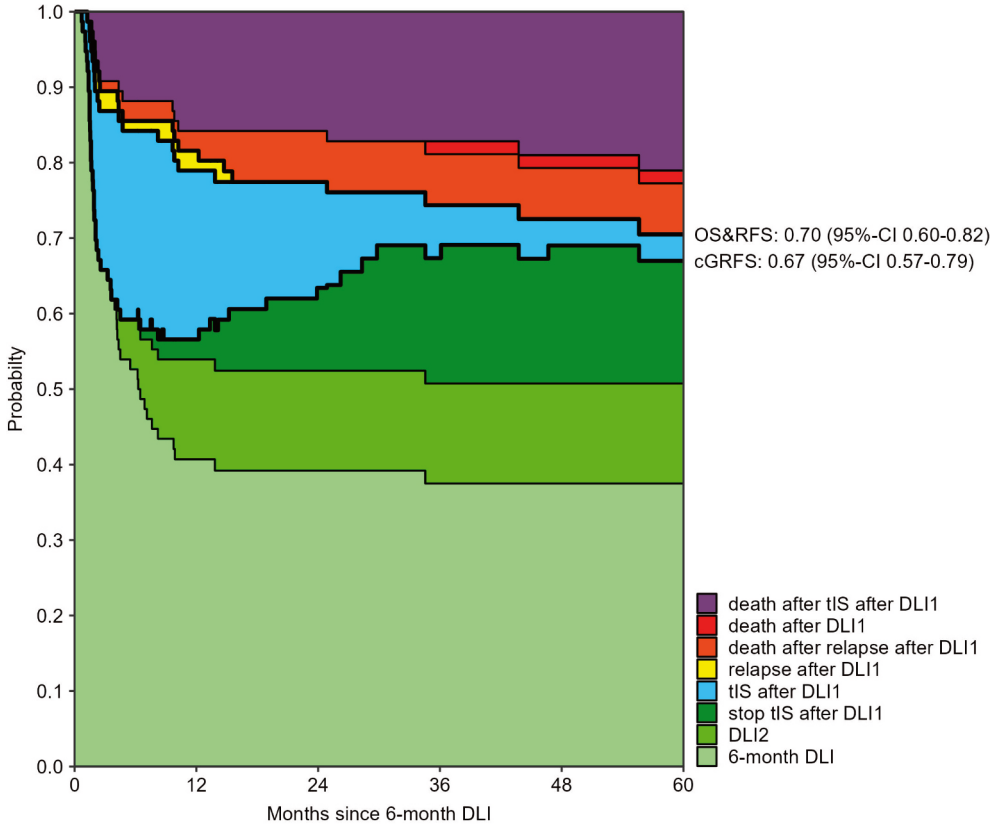


Figure 3. Outcomes after 6-month DLI. Stacked transition probabilities from state DLI1 (6-month DLI) estimated in the non-parametric model in Supplemental Figure 2. The difference between two adjacent curves represents the probability of being in the corresponding state. Nine patients required a second DLI because of MC. The legend only shows the states which were occupied within 5 years after the 6-month DLI. Bold lines show the overall survival (OS), relapse-free survival (RFS) and current GvHD-relapse-free survival (cGRFS), of which the 5-year probabilities with 95%-CI are stated next to the figure.

in the absence of viral infections, we compared the cumulative incidences of tIS for GvHD between ALC <1000x10^6/l and ≥1000x10^6/l in the 63 patients without any viral infection during the last week before the 3-month DLI. As we did not observe a significant difference (Supplemental Figure 9), there was no clear indication that viral infection acted as confounding factor. Together, these data show that viral infections close to the low-dose 3-month DLI increased the alloreactivity of this DLI leading to significantly more clinically relevant GvHD.

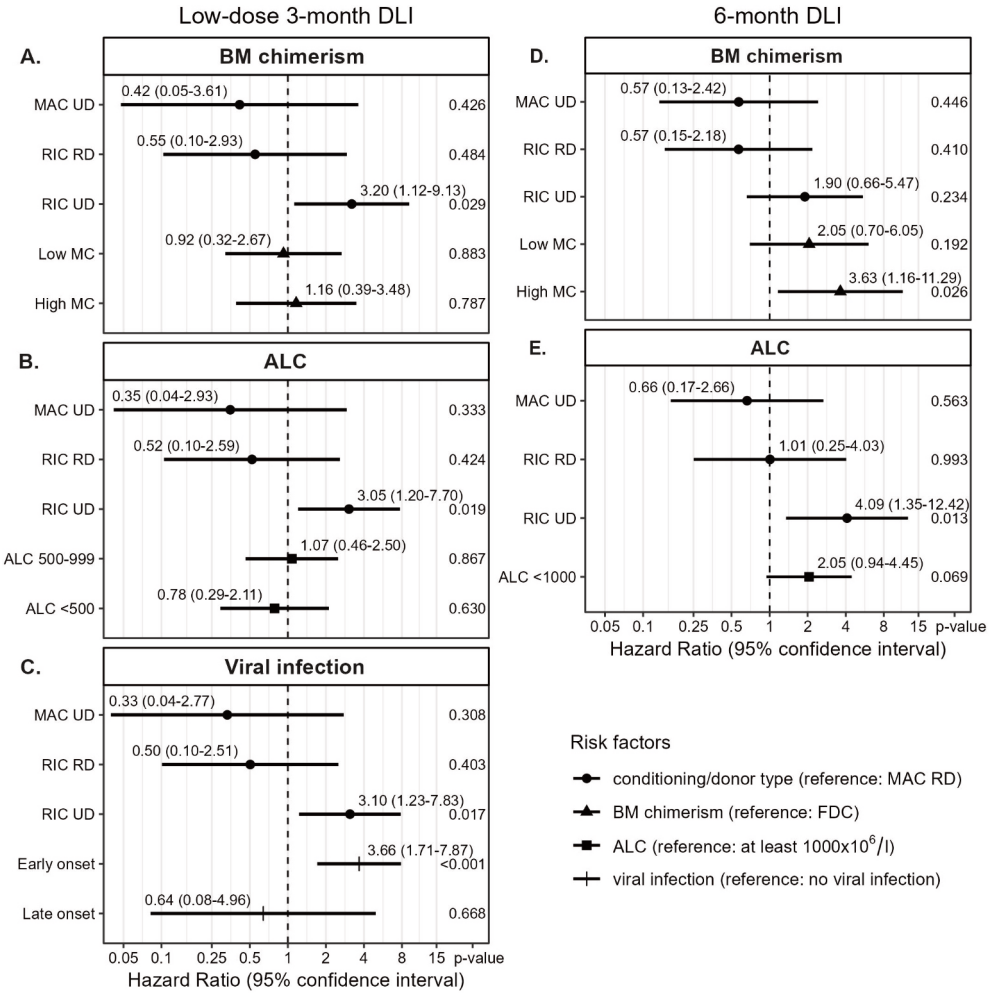


Figure 4. Cox proportional hazards models for the transition from first DLI to requiring tIS for GvHD (see Figure 1). Based on complete case analysis (A: n=84, B: n=87, C: n=88, D and E: n=74). Viral infection was treated as a time-varying covariate. DLI, donor lymphocyte infusion; BM, bone marrow; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; UD, unrelated donor; RD, related donor; low MC, 1-4% mixed chimerism; high MC, ≥5% mixed chimerism; FDC, full donor chimerism (no patient material detectable); ALC, absolute lymphocyte count (x10^6/l)

Mixed BM chimerism and lymphopenia increase the risk of GvHD after the 6-month DLI

We then investigated which risk factors were associated with the alloreactivity of the 6-month DLI. Viral infections were uncommon at the time of this DLI: of the 76 patients receiving this DLI, only 11 had a viral infection (3 early and 8 late-onset), most often EBV (n=3; Supplemental Figure 7B). The presence of high MC in the BM at time of DLI was a strong predictor for GvHD with a HR of 3.6 (95%-CI 1.2-11.3) compared to FDC, while patients with low MC had a nonsignificant higher risk of GvHD (HR 2.1, 95%-CI 0.7-6.1, p-value 0.19, Figure 4D). In the ALC model (Figure 4E), RIC UD was a significant risk factor for GvHD (HR 4.1, 95%-CI 1.3-12.4 compared to MAC RD). Additionally, a trend was observed for higher GvHD risk in lymphopenic patients compared to ALC $\geq 1000 \times 10^6$/l (HR 2.1, 95%-CI 0.9-4.5, p-value 0.07). Together, these data show for both the low-dose 3-month DLI and the 6-month DLI, with 50% dose reduction in case of an UD, comparable risks of GvHD between patients with RD and UD after MAC but not RIC. The data indicate that mixed BM chimerism increased the risk of clinically relevant GvHD after the 6-month DLI, and suggest a similar effect of lymphopenia.

Risk factors for death during treatment for GvHD after DLI

To identify risk factors for death during tIS for GvHD (Supplemental Figure 10), we first investigated the effect of patient age. As expected, older patients seemed to have a higher risk of dying from severe GvHD after the 6-month DLI (HR 2.1 per decade, 95%-CI 0.9-5.1, p-value 0.10). Remarkably, we did not observe this association after the low-dose 3-month DLI (p-value 0.7).

Next we investigated whether the main risk factors for clinically relevant GvHD also correlated with the risk of death among those who required treatment for GvHD. For the low-dose 3-month DLI, we considered the presence of an early viral infection. We observed a nonsignificant increase in the risk of dying during tIS for GvHD for patients with an early viral infection compared to those without an early viral infection close to DLI (HR 1.8, 95%-CI 0.6-5.6, p-value 0.28, Supplemental Figure 11). For the 6-month DLI we considered the presence of high mixed BM chimerism at time of DLI. Patients with high MC had a nonsignificant higher risk of death during tIS for GvHD compared to those with GvHD who had FDC or low MC at time of DLI (HR 2.0, 95%-CI 0.6-6.4, p-value 0.23, Supplemental Figure 12). In conclusion, among those who required tIS for GvHD, older patients had a higher risk of dying during treatment after the 6-month but not the low-dose 3-month DLI. We did not observe significant associations between the risk of death during tIS and BM chimerism or viral infections. However, only one of the 53 patients with FDC at time of the low-dose 3-month DLI or 6-month DLI developed lethal GvHD.

Impact of early viral infection and mixed BM chimerism on the cGRFS after the low-dose 3-month DLI and 6-month DLI

The probability of having clinically relevant GvHD at 6 months after the 3-month DLI was 15% (95%-CI 9-26) for the patients without any viral infection during the last week before DLI compared to 25% (95%-CI 14-46) for the patients with a viral infection

(Figure 5). The probability of death after start of tIS was 8% (95%-CI 4-17) compared to 32% (95%-CI 19-55), respectively. The cGRFS was 61% (95%-CI 50-73) and 31% (95%-CI 19-52), respectively.

For a MAC patient receiving a 6-month DLI from a RD, the predicted probability of having clinically relevant GvHD at 6 months after DLI was 14% (95%-CI 5-44) if the patient had FDC compared to 30% (95%-CI 11-80) if the patient had high MC, respectively (Figure 6). The probability of death after start of tIS was 4% (95%-CI 1-16) and 23% (95%-CI 9-58), respectively. The cGRFS for these reference patients was 77% (95%-CI 60-98) and 44% (95%-CI 19-100) at 6 months after DLI, respectively.

DISCUSSION

In this retrospective study we investigated the outcomes after prophylactic and preemptive DLI following alemtuzumab-based TCD alloSCT. The tenfold dose difference between the 3- and 6-month DLI resulted in comparable risks of GvHD. For both DLIs, the 50% dose reduction in case of an UD sufficed for patients with MAC but not RIC. We demonstrate that the risk factors for GvHD after DLI depend on the setting of the DLI: at time of the 3-month DLI, the occurrence of viral infections played a major role, while for the 6-month DLI the presence of high MC in the BM was an important risk factor. The strong impact of both factors on cGRFS underlines the clinical relevance of these findings. Additionally, we observed trends for higher GvHD risk in patients with low MC or lymphopenia at time of the 6-month DLI. The very low

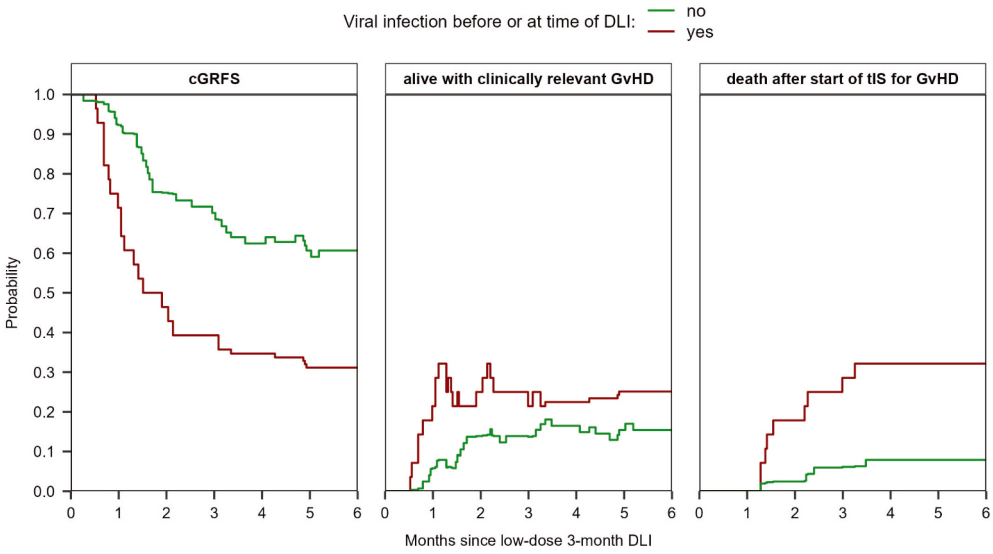


Figure 5. Estimated probabilities of cGRFS, being alive with clinically relevant GvHD, and of death after start of tIS for GvHD after the low-dose 3-month DLI based on the viral status at time of DLI (viral infection during the last week before DLI (n=25) or no viral infection until DLI (n=63)). The estimates are based on the non-parametric multi-state model in Supplemental Figure 3 which has two starting states ('DLI1 without early viral infection' and 'DLI1 with early viral infection'). See Supplemental Figure 13 for the probabilities of all states separately.

risk of lethal GvHD for patients with FDC at time of either DLI provides further evidence for the important role of patient-derived APCs and demonstrates the safety of DLI in these patients, consistent with the matched-pair analysis by Schmid et al.[32]

Viral infection and the concomitant antiviral immune response lead to tissue damage and upregulation of HLA class II expression by non-hematopoietic cells, and induce a pro-inflammatory environment promoting activation of professional APCs and immune cells. Miller et al. showed that the occurrence of any infection (bacterial, viral or fungal) increased the risk of acute GvHD after alloSCT.[33] We only considered viral infections, since these were most common in the relevant time period and most of the patients with a bacterial or fungal infection had a viral infection at the same time (data not shown). Other studies have reported associations specifically between CMV and GvHD.[19,34,35] Previously, we demonstrated activation of alloreactive HLA-DP1-specific CD4+ T cells leading to GvHD in two patients with a CMV reactivation after a CD4+ T-cell infusion from an HLA-DP1 mismatched donor.[19] Since about 80% of the patients with a 10/10 HLA-matched unrelated donor are HLA-DP mismatched[36,37] and CMV was the most common pathogen, this mechanism could play a role in our cohort. Due to the limited number of events, we could not differentiate between the different viral pathogens.

While the role of patient-derived professional APCs in the induction of alloreactivity has been clearly demonstrated in mice[38-41], results of human studies are conflicting.[42-48] This may be due to the cell subsets used for the chimerism measurement, possible bias by overrepresentation of patients with multiple DLIs, and the clinical setting. For example, Bar et al.[48] did not observe a significant correlation between BM chimerism and GvHD

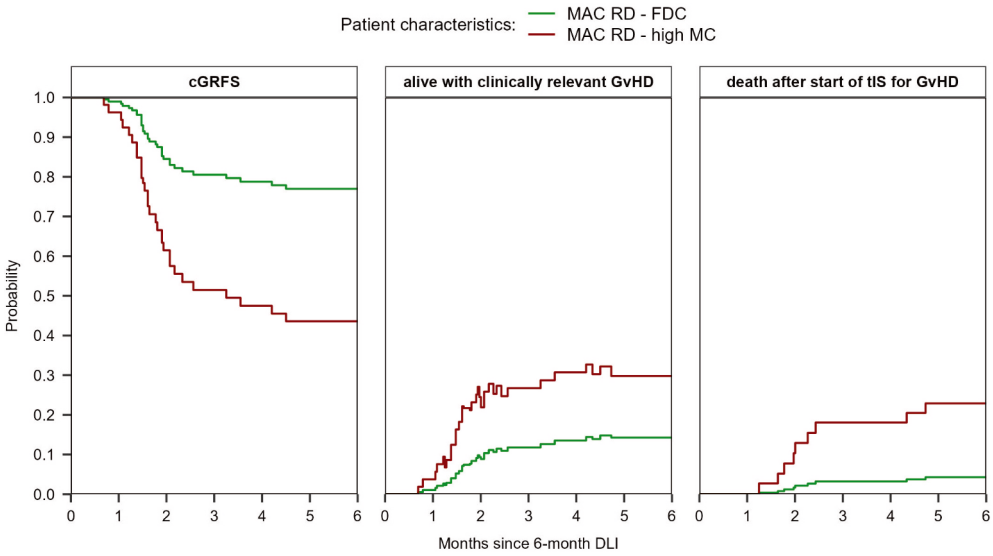


Figure 6. Prediction of cGRFS, being alive with clinically relevant GvHD, and of death after start of tIS for GvHD after the 6-month DLI for reference patients with different characteristics. The prediction is based on the multi-state model in Supplemental Figure 2 with semi-parametric transition-specific proportional hazards models with BM chimerism and conditioning/donor combination as covariates for the transition from 'DLI1' to 'tIS for GvHD after DLI1' and BM chimerism (high MC vs other) for the transition from 'tIS for GvHD after DLI1' to 'death after tIS after DLI1'. No covariates were assessed for the other transitions of the model.

(HR 1.26, p-value 0.46). They however analyzed therapeutic DLI in patients who often received disease-specific treatment or cytoreduction before DLI, which most likely resulted in a more pro-inflammatory environment at time of DLI. Under these circumstances, non-hematopoietic tissues from the patient express HLA class II molecules and can act as APCs to activate donor-derived alloreactive T cells.[49,50] The presence of a pro-inflammatory environment may also be an explanation for the absent association between BM chimerism and GvHD after the 3-month DLI, as tissue damage from the conditioning and recent viral infections may still be present. Another explanation may lie in the persistence of professional patient-derived APCs in the peripheral tissues at that time. The replacement of these APCs lags behind the donor-derived BM repopulation, as long as GvHD and severe inflammation as caused by myeloablative conditioning are absent.[51-53]

The relation between lymphopenia and alloreactivity of DLI has mostly been investigated in relapsed patients who often received (lymphodepleting) chemotherapy before DLI.[44,48,54] In this context, the effects of tissue damage and APC activation interfere with estimating the effect of the lymphopenia itself on the risk of GvHD. In our setting, patients received their DLI in the absence of relapse, tissue damage and chemotherapy. Here, we observed a trend for higher GvHD risk in lymphopenic patients at time of the 6-month DLI, but not at time of the 3-month DLI.

Multi-state modeling allowed us to not only estimate the effects of risk factors on the development of GvHD and death during treatment, but also assess the impact of these factors on the probabilities of different outcomes after DLI while taking into account the hazards of all clinical events. This is a major advantage compared to less advanced statistical methods since these probabilities are more relevant for patients than HRs. Multi-state models can capture recovery after GvHD and thereby model the current GvHD burden over time, which makes cGRFS a better estimate of treatment success than GvHD-relapse-free survival.[11,12] In 2016, we introduced the endpoint treatment success, which equals cGRFS.[10] During the last years, cGRFS and current immunosuppression-relapse-free survival have become more popular as outcome measures.[11,12,55-58] However, to our knowledge, we are the first who have applied semi-parametric multi-state modeling in this context. For this, detailed data collection regarding posttransplant events and interventions as performed in this study is essential.

Our observations may eventually lead to refinement of the DLI strategy. In the prophylactic or preemptive setting, there is room to lower the initial DLI dose, delay the DLI or start immunosuppressive treatment on early signs of GvHD based on the anticipated risk of severe GvHD. Before implementation, our results should be validated in other clinical settings, since BM chimerism, ALC and viral infections all depend on the conditioning, donor and use or method of TCD.[59-61] Larger cohorts with more events will allow for more precise prediction of alloimmune responses after DLI, not only GvHD but also the prevention of relapse. Especially the effect of BM chimerism on the risk of relapse should be investigated to confirm the correlation between MC and alloreactivity after DLI. If this is the case, the presence of MC at time of DLI can be considered for the determination of the dose of prophylactic or preemptive DLI.

ACKNOWLEDGEMENTS

Eva Koster has a Landkroon fellowship from the Doelfonds Leukemie of the Bontius Stichting. The authors would like to thank E.J.M. Bogers, S.M. Wallet-Malicka and H.A. Baelde for the data management and technical support.

REFERENCES

1. Horowitz, M., Gale, R., Sondel, P. et al. Graft-versus-leukemia reactions after bone marrow transplantation. Blood 75, 555-562 (1990). doi: 10.1182/blood.V75.3.555.555
2. Falkenburg, J. H. & Jedema, I. Allo-reactive T cells for the treatment of hematological malignancies. Molecular oncology 9, 1894-1903 (2015). doi: 10.1016/j.molonc.2015.10.014
3. Saad, A. & Lamb, L. S. Ex vivo T-cell depletion in allogeneic hematopoietic stem cell transplant: past, present and future. Bone Marrow Transplant 52, 1241-1248 (2017). doi: 10.1038/bmt.2017.22
4. Soiffer, R. J., Kim, H. T., McGuirk, J. et al. Prospective, Randomized, Double-Blind, Phase III Clinical Trial of Anti-T-Lymphocyte Globulin to Assess Impact on Chronic Graft-Versus-Host Disease-Free Survival in Patients Undergoing HLA-Matched Unrelated Myeloablative Hematopoietic Cell Transplantation. J Clin Oncol 35, 4003-4011 (2017). doi: 10.1200/jco.2017.75.8177
5. Castagna, L., Sarina, B., Bramanti, S. et al. Donor lymphocyte infusion after allogeneic stem cell transplantation. Transfus Apher Sci 54, 345-355 (2016). doi: 10.1016/j.transci.2016.05.011
6. Schmid, C., Labopin, M., Schaap, N. et al. Long-term results and GvHD after prophylactic and preemptive donor lymphocyte infusion after allogeneic stem cell transplantation for acute leukemia. Bone Marrow Transplant 57, 215-223 (2022). doi: 10.1038/s41409-021-01515-3
7. Weisdorf, D., Zhang, M. J., Arora, M. et al. Graft-versus-host disease induced graft-versus-leukemia effect: greater impact on relapse and disease-free survival after reduced intensity conditioning. Biol Blood Marrow Transplant 18, 1727-1733 (2012). doi: 10.1016/j.bbmt.2012.06.014
8. Yeshurun, M., Weisdorf, D., Rowe, J. M. et al. The impact of the graft-versus-leukemia effect on survival in acute lymphoblastic leukemia. Blood advances 3, 670-680 (2019). doi: 10.1182/bloodadvances.2018027003
9. Fraser, C. J., Bhatia, S., Ness, K. et al. Impact of chronic graft-versus-host disease on the health status of hematopoietic cell transplantation survivors: a report from the Bone Marrow Transplant Survivor Study. Blood 108, 2867-2873 (2006). doi: 10.1182/blood-2006-02-003954
10. Eefting, M., de Wreede, L. C., Halkes, C. J. et al. Multi-state analysis illustrates treatment success after stem cell transplantation for acute myeloid leukemia followed by donor lymphocyte infusion. Haematologica 101, 506-514 (2016). doi: 10.3324/haematol.2015.136846
11. Solomon, S. R., Sizemore, C., Zhang, X. et al. Current Graft-versus-Host Disease-Free, Relapse-Free Survival: A Dynamic Endpoint to Better Define Efficacy after Allogenic Transplant. Biol Blood Marrow Transplant 23, 1208-1214 (2017). doi: 10.1016/j.bbmt.2017.02.022
12. Holtan, S. G., Zhang, L., DeFor, T. E. et al. Dynamic Graft-versus-Host Disease-Free, Relapse-Free Survival: Multistate Modeling of the Morbidity and Mortality of Allotransplantation. Biol Blood Marrow Transplant 25, 1884-1889 (2019). doi: 10.1016/j.bbmt.2019.05.015
13. Biederstädt, A. & Rezvani, K. How I treat high-risk acute myeloid leukemia using preemptive adoptive cellular immunotherapy. Blood 141, 22-38 (2023). doi: 10.1182/blood.2021012411
14. Falkenburg, J., Schmid, C., Kolb, H., Locatelli, F. & Kuball, J. in The EBMT Handbook. Hematopoietic Stem Cell Transplantation and Cellular Therapies. (eds E. Carreras, C. Dufour, M. Mohty, & N. Kroger) 443-448 (Springer, 2019).
15. Mackinnon, S., Papadopoulos, E. B., Carabasi, M. H. et al. Adoptive immunotherapy evaluating escalating doses of donor leukocytes for relapse of chronic myeloid leukemia after bone marrow transplantation: separation of graft-versus-leukemia responses from graft-versus-host disease. Blood 86, 1261-1268 (1995).

16. Yun, H. D. & Waller, E. K. Finding the sweet spot for donor lymphocyte infusions. Biol Blood Marrow Transplant 19, 507-508 (2013). doi: 10.1016/j.bbmt.2013.02.005
17. Chakraverty, R., Flutter, B., Fallah-Arani, F. et al. The host environment regulates the function of CD8+ graft-versus-host-reactive effector cells. Journal of immunology (Baltimore, Md. : 1950) 181, 6820-6828 (2008). doi: 10.4049/jimmunol.181.10.6820
18. Ferrara, J. L., Levine, J. E., Reddy, P. & Holler, E. Graft-versus-host disease. Lancet 373, 1550-1561 (2009). doi: 10.1016/s0140-6736(09)60237-3
19. Stevanovic, S., van Bergen, C. A., van Luxemburg-Heijs, S. A. et al. HLA class II upregulation during viral infection leads to HLA-DP-directed graft-versus-host disease after CD4+ donor lymphocyte infusion. Blood 122, 1963-1973 (2013). doi: 10.1182/blood-2012-12-470872
20. Thiant, S., Yakoub-Agha, I., Magro, L. et al. Plasma levels of IL-7 and IL-15 in the first month after myeloablative BMT are predictive biomarkers of both acute GVHD and relapse. Bone Marrow Transplant 45, 1546-1552 (2010). doi: 10.1038/bmt.2010.13
21. Thiant, S., Labalette, M., Trauet, J. et al. Plasma levels of IL-7 and IL-15 after reduced intensity conditioned allo-SCT and relationship to acute GVHD. Bone Marrow Transplant 46, 1374-1381 (2011). doi: 10.1038/bmt.2010.300
22. Barge, R. M., Starrenburg, C. W., Falkenburg, J. H. et al. Long-term follow-up of myeloablative allogeneic stem cell transplantation using Campath "in the bag" as T-cell depletion: the Leiden experience. Bone Marrow Transplant 37, 1129-1134 (2006). doi: 10.1038/sj.bmt.1705385
23. von dem Borne, P. A., Beaumont, F., Starrenburg, C. W. et al. Outcomes after myeloablative unrelated donor stem cell transplantation using both in vitro and in vivo T-cell depletion with alemtuzumab. Haematologica 91, 1559-1562 (2006).
24. von dem Borne, P. A., Starrenburg, C. W., Halkes, S. J. et al. Reduced-intensity conditioning allogeneic stem cell transplantation with donor T-cell depletion using alemtuzumab added to the graft ('Campath in the bag'). Current opinion in oncology 21 Suppl 1, S27-29 (2009). doi: 10.1097/01.cco.0000357472.76337.0e
25. Schemper, M. & Smith, T. L. A note on quantifying follow-up in studies of failure time. Control Clin Trials 17, 343-346 (1996). doi: 10.1016/0197-2456(96)00075-x
26. Therneau, T. M. & Grambsch, P. M. Modeling Survival Data: Extending the Cox Model. (Springer, 2000).
27. Gerds, T. A. prodlim: Product-Limit Estimation for Censored Event History Analysis, <https://cran.r-project.org/package=prodlim> (2019).
28. Gray, B. cmprsk: Subdistribution Analysis of Competing Risks, <https://CRAN.R-project.org/package=cmprsk> (2022).
29. De Wreede, L. C., Fiocco, M. & Putter, H. mstate: An R Package for the Analysis of Competing Risks and Multi-State Models. Journal of Statistical Software 38, 1-30 (2011). doi: 10.18637/jss.v038.i07
30. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. (Springer-Verlag New York, 2016).
31. Krassowski, M. ComplexUpset, <https://doi.org/10.5281/zenodo.3700590> (2020).
32. Schmid, C., Labopin, M., Schaap, N. et al. Prophylactic donor lymphocyte infusion after allogeneic stem cell transplantation in acute leukaemia - a matched pair analysis by the Acute Leukaemia Working Party of EBMT. Br J Haematol 184, 782-787 (2019). doi: 10.1111/bjh.15691
33. Miller, H. K., Braun, T. M., Stillwell, T. et al. Infectious Risk after Allogeneic Hematopoietic Cell Transplantation Complicated by Acute Graft-versus-Host Disease. Biol Blood Marrow Transplant 23, 522-528 (2017). doi: 10.1016/j.bbmt.2016.12.630
34. Broers, A. E., van Der Holt, R., van Esser, J. W. et al. Increased transplant-related morbidity and mortality in CMV-seropositive patients despite highly effective prevention of CMV disease after allogeneic T-cell-depleted stem cell transplantation. Blood 95, 2240-2245 (2000). doi: 10.1182/blood.V95.7.2240
35. Cantoni, N., Hirsch, H. H., Khanna, N. et al. Evidence for a bidirectional relationship between cytomegalovirus replication and acute graft-versus-host disease. Biol Blood Marrow Transplant 16, 1309-1314 (2010). doi: 10.1016/j.bbmt.2010.03.020
36. Shaw, B. E., Gooley, T. A., Malkki, M. et al. The importance of HLA-DPB1 in unrelated donor hematopoietic cell transplantation. Blood 110, 4560-4566 (2007). doi: 10.1182/blood-2007-06-095265

37. Fleischhauer, K., Shaw, B. E., Gooley, T. et al. Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study. The Lancet. Oncology 13, 366-374 (2012). doi: 10.1016/s1470-2045(12)70004-9
38. Shlomchik, W. D., Couzens, M. S., Tang, C. B. et al. Prevention of graft versus host disease by inactivation of host antigen-presenting cells. Science (New York, N.Y.) 285, 412-415 (1999). doi: 10.1126/science.285.5426.412
39. Mapara, M. Y., Kim, Y. M., Wang, S. P. et al. Donor lymphocyte infusions mediate superior graft-versus-leukemia effects in mixed compared to fully allogeneic chimeras: a critical role for host antigen-presenting cells. Blood 100, 1903-1909 (2002). doi: 10.1182/blood-2002-01-0023
40. Merad, M., Hoffmann, P., Ranheim, E. et al. Depletion of host Langerhans cells before transplantation of donor alloreactive T cells prevents skin graft-versus-host disease. Nature medicine 10, 510-517 (2004). doi: 10.1038/nm1038
41. Xia, G., Truitt, R. L. & Johnson, B. D. Graft-versus-leukemia and graft-versus-host reactions after donor lymphocyte infusion are initiated by host-type antigen-presenting cells and regulated by regulatory T cells in early and long-term chimeras. Biol Blood Marrow Transplant 12, 397-407 (2006). doi: 10.1016/j.bbmt.2005.11.519
42. Levenga, H., Woestenenk, R., Schattenberg, A. V. et al. Dynamics in chimerism of T cells and dendritic cells in relapsed CML patients and the influence on the induction of alloreactivity following donor lymphocyte infusion. Bone Marrow Transplant 40, 585-592 (2007). doi: 10.1038/sj.bmt.1705777
43. Peggs, K. S., Thomson, K., Hart, D. P. et al. Dose-escalated donor lymphocyte infusions following reduced intensity transplantation: toxicity, chimerism, and disease responses. Blood 103, 1548-1556 (2004). doi: 10.1182/blood-2003-05-1513
44. Shaw, B. E., Byrne, J. L., Das-Gupta, E., Carter, G. I. & Russell, N. H. The impact of chimerism patterns and predonor leukocyte infusion lymphopenia on survival following T cell-depleted reduced intensity conditioned transplants. Biol Blood Marrow Transplant 13, 550-559 (2007). doi: 10.1016/j.bbmt.2006.12.451
45. Lutz, C., Massenkeil, G., Nagy, M. et al. A pilot study of prophylactic donor lymphocyte infusions to prevent relapse in adult acute lymphoblastic leukemias after allogeneic hematopoietic stem cell transplantation. Bone Marrow Transplant 41, 805-812 (2008). doi: 10.1038/sj.bmt.1705981
46. Liga, M., Triantafyllou, E., Tiniakou, M. et al. High alloreactivity of low-dose prophylactic donor lymphocyte infusion in patients with acute leukemia undergoing allogeneic hematopoietic cell transplantation with an alemtuzumab-containing conditioning regimen. Biol Blood Marrow Transplant 19, 75-81 (2013). doi: 10.1016/j.bbmt.2012.07.021
47. Caldemeyer, L. E., Akard, L. P., Edwards, J. R. et al. Donor Lymphocyte Infusions Used to Treat Mixed-Chimeric and High-Risk Patient Populations in the Relapsed and Nonrelapsed Settings after Allogeneic Transplantation for Hematologic Malignancies Are Associated with High Five-Year Survival if Persistent Full Donor Chimerism Is Obtained or Maintained. Biol Blood Marrow Transplant 23, 1989-1997 (2017). doi: 10.1016/j.bbmt.2017.07.007
48. Bar, M., Sandmaier, B. M., Inamoto, Y. et al. Donor lymphocyte infusion for relapsed hematological malignancies after allogeneic hematopoietic cell transplantation: prognostic relevance of the initial CD3+ T cell dose. Biol Blood Marrow Transplant 19, 949-957 (2013). doi: 10.1016/j.bbmt.2013.03.001
49. Collins, T., Korman, A. J., Wake, C. T. et al. Immune interferon activates multiple class II major histocompatibility complex genes and the associated invariant chain gene in human endothelial cells and dermal fibroblasts. Proceedings of the National Academy of Sciences of the United States of America 81, 4917-4921 (1984). doi: 10.1073/pnas.81.15.4917
50. Koyama, M., Kuns, R. D., Olver, S. D. et al. Recipient nonhematopoietic antigen-presenting cells are sufficient to induce lethal acute graft-versus-host disease. Nature medicine 18, 135-142 (2011). doi: 10.1038/nm.2597
51. Collin, M. P., Hart, D. N., Jackson, G. H. et al. The fate of human Langerhans cells in hematopoietic stem cell transplantation. The Journal of experimental medicine 203, 27-33 (2006). doi: 10.1084/jem.20051787
52. Mielcarek, M., Kirkorian, A. Y., Hackman, R. C. et al. Langerhans cell homeostasis and turnover after nonmyeloablative and myeloablative allogeneic hematopoietic cell transplantation.

Transplantation 98, 563-568 (2014). doi: 10.1097/tp.0000000000000097
53. van Balen, P., van der Zouwen, B., Kruisselbrink, A. B. et al. Tissue Damage Caused by Myeloablative, but Not Non-Myeloablative, Conditioning before Allogeneic Stem Cell Transplantation Results in Dermal Macrophage Recruitment without Active T-Cell Interaction. Front Immunol 9, 331 (2018). doi: 10.3389/fimmu.2018.00331
54. Miller, J. S., Weisdorf, D. J., Burns, L. J. et al. Lymphodepletion followed by donor lymphocyte infusion (DLI) causes significantly more acute graft-versus-host disease than DLI alone. Blood 110, 2761-2763 (2007). doi: 10.1182/blood-2007-05-090340
55. Kanakry, C. G., Bolaños-Meade, J., Kasamon, Y. L. et al. Low immunosuppressive burden after HLA-matched related or unrelated BMT using posttransplantation cyclophosphamide. Blood 129, 1389-1393 (2017). doi: 10.1182/blood-2016-09-737825
56. Kawamura, K., Nakasone, H., Kurosawa, S. et al. Refractory Graft-Versus-Host Disease-Free, Relapse-Free Survival as an Accurate and Easy-to-Calculate Endpoint to Assess the Long-Term Transplant Success. Biol Blood Marrow Transplant 24, 1521-1526 (2018). doi: 10.1016/j.bbmt.2018.02.004
57. Bluhmki, T., Schmoor, C., Finke, J. et al. Relapse- and Immunosuppression-Free Survival after Hematopoietic Stem Cell Transplantation: How Can We Assess Treatment Success for Complex Time-to-Event Endpoints? Biol Blood Marrow Transplant 26, 992-997 (2020). doi: 10.1016/j.bbmt.2020.01.001
58. Carnevale-Schianca, F., Caravelli, D., Gallo, S. et al. Post-Transplant Cyclophosphamide and Tacrolimus-Mycophenolate Mofetil Combination Governs GVHD and Immunosuppression Need, Reducing Late Toxicities in Allogeneic Peripheral Blood Hematopoietic Cell Transplantation from HLA-Matched Donors. Journal of clinical medicine 10 (2021). doi: 10.3390/jcm10061173
59. Antin, J. H., Childs, R., Filipovich, A. H. et al. Establishment of complete and mixed donor chimerism after allogeneic lymphohematopoietic transplantation: recommendations from a workshop at the 2001 Tandem Meetings of the International Bone Marrow Transplant Registry and the American Society of Blood and Marrow Transplantation. Biol Blood Marrow Transplant 7, 473-485 (2001). doi: 10.1053/bbmt.2001.v7.pm11669214
60. Ando, T., Tachibana, T., Tanaka, M. et al. Impact of graft sources on immune reconstitution and survival outcomes following allogeneic stem cell transplantation. Blood advances 4, 408-419 (2020). doi: 10.1182/bloodadvances.2019001021
61. Bejanyan, N., Brunstein, C. G., Cao, Q. et al. Delayed immune reconstitution after allogeneic transplantation increases the risks of mortality and chronic GVHD. Blood advances 2, 909-922 (2018). doi: 10.1182/bloodadvances.2017014464

SUPPLEMENTAL METHODS

Transplantation protocol

Myeloablative conditioning (MAC) consisted of cyclophosphamide (2 days 60 mg/kg intravenously) combined with either 9 Gy total body irradiation or busulfan (4 days 4x0.8 mg/kg intravenously). Reduced-intensity conditioning (RIC) consisted either of fludarabine (6 days 50 mg/m^2 orally or 30 mg/m^2 intravenously) and busulfan (2 days 4x0.8 mg/kg intravenously), or the FLAMSA regimen: fludarabine (5 days 30 mg/m^2 intravenously), cytarabine (4 days 2000 mg/m^2 intravenously), amsacrine (4 days 100 mg/m^2 intravenously) and busulfan (4 days 4x0.8 mg/kg intravenously).

Standard *in vitro* TCD was performed by adding 20 mg alemtuzumab (Sanofi Genzyme) to the graft. (1) Additional *in vivo* TCD depended on the donor type and conditioning regimen: MAC patients with a RD did not receive any *in vivo* TCD. All other patients received 15 mg alemtuzumab intravenously on days -6 and -5 (MAC) or on days -4 and -3 (RIC). Before June 2007, RIC patients with an UD received 10 mg/kg horse-derived anti-thymocyte globulin (Lymphoglobulin, Genzyme) additionally on day -4 until day -1. After Lymphoglobulin was withdrawn from the market, RIC patients with an UD first received no anti-thymocyte globulin (alloSCT between June 2007 and September 2009) and later received rabbit-derived anti-thymocyte globulin (Thymoglobulin, Sanofi Genzyme) additionally on day -2 (until April 2010 2mg/kg and thereafter 1mg/kg). Only MAC patients with an UD received posttransplant ciclosporin as GvHD prophylaxis, which was tapered from 1 month with the aim to stop within 3 months after alloSCT.

BM chimerism, ALC and viral infections

For the BM chimerism at time of DLI, we used the BM sample that was closest to DLI and taken within 5 weeks before and 1 week after DLI. BM chimerism was measured in total BM leukocytes by short-tandem-repeat PCR or, for patients transplanted before 2007 with a sex-mismatched donor, by FISH analysis using Vysis CEP X/Y probes. The lower detection limit of the chimerism analyses was 1-2%, depending on the method and the selected markers. For patients without any evaluable BM chimerism measurement during this period but whose last measurement before and first measurement after DLI belonged to the same chimerism category (FDC, low MC or high MC), this category was taken as the BM chimerism status at time of DLI.

ALC was calculated by the sum of the absolute numbers of circulating T cells, B cells and NK cells as measured on anticoagulated fresh venous blood by flow cytometry with bead calibration (Trucount tubes, Becton Dickinson) with a lower detection limit of $0.5x10^6$ cells/l. If these counts were unavailable, the lymphocyte count by manual blood smear was used. For the ALC at time of DLI, the closest measurement within 2 weeks before and 1 week after DLI was taken. For patients without any ALC measurement during this period but whose last ALC before and first ALC after DLI belonged to the same category (<500, 500-999 or $\geq 1000x10^6$/l for the low-dose 3-month DLI and <1000 or $\geq 1000x10^6$/l for the 6-month DLI), this category was taken as the ALC at time of DLI.

Cytomegalovirus (CMV) and Epstein-Barr virus (EBV) were monitored weekly by PCR on peripheral blood samples in all patients. Single positive values of CMV or EBV below log 2.4 were not considered. PCRs on other viruses were only performed in symptomatic patients. For the analyses, only the first viral infection was used.

Multi-state modelling

In a multi-state model patients move between states at the occurrence of clinical events or treatments. Transitions define which routes between states are allowed (for instance the transition from the state DLI to the state GvHD). (2) In a Markov model, the hazard of making a certain transition only depends on the current state and the time since start, which is in this case the first DLI. Each transition hazard can either be estimated without taking covariates into account (non-parametrically) or can be analyzed by means of a transition-specific Cox proportional hazards model (semi-parametric approach). The baseline hazards and the hazard ratios are the building blocks for the calculation of the transition probabilities, which represent the probabilities of being in each of the states over time. For example, in a semi-parametric model the probability of being alive with GvHD depends on the baseline hazard of GvHD, the effects of risk factors for GvHD, and the risks of death and disappearance of GvHD symptoms. Confidence intervals for the probabilities of cGRFS, death after start of tIS, being alive with clinically relevant GvHD, RFS and OS were calculated based on the estimated variance-covariance matrix of all transition probabilities.

References

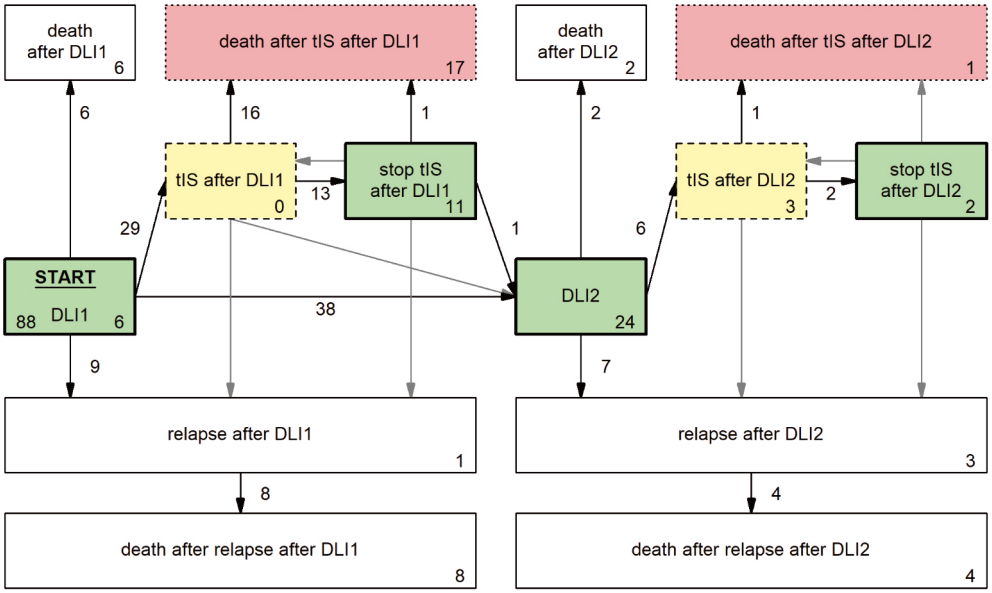
1. von dem Borne, P. A., Beaumont, F., Starrenburg, C. W. et al. Outcomes after myeloablative unrelated donor stem cell transplantation using both in vitro and in vivo T-cell depletion with alemtuzumab. Haematologica 91, 1559-1562 (2006).
2. Putter, H., Fiocco, M. & Geskus, R. B. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 26, 2389-2430 (2007). doi: 10.1002/sim.2712

SUPPLEMENTAL TABLES

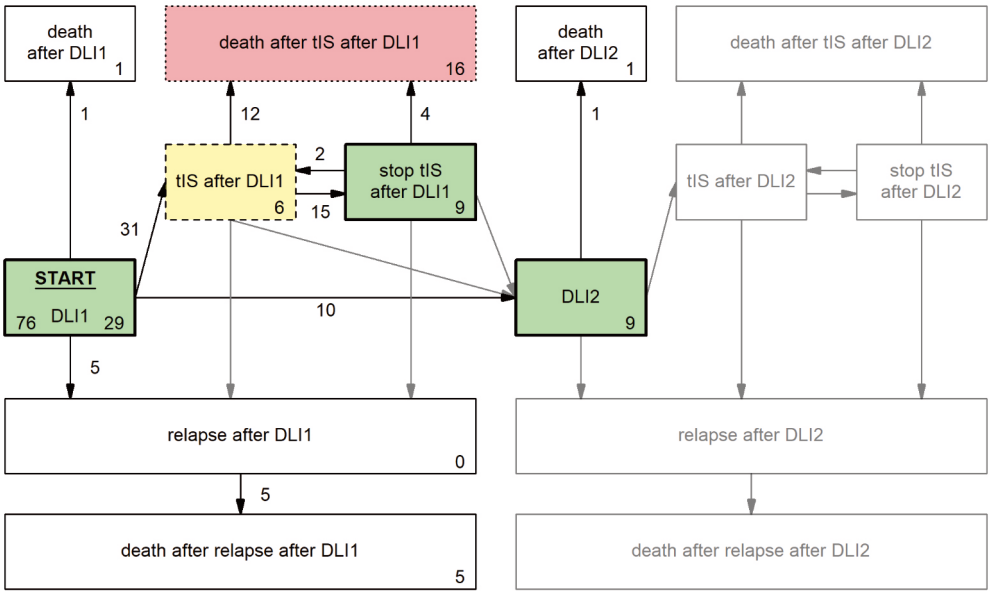
	Total cohort (all included patients with alloSCT; N = 388)
Age at alloSCT (years)	
median (range)	54 (18-78)
Disease	
acute myeloid leukemia	260 (67%)
acute lymphoblastic leukemia	85 (22%)
myelodysplastic syndrome	43 (11%)
Conditioning	
MAC: Cyclo/TBI	196 (51%)
MAC: Cyclo/Bu	9 (2%)
RIC: Flu/Bu*	167 (43%)
RIC: Flu/Bu/Ara-C/Amsa	16 (4%)
Donor	
RD	165 (43%)
UD	223 (57%)
Graft source	
G-CSF mobilized PBSC	368 (95%)
BM	20 (5%)
CMV serostatus patient/donor	
+/+	169 (44%)
+/-	70 (18%)
-/+	29 (7%)
-/-	120 (31%)
EBV serostatus patient/donor	
+/+	323 (83%)
+/-	30 (8%)
+/unknown	15 (4%)
-/+	18 (5%)
-/-	2 (1%)

Supplemental Table 1. Baseline characteristics of all included patients who received an alloSCT. alloSCT, allogeneic stem cell transplantation; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; Cyclo, cyclophosphamide; TBI, total body irradiation; Bu, busulfan; Flu, fludarabine; Ara-C, cytarabine; Amsa, amsacrine; RD, related donor; UD, unrelated donor; G-CSF, granulocyte-colony stimulation factor; PBSC, peripheral blood stem cells; BM, bone marrow; CMV, cytomegalovirus; EBV, Epstein-Barr virus. *One patient had not received a second consolidation course before transplant and received 2 days cyclophosphamide 750 mg/m^2 intravenously additionally to the conditioning regimen.

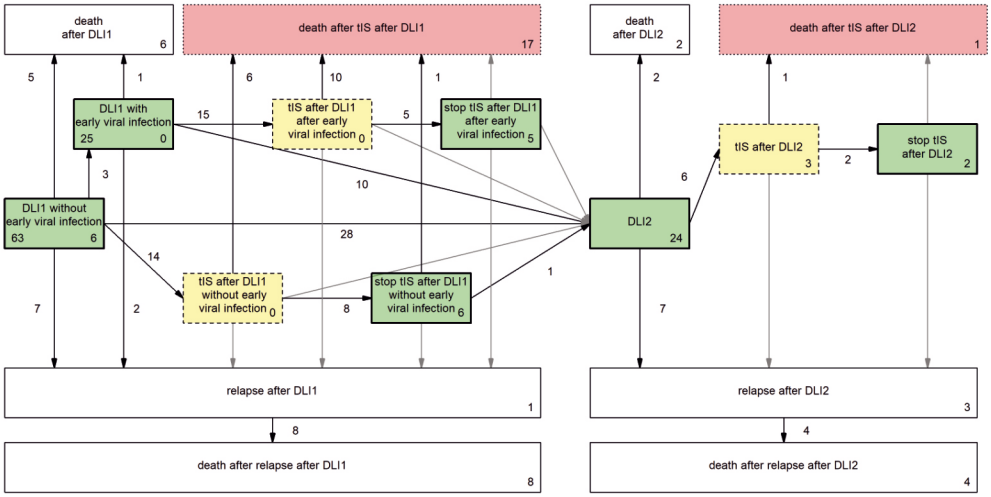
SUPPLEMENTAL FIGURES



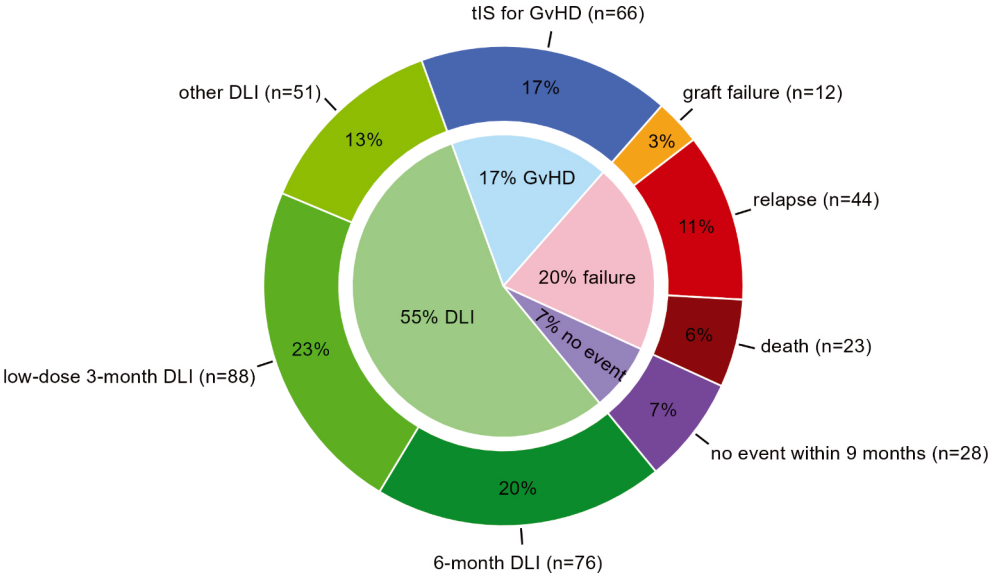
Supplemental Figure 1. Multi-state model for low-dose 3-month DLI. Boxes represent states and arrows represent the transitions between the states. Grey transitions were not used by any of the included patients and omitted from the final model. All patients started in the state 'DLI1'. The number at the bottom left corner of the starting state shows the number of patients included in the model. The numbers at the bottom right corner of the boxes show the numbers of the patients who were in that state at the end of their follow-up. The numbers next to the arrows show the numbers of the patients who made that transition during their follow-up. The cGRFS is the sum of the probabilities of all green (thick border) states, the probability of being alive with GvHD the sum of all yellow (dashed border) states, the probability of death after start of tIS for GvHD the sum of all red (dotted border) states, the RFS is the sum of all green (thick border) and yellow (dashed border) states, and the OS the sum of all non-death states. For these summarizing measures, no distinction was made between states after the first DLI or after multiple DLIs.



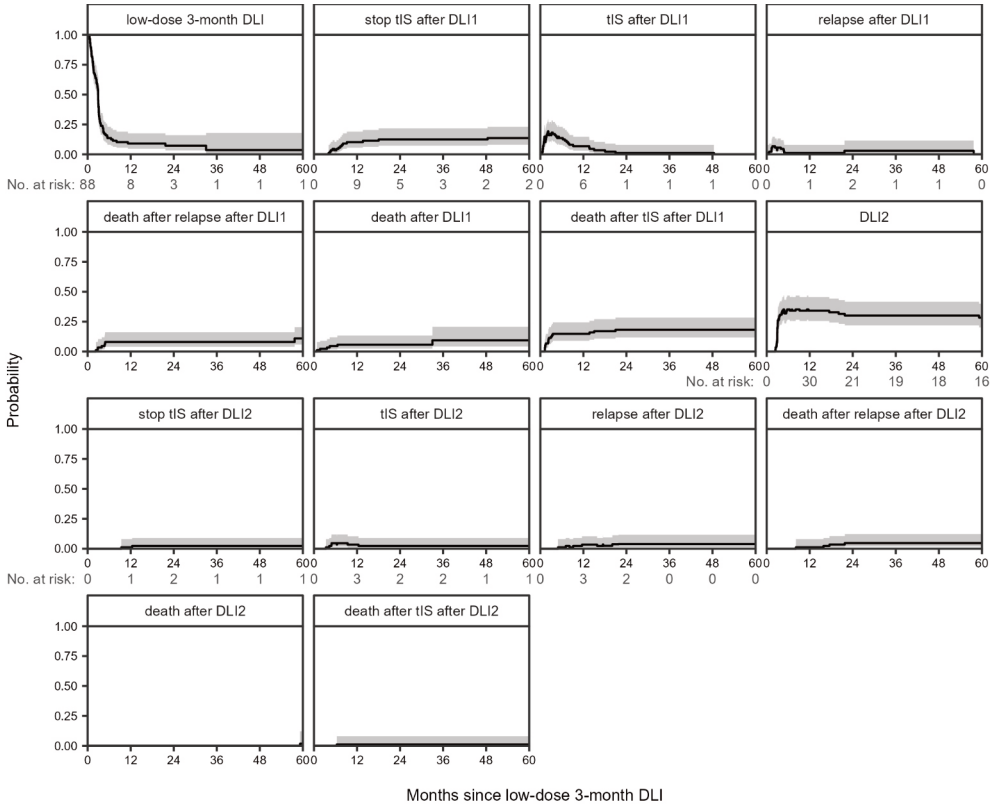
Supplemental Figure 2. Multi-state model for 6-month DLI. Boxes represent states and arrows represent the transitions between the states. Grey states and transitions were not used by any of the included patients and omitted from the final model. All patients started in the state 'DLI1'. The number at the bottom left corner of the starting state shows the number of patients included in the model. The numbers at the bottom right corner of the boxes show the numbers of the patients who were in that state at the end of their follow-up. The numbers next to the arrows show the numbers of the patients who made that transition during their follow-up. The cGRFS is the sum of the probabilities of all green (thick border) states, the probability of being alive with GvHD the sum of all yellow (dashed border) states, the probability of death after start of tIS for GvHD the sum of all red (dotted border) states, the RFS the sum of all green (thick border) and yellow (dashed border) states, and the OS the sum of all non-death states. For these summarizing measures, no distinction was made between states after the first DLI or after multiple DLIs.



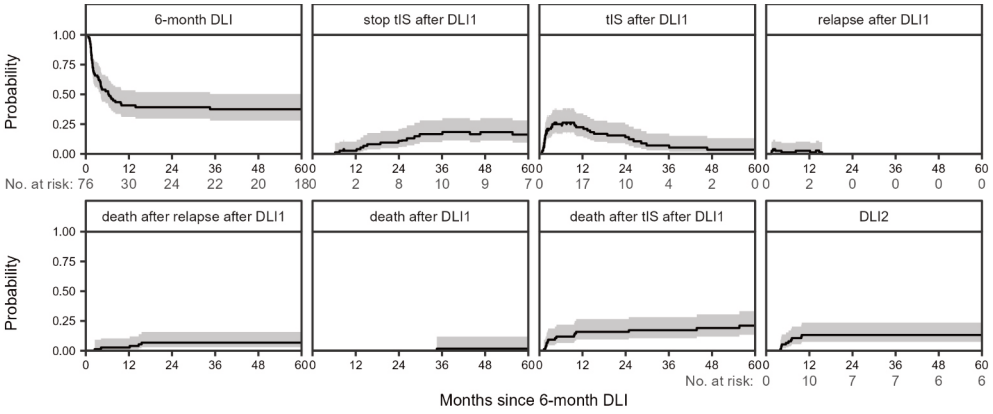
Supplemental Figure 3. Multi-state model for low-dose 3-month DLI considering early viral infections. Boxes represent states and arrows represent the transitions between the states. Grey transitions were not used by any of the included patients and omitted from the final model. 63 patients had no viral infection during the last week before DLI and started in the state 'DLI1 without early viral infection', while the 25 patients with a viral infection during the last week before DLI started in the state 'DLI1 with early viral infection' (see the numbers at the bottom left corner of the two starting states). Patients who had an early viral infection during the first 2 weeks after DLI without any prior event moved from 'DLI1 without early viral infection' to 'DLI1 with early viral infection' at time of the viral infection. The numbers at the bottom right corner of the boxes show the numbers of the patients who were in that state at the end of their follow-up. The numbers next to the arrows show the numbers of the patients who made that transition during their follow-up. The cGRFS is the sum of the probabilities of all green (thick border) states, the probability of being alive with GvHD the sum of all yellow (dashed border) states, the probability of death after start of tIS for GvHD the sum of all red (dotted border) states, the RFS the sum of all green (thick border) and yellow (dashed border) states, and the OS the sum of all non-death states. For these summarizing measures, no distinction was made between states after the first DLI or after multiple DLIs.



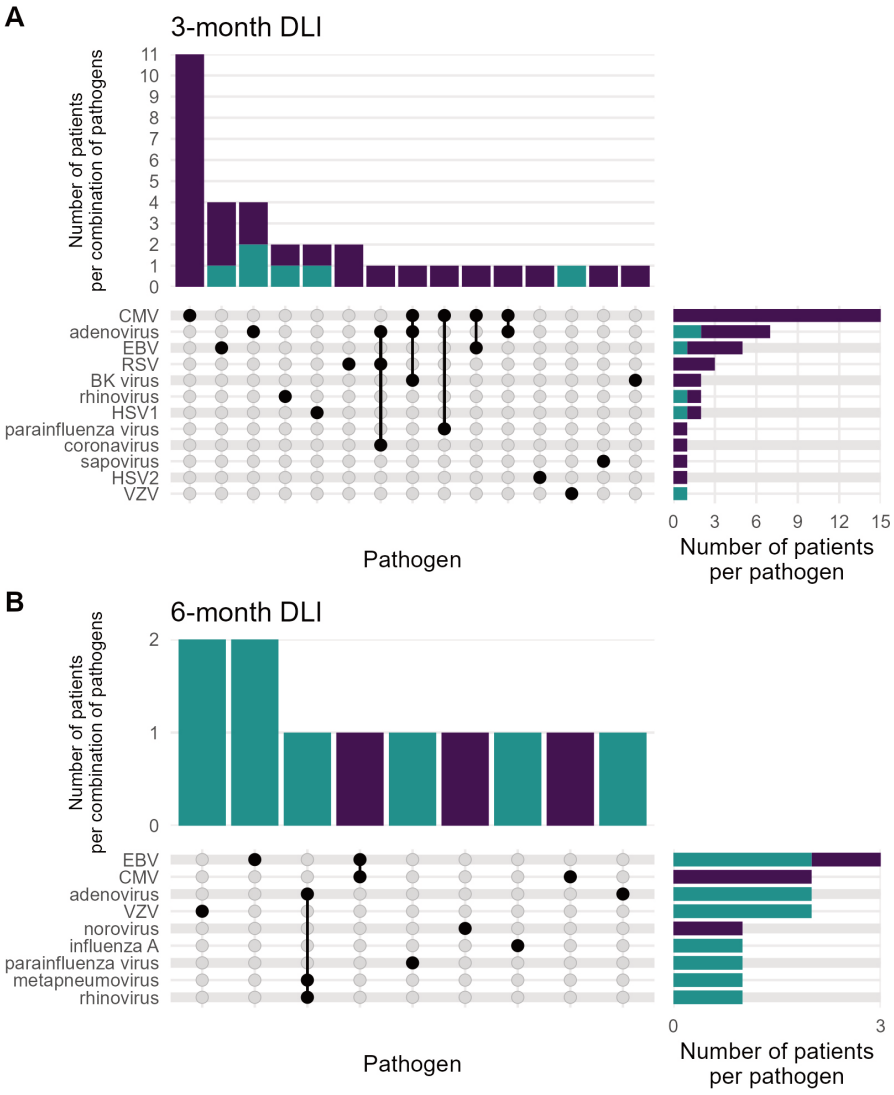
Supplemental Figure 4. Selection of the DLI cohorts. Events during the first 9 months after alloSCT for the total cohort. Per patient only the first occurring event was taken into account. The inner circle describes the main event categories (DLI, GvHD, treatment failure (i.e., death, relapse or graft failure), no event), while the outer circle further specifies the kind of DLI or treatment failure. The 88 patients who received the low-dose 3-month DLI and the 76 patients who received the 6-month DLI as first DLI were included in the DLI analyses.



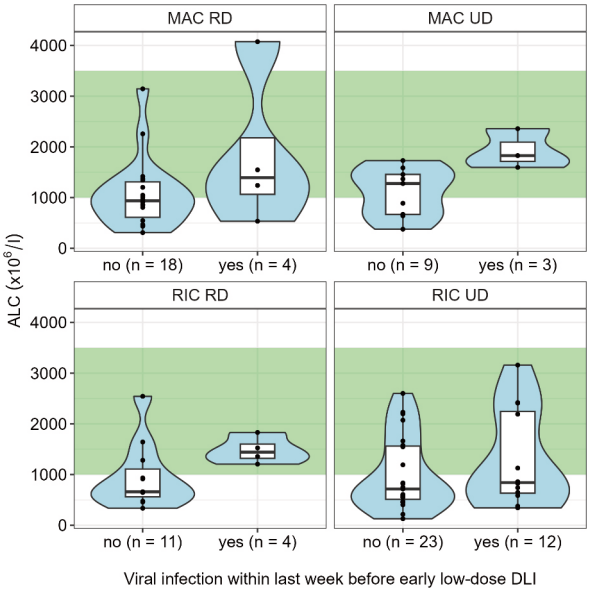
Supplemental Figure 5. Outcomes after low-dose 3-month DLI: probabilities with associated 95% confidence intervals per state. Probabilities with associated 95% confidence intervals for each state. The at risk numbers are shown for all non-death states and indicate the numbers of uncensored patients present in each state at different timepoints.



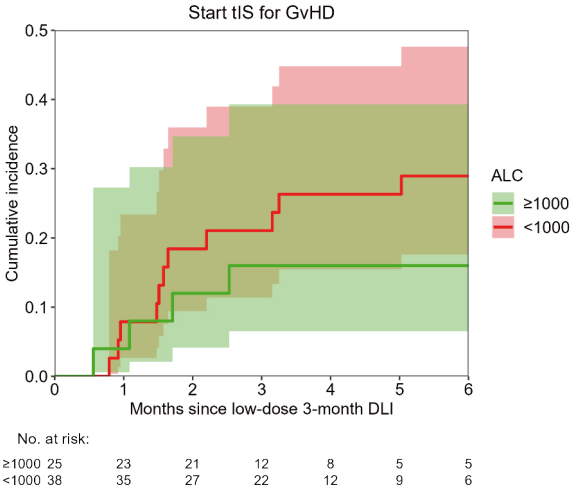
Supplemental Figure 6. Outcomes after 6-month DLI: probabilities with associated 95% confidence intervals per state. Probabilities with associated 95% confidence intervals for each state. The at risk numbers are shown for all non-death states and indicate the numbers of uncensored patients present in each state at different timepoints. Only states that were occupied within 5 years after 6-month DLI are shown.



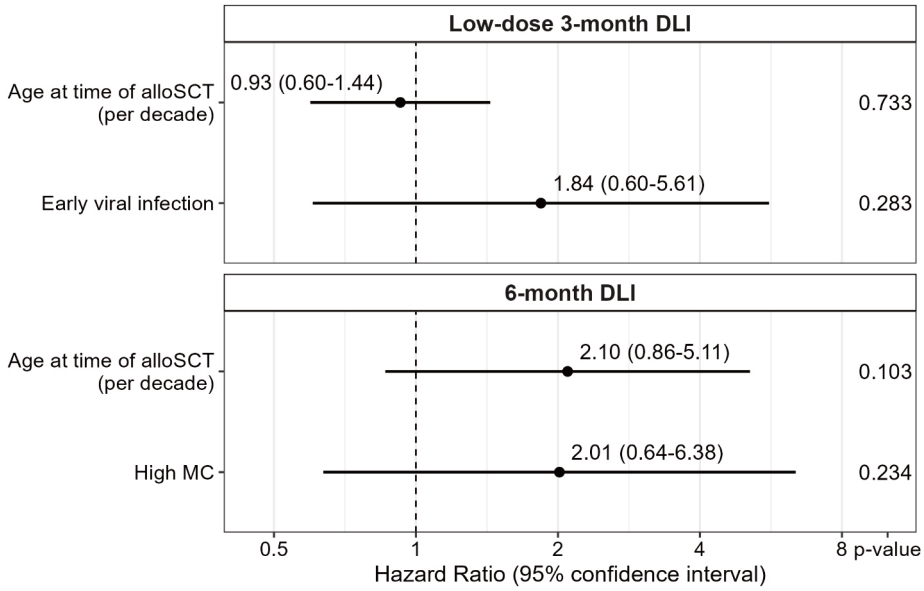
Supplemental Figure 7. Pathogens of viral infections close to the first DLI. UpSet plots of all viral pathogens present within 1 week before and 8 weeks after the low-dose 3-month DLI (panel A) or the 6-month DLI (panel B). The horizontal bar charts show for each of the pathogens the number of patients with this pathogen. As can be seen by the dot-connecting lines, some patients had multiple pathogens during this period. The vertical bar charts show the numbers of patients for each of the combinations. Purple indicates early onset (<2 weeks after DLI) viral infections, turquoise late onset (>2 weeks after DLI) infections. For instance, 3 patients had an EBV viremia close to the 6-month DLI, of whom two beyond 2 weeks after DLI without any other pathogen. The other patient had an early EBV viremia and a CMV viremia.



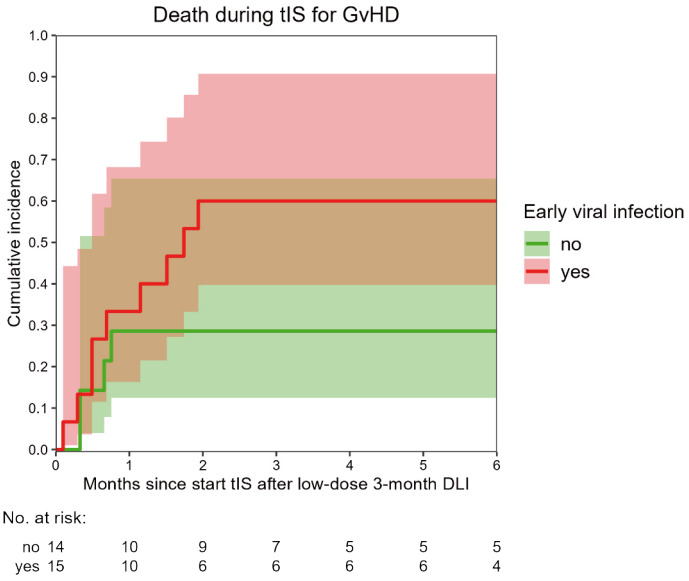
Supplemental Figure 8. ALC per conditioning/donor type and viral status at time of low-dose 3-month DLI. ALC at time of low-dose 3-month DLI per conditioning/donor type and the presence of a viral infection within the last week before this DLI. The boxplots are combined with violin plots showing the kernel probability density to visualize the distribution of the data. The lower and upper hinges of the boxplots correspond to the 25th and 75th percentiles. The green area shows the normal range used in our laboratory. Four patients for whom the exact ALC at time of DLI was unknown, were excluded (1 MAC RD without viral infection, 1 MAC RD with viral infection, 1 RIC UD without viral infection, 1 RIC UD with viral infection).



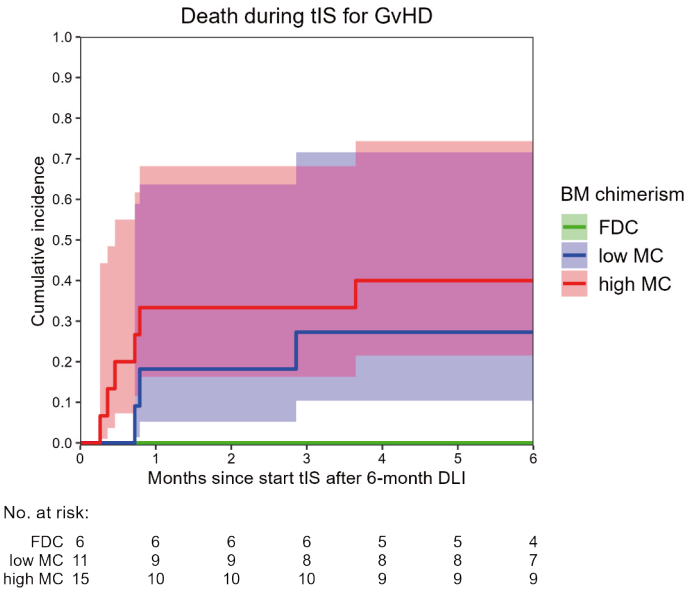
Supplemental Figure 9. Cumulative incidence of tIS for GvHD after low-dose 3-month DLI in the absence of viral infections within the last week before DLI. Cumulative incidences with associated 95% confidence intervals of tIS for GvHD after the low-dose 3-month DLI for patients with ALC ≥1000x10^6/l (n=25) or lower (n=38). This was calculated in a competing risks model starting at time of low-dose 3-month DLI with start tIS, relapse, death and DLI2 as competing events. The 25 patients with a viral infection during the last week before the low-dose 3-month DLI were excluded.



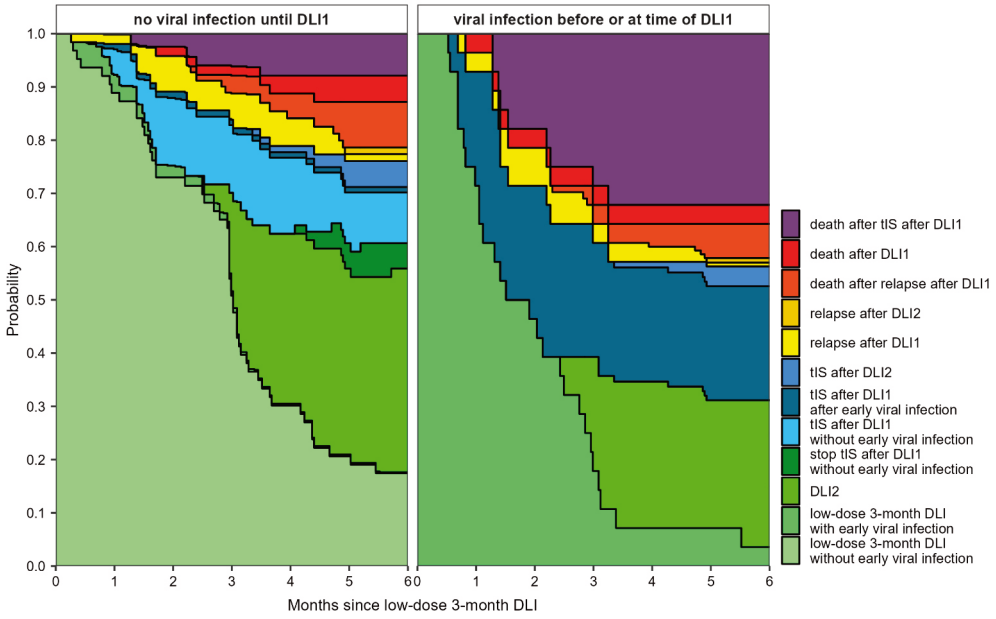
Supplemental Figure 10. Risk factors for death during tIS for GvHD. Cox proportional hazards models for the transition from tIS for GvHD after DLI1 to death (see Figure 1). Based on complete case analysis (n=29 for low-dose 3-month DLI and n=31 (age) or n=30 (chimerism) for 6-month DLI). DLI, donor lymphocyte infusion; alloSCT, allogeneic stem cell transplantation; high MC, ≥5% mixed chimerism in the bone marrow



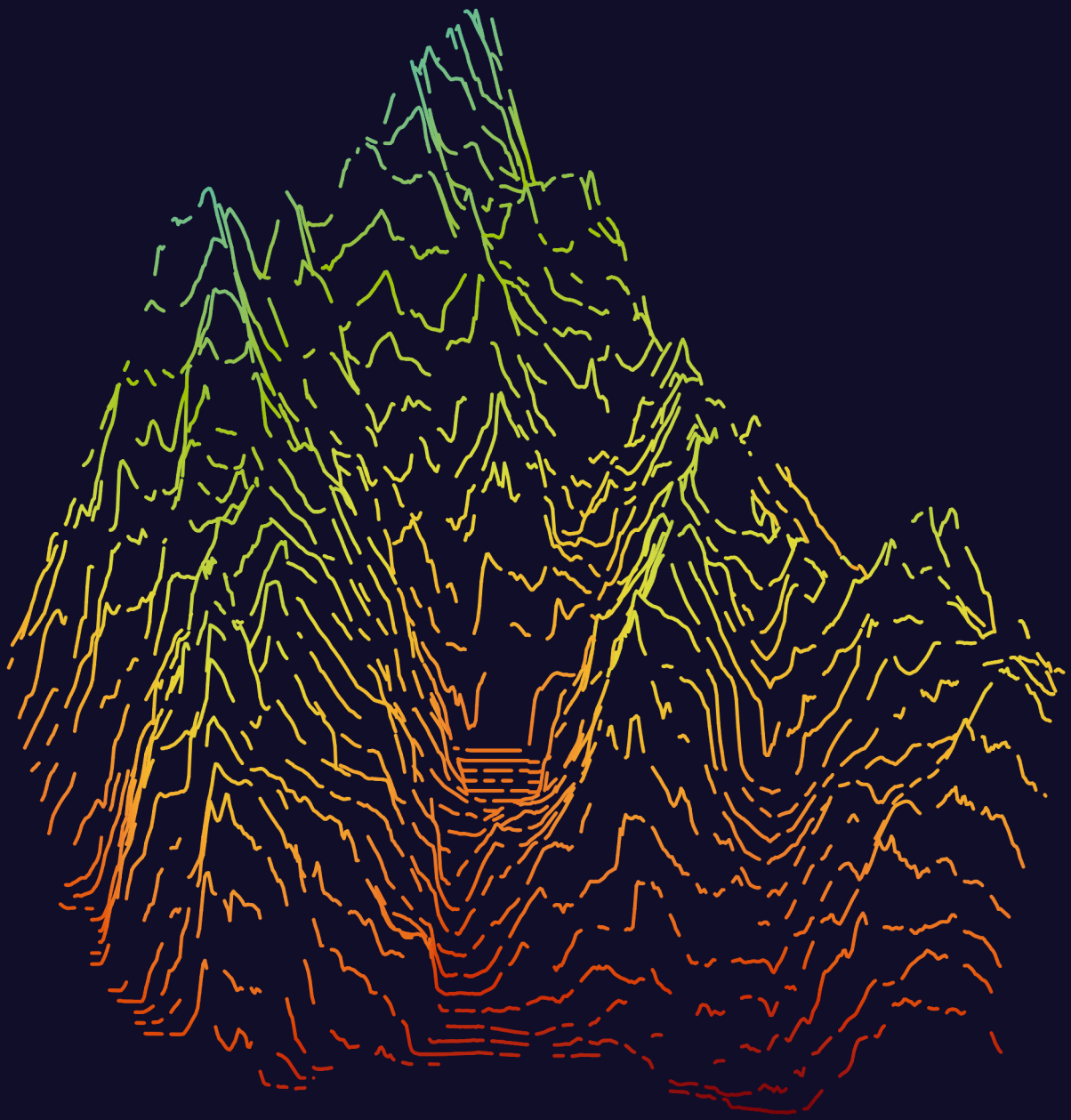
Supplemental Figure 11. Cumulative incidence of death during tIS for GvHD after low-dose 3-month DLI. Cumulative incidences with associated 95% confidence intervals of death during tIS for GvHD for patients who developed GvHD after an early viral infection and those without any early viral infection. This was calculated in a competing risks model starting at time of start tIS for GvHD after DLI with death, DLI2, relapse and stop tIS as competing events.



Supplemental Figure 12. Cumulative incidence of death during tIS for GvHD after 6-month DLI. Cumulative incidences with associated 95% confidence intervals of death during tIS for GvHD per BM chimerism status at time of DLI for patients who developed GvHD after the 6-month DLI. This was calculated in a competing risks model starting at time of start tIS for GvHD after DLI with death, DLI2, relapse and stop tIS as competing events. One patient with FDC and one with high MC had two tIS episodes and entered the risk set twice.



Supplemental Figure 13. Outcomes after low-dose 3-month DLI based on the viral status at time of DLI. Stacked state occupation probabilities after low-dose 3-month DLI based on the viral status at time of DLI (viral infection during the last week before DLI (n=25) or no viral infection until DLI (n=63). The estimates are based on the non-parametric multi-state model in Supplemental Figure 3 which has two starting states ('DLI1 without early viral infection' and 'DLI1 with early viral infection'). The difference between two adjacent curves represents the probability of being in the corresponding state. States that were not used within 6 months after DLI were omitted from the legend.



5

Transplantation strategy affects the risk of GvHD after prophylactic and preemptive donor lymphocyte infusion

Eva A.S. Koster, Peter A. von dem Borne, Joost G.K. van der Hem, Peter van Balen, Erik W.A. Marijt, Jennifer M.L. Tjon, Tjeerd J.F. Snijders, Daniëlle van Lammeren, Hendrik Veelken, J.H. Frederik Falkenburg, Liesbeth C. de Wreede, Constantijn J.M. Halkes

Ann Hematol. 2025 Oct;104(10):5379-5387.

ABSTRACT

Donor lymphocyte infusion (DLI) after allogeneic stem cell transplantation (alloSCT) can boost Graft-versus-Leukemia (GvL) reactivity but may induce Graft-versus-Host-Disease (GvHD). It is essential to understand which factors besides timing, donor type, and dose influence DLI alloreactivity. We previously identified viral infections, ≥5% patient cells in bone marrow chimerism, and lymphopenia at the time of DLI as relevant factors for GvHD after DLI following alemtuzumab-based T-cell depletion. Here, we investigated these factors and the alloreactivity after DLI following alloSCT with posttransplant cyclophosphamide in 83 patients with acute leukemia/myelodysplastic syndrome receiving a prophylactic or preemptive DLI. 5% had viral infections close to DLI, 6% had ≥5% mixed chimerism, and 17% had lymphopenia. 2-year cumulative incidence of GvHD requiring systemic treatment was low: 7% (95%-confidence interval 1-14%). 22 of the 28 patients with ≥1% mixed chimerism at the time of DLI (79%) converted to full-donor chimerism. None of these responders relapsed, indicating achievement of GvL despite the low incidence of GvHD. Our data show that DLI alloreactivity is determined by the conditions at the time of DLI which are influenced by the transplantation strategy. Adjusting the DLI dose based on these conditions may improve the balance between GvHD and GvL.

INTRODUCTION

Relapse remains an important cause of failure of allogeneic stem cell transplantation (alloSCT) in patients with acute leukemia. Unmodified donor lymphocyte infusions (DLI) can be given to boost the Graft-versus-Leukemia (GvL) effect to prevent relapse, but may induce Graft-versus-Host-Disease (GvHD). To improve the balance between GvHD and GvL and thereby the applicability of DLI, it is crucial to better understand which factors influence the alloreactivity of DLI.

Expert opinion recommends that dosing of prophylactic and preemptive DLI should at least be based on donor type and time after alloSCT to reduce the risk of severe GvHD.[1] In a recent study, we identified three other risk factors for the development of GvHD after DLI following alemtuzumab-based T-cell depleted (TCD) alloSCT: occurrence of viral infections (*de novo* or reactivation) close to DLI, presence of patient-derived antigen-presenting cells (APCs) in the bone marrow (BM), and lymphopenia.[2] Patient-derived APCs are highly capable of activating donor-derived alloreactive T cells.[3] After alloSCT, the professional APCs of the patient are gradually replaced by donor-derived APCs. We previously showed that the replacement of APCs in the skin occurs predominantly between 3 and 6 months after alloSCT.[4] Thus, from 6 months onwards the BM chimerism status should be a good indicator of the origin of the professional APCs in the peripheral tissues. Viral infections and lymphopenia promote the activation of (alloreactive) T cells.[5-7] The presence of these factors at the time of DLI depends on the transplantation strategy (i.e., conditioning intensity, use and type of TCD, and GvHD prophylaxis). Posttransplant cyclophosphamide (PTCY) preferentially targets activated alloreactive T cells and favors recovery of regulatory T cells. [8,9] This leads to relatively early lymphocyte recovery and better protection against severe infections compared to other TCD strategies.[10-12] Additionally, most patients achieve full-donor chimerism (FDC, <1% patient cells) within two months after PTCY alloSCT.[13] This profile could therefore be associated with low alloreactivity of DLI following PTCY alloSCT.[2] Indeed, the risk of GvHD appears to be similar between haploidentical DLI following PTCY alloSCT and DLI from HLA-matched donors after non-PTCY alloSCT despite the larger genetic disparity.[14] In the non-haploidentical PTCY setting, only two studies have reported outcomes after DLI. Carnevale-Schianca et al. investigated 14 patients receiving therapeutic DLI after which none developed grade III-IV acute GvHD and 1 patient developed chronic GvHD.[15] They reported an overall response rate of 57%. However, as more than half of the patients also received systemic therapy or radiotherapy, the contribution of the DLI itself on disease control is unclear.[15] Shanmugasundaram et al. investigated 38 DLIs given to 21 patients after PTCY, of whom 8 with a non-haploidentical donor, and observed low risks of acute (8%) and chronic (3%) GvHD but limited efficacy with 11% and 15% complete response after DLI for relapse and mixed chimerism, respectively.[16] These reported risks of GvHD are considerably lower than those observed after non-haploidentical DLI following other transplantation strategies.[2,17,18] However, both studies involved a wide variety of conditioning regimens and DLI settings (i.e., timing since alloSCT, DLI dose and pre-DLI treatments such as chemotherapy and steroids), making it hard to investigate the impact of the transplantation strategy and DLI circumstances on the alloreactivity of DLI.

In the current study, we investigated DLI after non-haploidentical PTCY alloSCT in a more homogeneous cohort treated according to a standardized DLI protocol: all patients were scheduled for prophylactic DLI at 4 or 6 months after alloSCT with fixed doses based on timing and donor type. We analyzed the conditions at the time of DLI and assessed the alloreactivity after DLI, i.e. development of clinically relevant GvHD, conversion of mixed chimerism (MC, ≥1% patient cells) to FDC and the risk of relapse. By following the same systematic approach we used in the setting of DLI after alemtuzumab-based TCD alloSCT, the impact of the transplantation strategy on the DLI conditions and alloreactivity can be investigated.

METHODS

Study population

This observational study included all adult patients with acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL) or myelodysplastic syndrome with excess blasts (MDS-EB2) in complete morphologic remission who received PTCY alloSCT from a ≥8/10 HLA-matched donor at Leiden University Medical Center (LUMC, Leiden, The Netherlands) between April 2020 and December 2022. The DLI cohort consisted of all patients who received a first DLI scheduled at 4 or 6 months after alloSCT (actually administered at 3.7-5.2 months and 5.3-9.0 months, respectively) without prior relapse or therapeutic systemic immunosuppression (tIS) for GvHD. The study was approved by the Medical Ethics Committee Leiden The Hague Delft (RP 22.002). All patients signed informed consent for data collection and analysis. Data were analyzed as of March 2024.

Transplantation and DLI protocol

Myeloablative conditioning consisted either of cyclophosphamide (2 days 60 mg/kg iv) and total body irradiation (3 days 2x2 Gy), or of thiotepa (2 days 5 mg/kg iv), fludarabine (3 days 50 mg/m^2 iv) and busulfan (3 days 4x0.8 mg/kg iv). Reduced-intensity conditioning consisted of fludarabine (5 days 30 mg/m^2 iv), cyclophosphamide (2 days 14.5 mg/kg iv) and total body irradiation (1 day 2 Gy). All patients received 40 mg/kg cyclophosphamide intravenously on days +3 and +4, 3x15 mg/kg mycophenolate from day +5 until +28, and tacrolimus titrated at 5-10 ng/ml from day +5 until +84, after which it was tapered with the aim to stop by day +120 or +150, depending on the timing of the first scheduled DLI (i.e., at 4 or 6 months, respectively). Patients had to be off GvHD prophylaxis for at least 2 weeks before a DLI could be administered. Four CMV seropositive patients with a CMV negative donor who were transplanted after October 2021 received letermovir prophylaxis.

In the absence of GvHD requiring tIS, patients considered to have a high risk of early relapse were scheduled to receive a 4-month DLI ($0.3x10^6$ or $0.15x10^6$ T cells/kg in case of related donor [RD] or unrelated donor [UD], respectively). Reasons for prophylactic 4-month DLI were high-risk disease characteristics or incomplete pretransplant treatment. Preemptive 4-month DLI was given if minimal residual disease (MRD) was present at 2 months after alloSCT or in case of rapidly increasing MC between 2 and 4 months after alloSCT. All patients without GvHD requiring tIS, including those who had received a 4-month DLI, were scheduled to receive a prophylactic 6-month DLI, i.e.,

regardless of their anticipated relapse risk and chimerism or MRD status ($3x10^6$ or $1.5x10^6$ T cells/kg, respectively). None of the patients received GvHD prophylaxis after DLI. Patients with persisting or increasing MC or MRD after the 6-month DLI could receive additional preemptive DLIs in escalating doses with a minimum interval of 3 months between DLIs. Patients with insufficient response despite multiple DLIs could receive interferon treatment.

BM chimerism, absolute lymphocyte count, viral infections and definitions of clinical events

BM chimerism, absolute lymphocyte count (ALC) and viral infections were measured and defined as described previously.[2] The three chimerism categories were FDC, low MC (1-4% patient cells), and high MC (≥5% patient cells). The three ALC categories were ALC $<500x10^6$/l, ALC between 500 and $999x10^6$/l and ALC $\geq 1000x10^6$/l. All viral infections confirmed by PCR that occurred within 1 week before and2 weeks after DLI without any prior relapse, second DLI or tIS were considered. Relapse was defined as recurrence of at least 5% blasts on cytomorphologic BM examination, at least 1% blasts in the peripheral blood or the development of extramedullary disease. Clinically relevant GvHD was defined as GvHD for which tIS was administered for at least 14 days.[2]

Analyses

Chimerism response after DLI was evaluated as described previously[19]. Briefly, an algorithm was used to assess the BM chimerism response after DLI in all patients who had MC at the time of their first DLI. A complete response was defined as conversion to FDC, and a partial response as a relative decrease in patient chimerism of 50% or an absolute decrease of 20%, 10% or 5% depending on the level of patient chimerism at the time of first DLI: ≥50%, 20-50% or <20% MC.

The cumulative incidence of clinically relevant GvHD was calculated using a competing risks model starting at the time of first DLI with start of tIS as event of interest and relapse and death as competing events.

The current GvHD-relapse free survival (cGRFS) was calculated using two time-inhomogeneous Markov multi-state models starting at time of alloSCT (total cohort, Supplemental Figure 1) or first DLI (DLI cohort, Supplemental Figure 2). cGRFS was introduced by Solomon et al. and takes into account that patients can recover from GvHD, providing a more accurate measure of long-term treatment success than the GvHD-relapse free survival.[20] However, the cGRFS defined by Solomon et al. only considers moderate-severe chronic GvHD. To get insight in the total burden of clinically relevant GvHD, we considered the use of tIS for any GvHD instead.[21]

In a multi-state model, patients move between states at the occurrence of clinical events. In the absence of relapse, patients could move between the states 'tIS for GvHD' and 'cGRFS' based on whether and when they used tIS for GvHD. From both states, patients could move to the states 'relapse' at time of relapse and 'non-relapse mortality' at time of death without relapse. The 'relapse' and 'non-relapse mortality' states were absorbing, meaning that patients could never leave these states; the probabilities of these two states represent the respective cumulative incidences. As long as no event occurred, patients remained in their current state until end of follow-up.

All analyses were performed in R version 4.4.0 using the packages prodlim[22], mstate[23], ggplot2[24], ggalluvial[25] and ComplexUpset[26].

RESULTS

Cohort

108 patients were included in this study. At 2 years after alloSCT, the cGRFS was 66% (95%-confidence interval [95%-CI] 57-77) and the cumulative incidences of relapse and non-relapse mortality were 22% (95%-CI 15-33) and 7% (95%-CI 3-13), respectively (Supplemental Figure 3). 83 patients were included in the DLI analyses: 37 received the low-dose 4-month DLI and 46 the 6-month DLI as first DLI (Table 1). The other 25 patients did not receive a standard DLI because of early relapse (n=9), GvHD (n=7, of whom 6 required tIS), death without relapse or tIS (n=4), or (temporary) donor unavailability (n=5).

For the total DLI cohort, the cGRFS was 79% (95%-CI 70-89%) at 2 years after the first DLI. At this time, the probability of using tIS was 4% (95%-1-12) and the cumulative incidences of relapse and non-relapse mortality were 14% (95%-CI 8-26%) and 3% (95%-CI 1-10%), respectively (Figure 1).

Conditions at time of DLI

First, we examined the risk factors for GvHD that we had identified previously in the setting of DLI after alemtuzumab-based TCD: viral infections, BM chimerism (as measure for patient-derived APCs), and lymphopenia at the time of first DLI (Table 2, Supplemental Table 1). Four patients (5%) had viral infections during the week before or first two weeks after DLI. 55 patients (66%) had FDC at the time of DLI and only 5 (6%) had MC with ≥5% patient cells. Minimum ALC at the time of DLI was $477x10^6$ cells/l; 17% of the patients had lymphopenia of $<1000x10^6$ lymphocytes/l.

Alloreactivity after DLI

We then investigated the development of GvHD after DLI. Only 5 patients developed clinically relevant GvHD after DLI, resulting in a cumulative incidence of 7% (95%-CI 1-14%) at 2 years after the first DLI. None of the 5 GvHD patients had lymphopenia or a viral infection close to DLI (Supplemental Table 2). Two patients had mixed BM chimerism, 1% and 14% patient cells, at the time of their 6-month DLI. The latter had also received a 4-month DLI while having 12% MC, but did not have any GvHD symptoms until 1 month after the 6-month DLI, after which grade 4 acute GvHD developed. Despite tIS including prednisone and ruxolitinib, this patient died from GvHD 4 months after the 6-month DLI. The other three patients developed GvHD after receiving a DLI from an UD, of whom two with a 9/10 HLA-matched donor.

To investigate whether DLI could induce conversion from MC to FDC, we examined the BM chimerism kinetics of the subset of patients with ≥1% MC at the time of DLI during the first year after DLI (n = 28, Figure 2). 22 patients (79%) converted to FDC, including the two patients with MC who developed clinically relevant GvHD. One of the other complete responders received interferon before conversion. There were no relapses or

deaths during follow-up in the complete responders except the patient with lethal GvHD (median follow-up since their first DLI: 15 months, interquartile range 12-20). Six patients did not convert to FDC: 4 relapsed and 2 did not relapse before censoring at 14

	DLI cohort (N = 83)
Age at the time of first DLI (years)	
median (range)	60 (20-77)
Sex	
Male	50 (60%)
Female	33 (40%)
Disease	
AML*	63 (76%)
ELN adverse risk	34
ELN intermediate risk	15
ELN favorable risk (reason alloSCT: MRD+, no CR after first remission induction course, MRD+ after 2 remission induction courses)	9
relapsed AML	5
ALL	11 (13%)
B-ALL with t(9;22)	3
B-ALL, NOS	5
T-ALL	3
MDS-EB2	9 (11%)
Conditioning	
MAC: thiotepa, Flu and Bu	19 (23%)
MAC: Cy and TBI	1 (1%)
RIC: Flu, Cy and TBI	63 (76%)
Interval between stop GvHD prophylaxis and first DLI (days)	
4-month DLI patients: median (range)	33 (15-89)
6-month DLI patients: median (range)	71 (33-145)
Donor	
10/10 HLA-matched RD	15 (18%)
10/10 HLA-matched UD	50 (60%)
9/10 HLA-matched UD	17 (20%)
8/10 HLA-matched UD	1 (1%)
CMV serostatus patient/donor	
+/+	34 (41%)
+/-	9 (11%)
-/+	7 (8%)
-/-	33 (40%)
EBV serostatus patient/donor	
+/+	67 (81%)
+/-	8 (10%)
-/+	4 (5%)
-/-	4 (5%)

Table 1. Baseline characteristics of the 83 patients in the DLI cohort. DLI, donor lymphocyte infusion; AML, acute myeloid leukemia; alloSCT, allogeneic stem cell transplantation; MRD, minimal residual disease; ALL, acute lymphoblastic leukemia; MDS-EB2, myelodysplastic syndrome with excess blasts; MAC, myeloablative conditioning; RIC, reduced-intensity conditioning; Flu, fludarabine; Bu, busulfan; Cy, cyclophosphamide; TBI, total body irradiation; RD, related donor; UD, unrelated donor; CMV, cytomegalovirus; EBV Epstein-Barr virus. *AML risk scores are based on the 2022 ELN risk classification.

months after the first DLI. Notably, only 3 of 17 patients receiving the 4-month DLI converted before the 6-month DLI was administered. Together, these data show a low risk of GvHD following DLI in this transplantation setting (one case of lethal GvHD), but indicate achievement of a meaningful GvL effect in the majority of the patients.

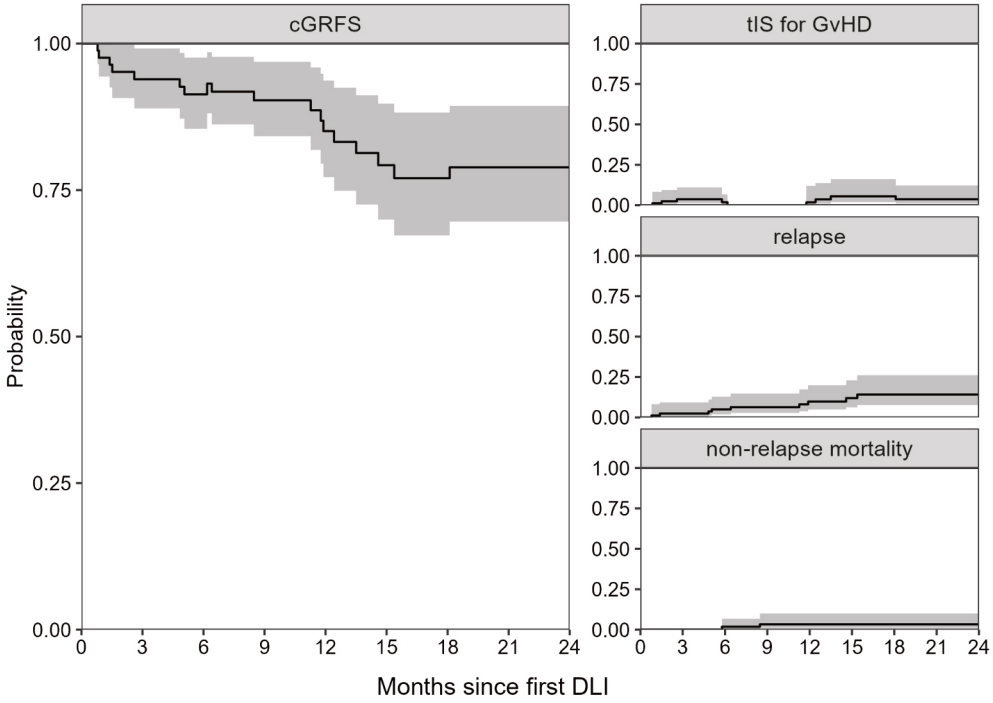


Figure. 1 Probability of cGRFS, current use of tIS for GvHD, relapse and non-relapse mortality for all patients receiving DLI (n = 83). Outcome of the multi-state model over time since first DLI. The 'relapse' and 'non-relapse mortality' states are absorbing: these curves represent cumulative incidences. The structure of the model is shown in Supplemental Figure 2.

	4-month DLI (N = 37)	**6-month DLI (N = 46)**
Viral infection within 1 week before until 2 weeks after DLI		
Yes	5%	4%
No	95%	96%
BM chimerism		
High mixed chimerism: ≥5% patient cells	14%	0%
Low mixed chimerism: 1-4% patient cells	32%	24%
Full donor: <1% patient cells	54%	76%
Absolute lymphocyte count		
$<500x10^6$/l	3%	0%
$500\text{-}999x10^6$/l	16%	15%
$\geq 1000x10^6$/l	81%	85%

Table 2. Presence of viral infections, mixed BM chimerism and lymphopenia at the time of first DLI. DLI, donor lymphocyte infusion; BM, bone marrow

DISCUSSION

The low risk of clinically relevant GvHD after DLI following PTCY alloSCT from HLA-matched and HLA-mismatched donors observed in our study and by others[15,16] shows that application of non- haploidentical DLI after PTCY is relatively safe. The 4% cumulative incidence at 3 months is strikingly lower than the 30% we observed after DLI following alemtuzumab-based TCD alloSCT.[2] Against the background of our previous study[2], infrequent occurrence of DLI-induced GvHD after non-haploidentical PTCY alloSCT can be explained by the relatively high prevalence of FDC at the time of DLI, absence of deep lymphopenia, and low incidence of viral infections around the time of DLI. A combined interpretation of the results of this study and our study on DLI after alemtuzumab-based TCD[2] indicates that transplantation strategies have a profound impact on the conditions at the time of DLI, which in turn influence the alloreactive potential of DLI. The impact of the conditioning regimen on DLI alloreactivity was also noted by Shanmugasundaram et al., who observed GvHD only in patients who received alemtuzumab or anti-thymocyte globulin in addition to the PTCY.[16]

In both our studies[2], none of the FDC patients receiving DLI developed lethal GvHD. Together with the results of a matched-pair analysis by Schmid et al.[27], this demonstrates the safety of prophylactic non-haploidentical DLI. The patient with high MC developing lethal DLI-induced GvHD illustrates the relevance of high presence of patient-derived

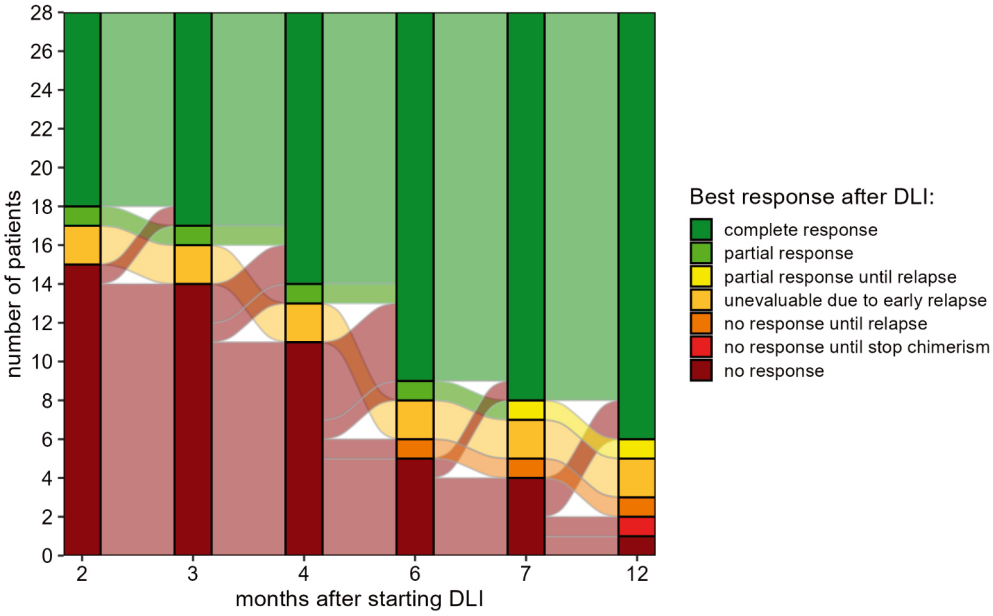


Figure. 2 BM chimerism response after DLI for the patients with mixed chimerism at the time of first DLI (n = 28). The best BM chimerism response achieved at different time points after the first DLI (complete response: conversion to full-donor chimerism, partial response: decreasing mixed chimerism, no response: stable/increasing mixed chimerism). Two patients relapsed before the first chimerism measurement after DLI (relapse at 0.8 and 1.4 months after low-dose 4-month DLI) and two patients relapsed before reaching a complete response (relapse at 4.8 and 6.4 months after low-dose 4-month DLI, both also received the 6-month DLI before relapse). One other patient converted to full-donor chimerism after start of interferon. Events after reaching a complete response are not shown.

APCs and shows that under certain conditions DLI after PTCY can induce lethal GvHD. Of the other patients developing GvHD after DLI, 2 had an HLA mismatch. While PTCY may reduce the impact of having an HLA mismatch on the GvHD risk after non-haploidentical alloSCT[28], this effect is likely smaller when fresh alloreactive lymphocytes are infused by a DLI several months thereafter when the degree of genetic disparity may play an important role in the development of GvHD. The low number of GvHD cases in our cohort did not allow us to estimate the effect sizes of mixed chimerism and HLA mismatch on the risk of GvHD.

The low-dose 4-month DLI after PTCY alloSCT rarely induced chimerism conversion or clinically relevant GvHD, suggesting limited alloreactive potential in contrast to the 3-month low-dose DLI after alemtuzumab-based TCD alloSCT.[2,19] This is likely due to the different conditions at the time of DLI: the faster lymphocyte recovery after PTCY alloSCT compared to alemtuzumab-based TCD leads to less viral infections and therefore less inflammation during the months after alloSCT. Combined with the later timing of the low-dose DLI after PTCY alloSCT (one month later than after alemtuzumab-based TCD alloSCT), this leads to a less pro-inflammatory environment at the time of DLI and a low alloreactive potential of the 4-month DLI after PTCY alloSCT with the current DLI dose. However, the total DLI strategy led to similar conversion rates for DLI after alemtuzumab or PTCY.[19] In both settings[19], conversion from MC to FDC after DLI occurred often in the absence of clinically relevant GvHD, but the GvHD/GvL balance seems to be better in the PTCY setting: the doses of the 6-month DLI and any subsequent DLI were apparently sufficient to induce chimerism conversion, but with a lower GvHD risk than in the alemtuzumab setting. This supports the conclusions of Van Bergen et al. that whether or not GvL is companied with GvHD not only depends on the diversity of the alloreactive T cells but also on the inflammatory conditions.[29] Differences in the timing and doses of DLI and the conditions at the time of infusion might explain why the chimerism conversion rates in our studies differ from those reported by Shanmugasundaram et al..[16]

The aim of prophylactic and preemptive DLI is to prevent relapse without causing excessive toxicity. With our total strategy of TCD alloSCT followed by standard DLI, the 2-year cumulative incidence of relapse was 22%. This is still close to the estimates reported in studies on non-haploidentical PTCY alloSCT for acute leukemia without DLI, which range from 19% in a single-center study to 28% in a 9/10 HLA-matched UD registry cohort.[28,30,31] The 2-year non-relapse mortality in our study (7%) seems be a bit lower than in the other studies (15-20%).[28,30,31] Comparing studies is notoriously difficult because of differences in transplantation strategy and characteristics of the patients, diseases and donors. However, in our study none of the patients who converted to FDC after DLI experienced relapse, indicating that a meaningful GvL effect was achieved. Together with the low toxicity, this strongly suggests that application of DLI after non-haploidentical PTCY alloSCT can have a beneficial clinical effect. In our cohort, about half of the relapsing patients relapsed between 3 and 6 months after alloSCT. Considering the low toxicity and efficacy of our 4-month low-dose DLI, it might be possible to increase the dose of this DLI or to administer the current dose at an earlier time to reduce the relapse risk during this period without inducing severe GvHD.

A limitation of our study is that we do not have a control group of patients not receiving standard DLI. Since most alloreactivity was observed after the 6-month DLI, several

months after cessation of double GvHD prophylaxis, we assume that the observed alloreactivity is DLI-induced, but cannot rule out some effect of the tapering of GvHD prophylaxis. However, after PTCY alloSCT combined with double GvHD prophylaxis using HLA-mismatched donors and no DLI, Soltermann et al. observed a cumulative incidence of only 15% acute GvHD grade II-IV, predominantly occurring during the first 2 months.[32] This suggests that the GvHD we observed after the 4- and 6-month DLI is most likely related to the DLI. We are currently planning a clinical trial to investigate the optimal timing and dose of prophylactic DLI and to compare alloSCT with or without standard prophylactic DLI.

In conclusion, our data show that non-haploidentical prophylactic and preemptive DLI following PTCY alloSCT give a low risk of clinically relevant GvHD but still a meaningful GvL effect. The conditions in which DLI are more likely to induce severe GvHD are known. Careful tailoring the DLI dose to the conditions at the time of the DLI could therefore improve the balance between GvHD and GvL and increase the safety and efficacy of DLI.

ACKNOWLEDGEMENTS

Eva Koster has a Landkroon fellowship from the Doelfonds Leukemie of the Bontius Stichting. The authors would like to thank E.J.M. Bogers, S.M. Wallet-Malicka and H.A. Baelde for the data management and technical support.

REFERENCES

1. Falkenburg, J. H. F., Schmid, C., Kolb, H. J. & Kuball, J. in The EBMT Handbook: Hematopoietic Cell Transplantation and Cellular Therapies (eds Anna Sureda et al.) 531-539 (Springer International Publishing, 2024).
2. Koster, E. A. S., von dem Borne, P. A., van Balen, P. et al. Risk factors for graft-versus-host-disease after donor lymphocyte infusion following T-cell depleted allogeneic stem cell transplantation. Frontiers in Immunology 15, 1335341 (2024). doi: 10.3389/fimmu.2024.1335341
3. Chakraverty, R. & Sykes, M. The role of antigen-presenting cells in triggering graft-versus-host disease and graft-versus-leukemia. Blood 110, 9-17 (2007). doi: 10.1182/blood-2006-12-022038
4. van Balen, P., van der Zouwen, B., Kruisselbrink, A. B. et al. Tissue Damage Caused by Myeloablative, but Not Non-Myeloablative, Conditioning before Allogeneic Stem Cell Transplantation Results in Dermal Macrophage Recruitment without Active T-Cell Interaction. Front Immunol 9, 331 (2018). doi: 10.3389/fimmu.2018.00331
5. Thiant, S., Labalette, M., Trauet, J. et al. Plasma levels of IL-7 and IL-15 after reduced intensity conditioned allo-SCT and relationship to acute GVHD. Bone Marrow Transplant 46, 1374-1381 (2011). doi: 10.1038/bmt.2010.300
6. Thiant, S., Yakoub-Agha, I., Magro, L. et al. Plasma levels of IL-7 and IL-15 in the first month after myeloablative BMT are predictive biomarkers of both acute GVHD and relapse. Bone Marrow Transplant 45, 1546-1552 (2010). doi: 10.1038/bmt.2010.13
7. Stevanovic, S., van Bergen, C. A., van Luxemburg-Heijs, S. A. et al. HLA class II upregulation during viral infection leads to HLA-DP-directed graft-versus-host disease after CD4+ donor lymphocyte infusion. Blood 122, 1963-1973 (2013). doi: 10.1182/blood-2012-12-470872
8. Nunes, N. S. & Kanakry, C. G. Mechanisms of Graft-versus-Host Disease Prevention by Post-transplantation Cyclophosphamide: An Evolving Understanding. Front Immunol 10, 2668 (2019). doi: 10.3389/fimmu.2019.02668

9. Zhao, C., Bartock, M., Jia, B. et al. Post-transplant cyclophosphamide alters immune signatures and leads to impaired T cell reconstitution in allogeneic hematopoietic stem cell transplant. Journal of Hematology & Oncology 15, 64 (2022). doi: 10.1186/s13045-022-01287-3
10. Ciurea, S. O., Mulanovich, V., Saliba, R. M. et al. Improved early outcomes using a T cell replete graft compared with T cell depleted haploidentical hematopoietic stem cell transplantation. Biol Blood Marrow Transplant 18, 1835-1844 (2012). doi: 10.1016/j.bbmt.2012.07.003
11. Tischer, J., Engel, N., Fritsch, S. et al. Virus infection in HLA-haploidentical hematopoietic stem cell transplantation: incidence in the context of immune recovery in two different transplantation settings. Ann Hematol 94, 1677-1688 (2015). doi: 10.1007/s00277-015-2423-y
12. Montoro, J., Roldán, E., Piñana, J. L. et al. Ex vivo T-cell depletion vs post-transplant cyclophosphamide, sirolimus, and mycophenolate mofetil as graft-vs-host disease prophylaxis for allogeneic hematopoietic stem cell transplantation. European journal of haematology 106, 114-125 (2021). doi: 10.1111/ejh.13529
13. Rappazzo, K. C., Zahurak, M., Bettinotti, M. et al. Nonmyeloablative, HLA-Mismatched Unrelated Peripheral Blood Transplantation with High-Dose Post-Transplantation Cyclophosphamide. Transplant Cell Ther 27, 909.e901-909.e906 (2021). doi: 10.1016/j.jtct.2021.08.013
14. Dholaria, B., Savani, B. N., Labopin, M. et al. Clinical applications of donor lymphocyte infusion from an HLA-haploidentical donor: consensus recommendations from the Acute Leukemia Working Party of the EBMT. Haematologica 105, 47-58 (2020). doi: 10.3324/haematol.2019.219790
15. Carnevale-Schianca, F., Caravelli, D., Gallo, S. et al. Post-Transplant Cyclophosphamide and Tacrolimus-Mycophenolate Mofetil Combination Governs GVHD and Immunosuppression Need, Reducing Late Toxicities in Allogeneic Peripheral Blood Hematopoietic Cell Transplantation from HLA-Matched Donors. Journal of clinical medicine 10 (2021). doi: 10.3390/jcm10061173
16. Shanmugasundaram, K., Napier, S., Dimitrova, D. et al. Safety but limited efficacy of donor lymphocyte infusion for post-transplantation cyclophosphamide-treated patients. Bone Marrow Transplant 59, 1513-1524 (2024). doi: 10.1038/s41409-024-02312-4
17. Castagna, L., Sarina, B., Bramanti, S. et al. Donor lymphocyte infusion after allogeneic stem cell transplantation. Transfus Apher Sci 54, 345-355 (2016). doi: 10.1016/j.transci.2016.05.011
18. Ros-Soto, J., Snowden, J. A., Szydlo, R. et al. Outcomes After Donor Lymphocyte Infusion in Patients With Hematological Malignancies: Donor Characteristics Matter. Transplant Cell Ther 28, 183.e181-183.e188 (2022). doi: 10.1016/j.jtct.2022.01.022
19. Koster, E. A. S., von dem Borne, P. A., van Balen, P. et al. Competitive Repopulation and Allo-Immunologic Pressure Determine Chimerism Kinetics after T Cell-Depleted Allogeneic Stem Cell Transplantation and Donor Lymphocyte Infusion. Transplant Cell Ther 29, 268.e261-268.e210 (2023). doi: 10.1016/j.jtct.2022.12.022
20. Solomon, S. R., Sizemore, C., Zhang, X. et al. Current Graft-versus-Host Disease-Free, Relapse-Free Survival: A Dynamic Endpoint to Better Define Efficacy after Allogenic Transplant. Biol Blood Marrow Transplant 23, 1208-1214 (2017). doi: 10.1016/j.bbmt.2017.02.022
21. Eefting, M., de Wreede, L. C., Halkes, C. J. et al. Multi-state analysis illustrates treatment success after stem cell transplantation for acute myeloid leukemia followed by donor lymphocyte infusion. Haematologica 101, 506-514 (2016). doi: 10.3324/haematol.2015.136846
22. Gerds, T. A. prodlim: Product-Limit Estimation for Censored Event History Analysis, <https://cran.r-project.org/package=prodlim> (2019).
23. De Wreede, L. C., Fiocco, M. & Putter, H. mstate: An R Package for the Analysis of Competing Risks and Multi-State Models. Journal of Statistical Software 38, 1-30 (2011). doi: 10.18637/jss.v038.i07
24. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. (Springer-Verlag New York, 2016).
25. Brunson, J. C. & Read, Q. D. ggalluvial: Alluvial Plots in 'ggplot2'. <http://corybrunson.github.io/ggalluvial/> (2023).
26. Krassowski, M. ComplexUpset, <https://doi.org/10.5281/zenodo.3700590> (2020).
27. Schmid, C., Labopin, M., Schaap, N. et al. Prophylactic donor lymphocyte infusion after allogeneic stem cell transplantation in acute leukaemia - a matched pair analysis by the Acute Leukaemia Working Party of EBMT. Br J Haematol 184, 782-787 (2019). doi: 10.1111/bjh.15691

28. Lorentino, F., Labopin, M., Ciceri, F. et al. Post-transplantation cyclophosphamide GvHD prophylaxis after hematopoietic stem cell transplantation from 9/10 or 10/10 HLA-matched unrelated donors for acute leukemia. Leukemia 35, 585-594 (2021). doi: 10.1038/s41375-020-0863-4
29. van Bergen, C. A., van Luxemburg-Heijs, S. A., de Wreede, L. C. et al. Selective graft-versus-leukemia depends on magnitude and diversity of the alloreactive T cell response. The Journal of clinical investigation 127, 517-529 (2017). doi: 10.1172/jci86175
30. Moiseev, I. S., Pirogova, O. V., Alyanski, A. L. et al. Graft-versus-Host Disease Prophylaxis in Unrelated Peripheral Blood Stem Cell Transplantation with Post-Transplantation Cyclophosphamide, Tacrolimus, and Mycophenolate Mofetil. Biol Blood Marrow Transplant 22, 1037-1042 (2016). doi: 10.1016/j.bbmt.2016.03.004
31. Nagler, A., Labopin, M., Dholaria, B. et al. Comparison of Haploidentical Bone Marrow versus Matched Unrelated Donor Peripheral Blood Stem Cell Transplantation with Posttransplant Cyclophosphamide in Patients with Acute Leukemia. Clin Cancer Res 27, 843-851 (2021). doi: 10.1158/1078-0432.Ccr-20-2809
32. Soltermann, Y., Heim, D., Medinger, M. et al. Reduced dose of post-transplantation cyclophosphamide compared to ATG for graft-versus-host disease prophylaxis in recipients of mismatched unrelated donor hematopoietic cell transplantation: a single-center study. Ann Hematol 98, 1485-1493 (2019). doi: 10.1007/s00277-019-03673-3

SUPPLEMENTAL TABLES

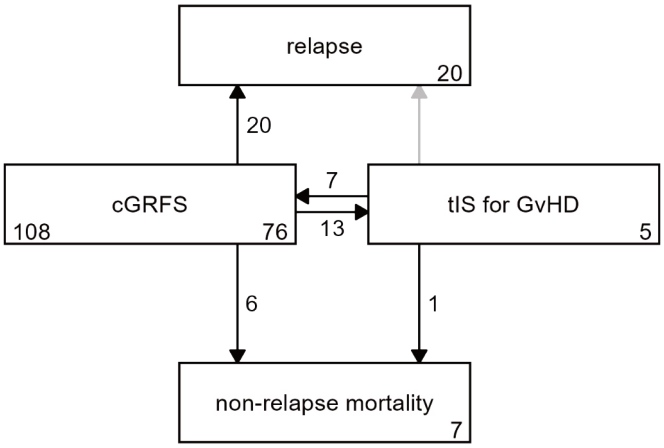
		BM chimerism: FDC	BM chimerism: 1-4% MC	BM chimerism: ≥5% MC
No viral infection close to DLI*	ALC ≥1000	43	18	4
	ALC 500-999	10	2	1
	ALC <500	0	1	0
Viral infection close to DLI*	ALC ≥1000	2 (both COVID-19)	2 (CMV, rhinovirus)	0
	ALC 500-999	0	0	0
	ALC <500	0	0	0

Supplemental Table 1. Numbers of patients for each combination of characteristics of BM chimerism, ALC and viral infection. BM, bone marrow; FDC, full-donor chimerism; MC, mixed chimerism; ALC, absolute lymphocyte count (x10^6/l); DLI, donor lymphocyte infusion. *Within 1 week before until 2 weeks after DLI

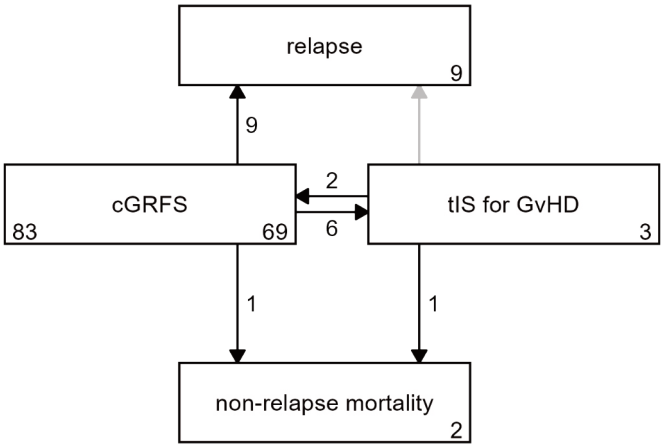
Donor	Last DLI before onset of GvHD	BM chimerism at time of DLI	ALC at time of DLI	Viral infection close to DLI	GvHD requiring tIS*	Outcome
MM UD	4-month DLI	FDC	≥1000	No	• aGVHD liver grade 1 • extensive cGVHD liver, muscles	Resolved
MM UD	6-month DLI after 4-month DLI	FDC	≥1000	No	• extensive cGVHD eyes, nails	Ongoing tIS 1 year after DLI
UD	6-month DLI	FDC	≥1000	No	• Extensive cGVHD lungs, muscles	Ongoing tIS 1 year after DLI
RD	6-month DLI	1% MC	≥1000	No	• Extensive cGvHD skin	Ongoing tIS 9 months after DLI
RD	6-month DLI after 4-month DLI	12% MC	≥1000	No	• aGVHD skin, liver, GI grade 4	Death from GvHD

Supplemental Table 2. Characteristics of the 5 patients who developed GvHD after DLI. MM, HLA-mismatched (else: 10/10 HLA-matched); RD, related donor; UD, unrelated donor; BM, bone marrow; FDC, full-donor chimerism; MC, mixed chimerism; ALC, absolute lymphocyte count; aGvHD, acute GvHD; cGvHD, chronic GvHD; tIS, therapeutic systemic immunosuppression. *Grading of acute and chronic GvHD according to the modified Glucksberg and the Seattle criteria, respectively.

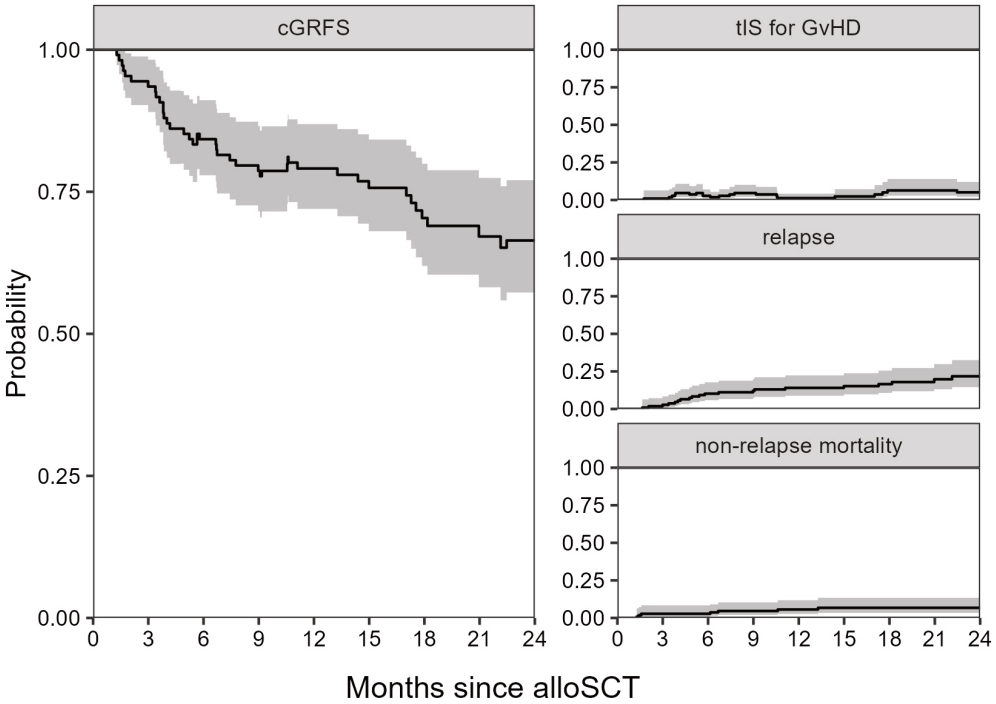
SUPPLEMENTAL FIGURES



Supplemental Figure 1. Multi-state model starting from alloSCT (total cohort). Boxes represent states and arrows represent the transitions between the states. The grey transition was not used by any of the included patients. All patients started in the state 'cGRFS' at the time of alloSCT. The number at the bottom left corner of the starting state shows the number of patients included in the model. The numbers at the bottom right corner of the boxes show the numbers of the patients who were in that state at the end of their follow-up. The numbers next to the arrows show the numbers of the patients who made that transition during their follow-up.

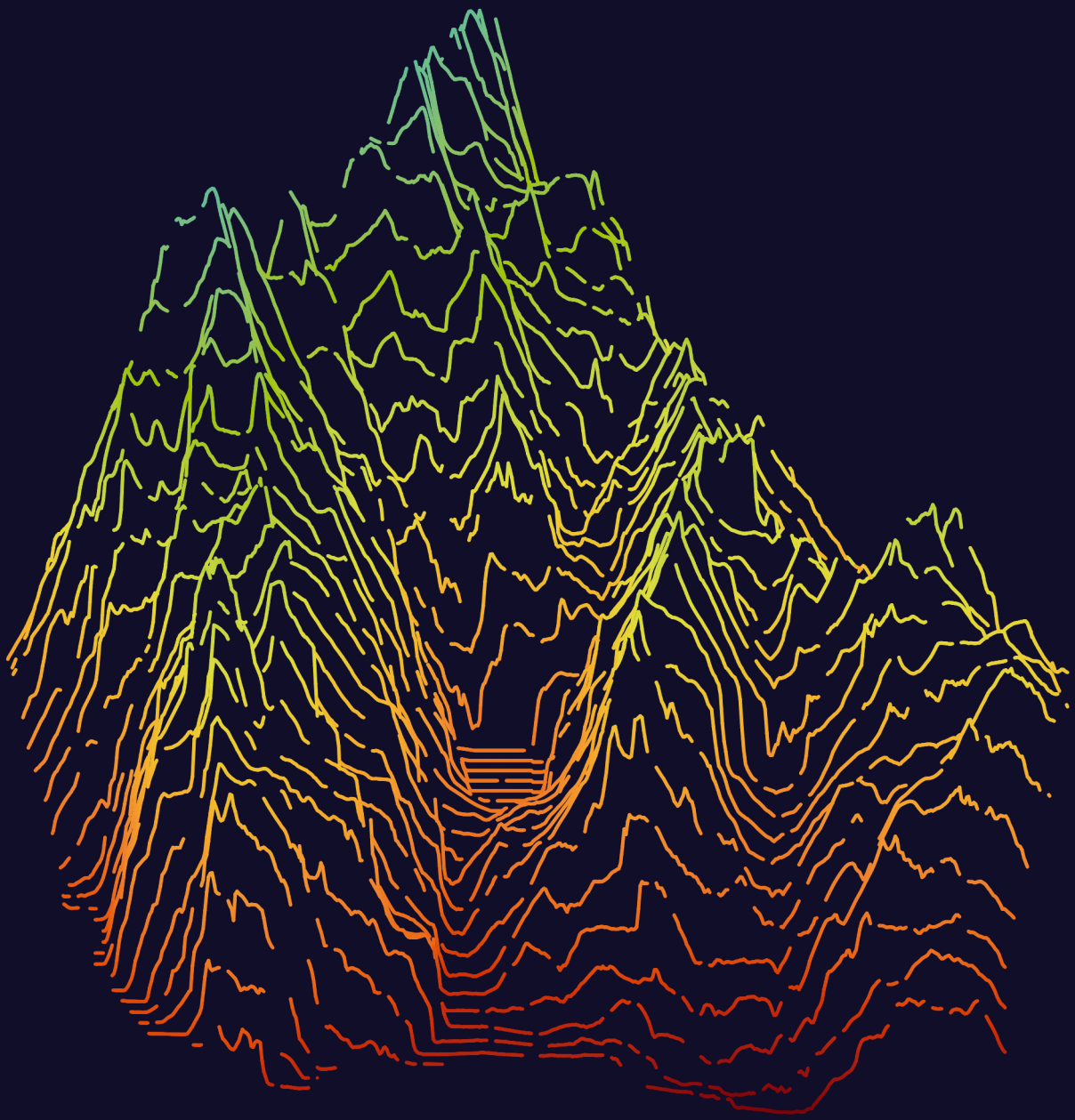


Supplemental Figure 2. Multi-state model starting from first DLI (DLI cohort). Boxes represent states and arrows represent the transitions between the states. The grey transition was not used by any of the included patients. All patients started in the state 'cGRFS' at the time of their first DLI. The number at the bottom left corner of the starting state shows the number of patients included in the model. The numbers at the bottom right corner of the boxes show the numbers of the patients who were in that state at the end of their follow-up. The numbers next to the arrows show the numbers of the patients who made that transition during their follow-up.



Supplemental Figure 3. Probability of cGRFS, current use of tIS for GvHD, relapse and non-relapse mortality after alloSCT (total cohort). The 'relapse' and 'non-relapse mortality' states are absorbing: these curves represent cumulative incidences. The structure of the model is shown in Supplemental Figure 1.

5



6

Anti-thymocyte globulin-based treatment frequently leads to enduring treatment success in both old and young adult patients with aplastic anemia: a real-world analysis from the Dutch Aplastic Anemia Registry

Constantijn J.M. Halkes*, Eva A.S. Koster*, Lisette J.M. Bogers, Floor C.J.I. Heubel-Moenen, Laura G.M. Daenen, Saskia K. Klein, Saskia M.C. Langemeijer, Erfan Nur, Marc H.G. Raaijmakers, Tjeerd J.F. Snijders, Jennifer M.L. Tjon, Liesbeth C. de Wreede

* shared first authorship

In submission

ABSTRACT

Discussion remains concerning the safety and tolerability of anti-thymocyte globulin (ATG)-based immunosuppressive therapy (IST) in older patients with aplastic anemia (AA). Using data of 127 consecutive patients from the Dutch adult AA registry, we evaluated long-term treatment success of standard ATG-based IST as first-line treatment with a multi-state model. Only one death was potentially associated with ATG. We defined Transplantation-, Treatment- and Disease-Free Survival (TT-DFS) as the ultimate success of this treatment. This means that a patient is alive, is currently transfusion-independent without having received an allogeneic stem cell transplantation, has not developed AML or MDS, and has stopped all medication for AA. The probability of TT-DFS was 42% at 5 years after start of IST. In patients younger than 40 years (n=36) and in patients aged 60 or above (n=53), this was 58% and 34%, respectively. Older age, more severe AA and absence of a PNH-clone of ≥1% all reduced the likelihood of reaching TT-DFS. These analyses on unselected nationwide data indicate that ATG-based IST is effective and safe also in older patients. They suggest that age, AA severity and presence of a PNH-clone should be taken into account when considering this treatment in older patients.

INTRODUCTION

Acquired aplastic anemia (AA) is a rare hematological disease characterized by pancytopenia and a hypocellular bone marrow. The exact pathogenesis is unknown but most likely involves an autoimmune reaction.[1,2] The preferred first-line treatment in the majority of adult patients is intensive immunosuppressive therapy (IST) based on anti-thymocyte globulin (ATG) and ciclosporin (CsA) with or without eltrombopag.[3] As this treatment is associated with an increased risk of acute cardiac toxicity and infusion reactions and conflicting results are published concerning its effectivity in patients aged 60 years or older, there is debate whether ATG-based IST should be offered as preferred first-line treatment to older AA patients instead of a less intensive treatment.[3-6]

Recently, an analysis based on two single-center clinical trials which ran between 2005 and 2022, showed that acute toxicity after IST, response rate at six months and overall survival (OS) were similar in responding patients aged below 60 and of at least 60 years.[5] The authors concluded that ATG-based IST should be the preferred first-line treatment in all AA patients aged 60 and older. In contrast, another analysis on patients treated between 1976 and 2024 in 4 tertiary hematologic centers showed an inferior response rate after ATG-based IST in AA patients aged 60 and older compared to patients younger than 60. The authors concluded that older patients had an unfavorable risk/benefit ratio for ATG-based therapy and that this treatment should only be offered to AA patients younger than 60 and to patients between 60 and 65 years without comorbidities.[6]

When considering ATG-based IST as first-line treatment, not only the risk of severe side effects in older patients with comorbidities, but also the probability to achieve an enduring response should be taken into account. Hematological recovery at six months after start of IST is commonly used as a measure for treatment success of IST. Younger age, non-severe AA and the presence of a glycosyl phosphatidylinositol (GPI)-deficient cell clone (Paroxysmal Nocturnal Hemoglobinuria (PNH)-clone) at time of diagnosis are associated with reaching this short-term treatment success in several studies.[7-9] Long-term treatment success and survival after IST in patients with AA is hampered by disease related mortality due to infections or bleedings, toxicity of second-line treatments like allogeneic stem cell transplantation (alloSCT) and the development of other bone marrow diseases like MDS or AML. The most common treatment failure is relapse of AA, after which a renewed response can often be achieved by reintroduction of IST. These complex dynamics that are all relevant for achieving and maintaining long-term IST treatment success cannot be captured by the fixed 6-month timepoint for hematological response and long-term OS.

Therefore, we propose a comprehensive measure of treatment success over time after ATG-based IST for AA. The main component of this measure is Disease-free survival (DFS): being alive and currently transfusion-independent without having developed other bone marrow diseases like MDS or AML. We consider this the broad aim of ATG-based IST. DFS should preferably be achieved without needing an alloSCT as second-line treatment: Transplantation-free DFS (T-DFS). The ultimate aim of IST is to stop all medication after achieving a response, leading to the endpoint Transplantation- and Treatment-free DFS (TT-DFS). All these endpoints are dynamic, i.e. during time since start of treatment they are continuously updated when responses are gained, lost and

regained or other relevant events take place. These complex endpoints can be assessed by multi-state models, which can also serve to investigate the impact of risk factors on different steps of the recovery/disease process. We used DFS, T-DFS and TT-DFS to evaluate treatment success in patients from the Dutch adult aplastic anemia registry who were treated uniformly with horse-derived ATG (ATGAM)-based IST. This registry describes a unique real-world nationwide cohort containing data of 144 consecutive adult AA patients who were treated with ATGAM-based IST in the Netherlands between 2012 and 2021.

MATERIALS AND METHODS

Patients and treatment

The Dutch national AA registry is a population-based registry in which data is collected from all consecutive patients in the Netherlands who received ATGAM in combination with CsA as first-line treatment for AA according to the national guideline[10] (Supplemental Methods). Treatment is given in a limited number of Dutch hospitals, that all contribute data to the registry (Supplemental Table 1). Completeness of patient cohorts per hospital was checked using delivery data of ATGAM to the hospital pharmacies. Clinical data and laboratory results were collected by the treating hematologists or local data managers at baseline and at regular intervals after start of ATGAM and checked by a central data manager. The Institutional review board of the Leiden University Medical Centre approved data collection and analysis (protocol nr. C13.014). The dataset was closed on July 1st 2022.

Definitions and endpoints

AA diagnosis was classified according to the Camitta criteria.[11] Transfusion independency was defined as having been free of transfusion for at least 4 weeks. Non-transplant therapy was defined as any treatment to improve hematopoiesis in AA except alloSCT. This could include multiple courses of ATG, CsA, other immunosuppressive drugs or drugs with another working mechanism like eltrombopag or Danazol. Disease-free survival was defined as being alive and currently transfusion-independent without having started treatment for a secondary bone marrow disease, i.e., AML or MDS. Transplantation-free DFS was defined as DFS without having received an alloSCT. Treatment- and Transplantation-free DFS was defined as T-DFS without having received any non-transplant therapy for at least 2 weeks. The primary endpoint of this analysis was the probability of TT-DFS until 5 years after start of ATGAM. Secondary endpoints were the cumulative incidence of achieving transfusion independency and probabilities of DFS, T-DFS and OS.

Statistical analysis

OS from start of first IST with 95% confidence intervals (95%-CI) was calculated using the Kaplan-Meier method. Follow-up was quantified using the reverse Kaplan-Meier method which censors patients when they die.[12] The cumulative incidence of achieving transfusion independency after first-line treatment was estimated using a competing risks model with death and start of second-line treatment or treatment for AML or MDS as

competing events. The cumulative incidence of achieving transfusion independency regardless of the number of treatments was estimated using a competing risks model with death and treatment for AML or MDS as competing events. To estimate DFS, T-DFS and TT-DFS after IST and to identify which factors influence the treatment response, a Markov time-inhomogeneous multi-state model was constructed. In a multi-state model patients transit between states at the occurrence of clinical events. Each transition hazard can either be estimated without taking covariates into account (non-parametrically) or analyzed by means of a transition-specific Cox proportional hazards model (semi-parametric approach). The baseline hazards and the hazard ratios (HRs) are the building blocks for the calculation of the transition probabilities, which represent the probabilities of being in each of the states over time.[13,14]

The structure of the model is shown in Figure 1. Starting state and time of the model is the start of the IST (ATGAM and ciclosporin). As all patients are also transfusion-

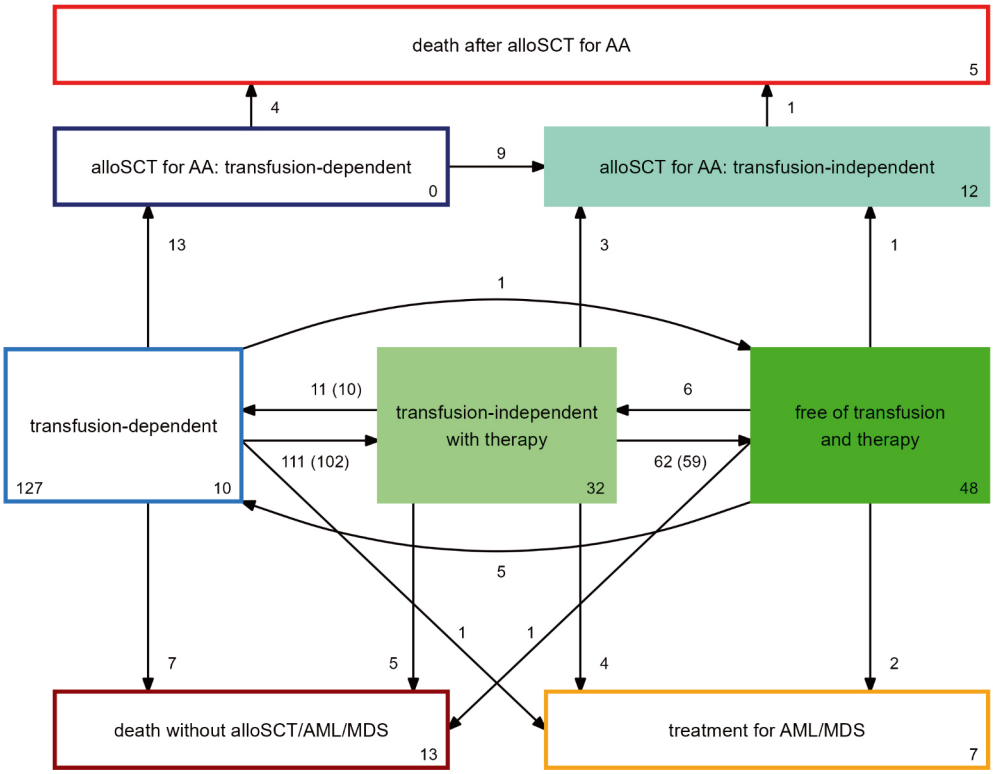


Figure 1. Structure of the multi-state model. Boxes represent states, arrows transitions. All patients start in the state 'transfusion-dependent'. The number in the left corner of this box shows the number of patients included in the non-parametric model. The numbers at the bottom right corner of the boxes show the numbers of patients who were in that state at the end of their follow-up. The numbers next to the arrows show the numbers of transitions that were observed during follow-up. A few patients made a transition multiple times during their follow-up. For the transitions for which this was the case, the numbers inside the brackets show the numbers of unique patients who made these transitions. TT-DFS equals the probability of being in the state 'free of transfusion and therapy' (dark green), T-DFS the probability of being either in this state or in the state 'transfusion-independent with therapy' (light green), DFS the probability of being in any of the three filled states.

dependent at that time, this state is called 'transfusion-dependent'. If a patient has been transfusion-independent for 4 weeks, this patient will move to the state 'transfusion-independent with therapy' after these 4 weeks. If a patient in this state has stopped all non-transplant therapy for 2 weeks and remains transfusion-independent, the patient will move to the state 'free of transfusion and therapy'. Transfusion-independent patients who become transfusion-dependent and patients who have to restart non-transplant treatment will move to the appropriate states. From these three states, patients can experience three failure events: requirement of alloSCT, start of treatment for another bone marrow disease (AML or MDS) and death. The latter two are absorbing states, implicating that patients can never leave the state. AlloSCT is split into transfusion dependency after alloSCT, transfusion independency after alloSCT and death after alloSCT to evaluate the outcome after alloSCT as non-first-line treatment.

The effects of patient age, presence of a PNH-clone and severity of the AA on the hazards of several transitions of the multi-state model were estimated using transition-specific Cox models (Supplemental Methods). To translate the obtained HRs into clinically relevant measures, we calculated multi-state model-based outcomes for reference patients with different baseline characteristics. This shows the impact of age, the presence of a PNH-clone and the severity of the disease on T-DFS and TT-DFS.

Confidence intervals for DFS and T-DFS, which are combinations of multiple states, were calculated based on the estimated variance-covariance matrix of the transition probabilities.[14]

Software

All analyses were performed in R version 4.5.0 using the packages survival[15], prodlim[16], mstate[14], ggplot2[17], and ggh4x[18].

RESULTS

The majority of AA patients become transfusion-independent after IST

In total, 144 patients who received ATGAM-based IST between 2012 and 2021 were included in the Dutch AA registry. Seventeen patients received IST with the addition of eltrombopag as first-line treatment in the prospective randomized EBMT RACE study[9]. To preserve a uniform treatment cohort we excluded these 17 patients. Analyses were performed on the remaining 127 patients.

Baseline characteristics of the patients are shown in Table 1. The median follow-up time was 56 months (interquartile range 25-80). Figure 1 shows the number of transitions between the different states. At start of IST, all 127 patients were transfusion-dependent. The cumulative incidence of achieving transfusion independency without any second-line treatment was 64% (95%-CI 55-72) at 1 year after start of ATGAM and CsA (Supplemental Figure 1). For patients who did not become transfusion-independent, second-line treatment was eltrombopag in 21 patients (of whom 15 became transfusion-independent), rabbit-derived ATG (Thymoglobulin) in 7 patients (of whom 3 became transfusion-independent) and 8 patients received an alloSCT. See Supplemental Figure 2 for the respective third- and fourth-line therapies given after ATGAM with CsA. The

cumulative incidence of becoming transfusion-independent regardless of the number of treatments was 88% (95%-CI 82-94) at 2 years after start of IST (Supplemental Figure 1).

Low treatment-related mortality after ATG-based IST

The OS at 5 years after start of IST was 79% (95%-CI 70-87%, Supplemental Figure 3A). Twenty-one patients died within 5 years after start of ATG (5 after alloSCT as non-first-line treatment for AA (median time between start IST and death 12 months) and 3 due to AML or MDS (median time between start IST and death 45 months). Seven patients died due to cytopenia-related complications after a median time of 32 months (causes of death: hemorrhage, pneumonia or refractory aplastic anemia). One patient (age 77 years) died due to cardiac failure ten days after start of IST, which could potentially have been associated with a newly diagnosed atrial fibrillation. All causes of death in the first 5 years after start of IST are shown in the Supplemental Table 2.

At 5 years, 42% of patients are Transplantation-, Treatment- and Disease-free

Figure 2 shows the outcomes for patients during the first 5 years after start of first-line IST. Most patients became transfusion-independent during the first year and from the first year on, patients were also able to stop their medication for AA. Five years after start of IST, 70% (95%-CI 61-81) of the patients were transfusion-independent and had not

	Total cohort
Median Age (range)	54 (18-79) years
Age category- no (%)	
Age 18-39 years	36 (28%)
Age 40-59 years	38 (30%)
Age > 59 years	53 (42%)
Sex- no (%)	
Male	81 (64%)
Female	46 (36%)
Disease severity- no(%)	
Non-severe	42 (33%)
Severe	54 (43%)
Very severe	31 (24%)
PNH-clone at diagnosis-no (%)	
<1%	62 (55%)
≥1%	50 (45%)
Missing	15
Cytogenetic abnormalities- no (%)	
Normal	75/88 (85)
Abnormal karyotype*	8/88 (9)
Karyotype analysis failed	5/88 (6)
Missing	39
Median time between diagnosis and start ATGAM (range)**	1.6 (0.1-191) months

Table 1. Baseline characteristics of the 127 patients with ATGAM-based IST as first-line treatment *Abnormal karyotype included loss of Y chromosome (6 patients), -13q (1 patient) and extra marker chromosome of unknown origin (1 patient) ** Time of diagnosis not known in 13 patients.

developed MDS or AML (DFS). The Transplantation-free DFS (T-DFS) was 60% (95%-CI 51-71) and the Transplantation- and Treatment-free DFS (TT-DFS) was 42% (95%-CI 33-54) (Figure 3A). Supplemental Figure 3B-D and Figure 3B-D show OS, DFS, T-DFS and TT-DFS per age category. Patients aged 18-39 years at start of treatment had a 5-year OS of 97% (95%-CI 91-100), and a TT-DFS of 58% (95%-CI 41-81). For patients aged 40-59, the 5-year OS and TT-free DFS were 78% (95%-CI 63-93) and 38% (95%-CI 22-64), respectively. For patients of 60 years and older, the 5-year OS and TT-free DFS were 65% (95%-CI 48-81) and 34% (95%-CI 22-54), respectively.

Age, disease severity and concomitant PNH-clone at diagnosis predict TT-DFS

Patients with a PNH-clone (>1%) at diagnosis had a HR of 2.2 (95%-CI 1.4-3.4) to become transfusion-independent during AA treatment compared to patients without a PNH-clone. Patients aged 40-59 and older than 59 were less likely to become transfusion-independent compared to younger patients (HR 0.4 [95%-CI 0.2-0.7] and 0.4 [95%-CI 0.3-0.7], respectively). The same was observed for patients with (very)

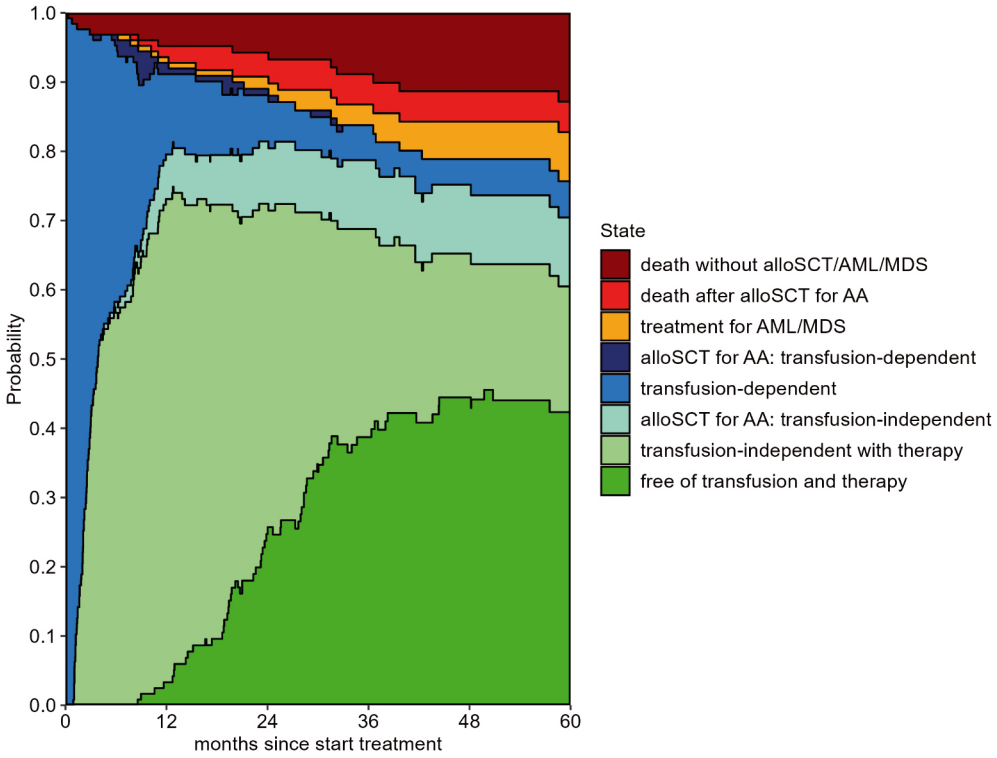


Figure 2. Outcomes after start of IST. Stacked transition probabilities from state 'transfusion-dependent with therapy': the difference between two adjacent curves represents the probability of being in the corresponding state. The probability of being in the state 'free of transfusion and therapy' represents the TT-DFS, the probability of being in this state or in the state 'transfusion-independent with therapy' the T-DFS. DFS is the sum of these two states and 'alloSCT for AA: transfusion-independent'. At 5 years after alloSCT, the probabilities of DFS, T-DFS and TT-DFS were 70% (95%-CI 61-81), 60% (95%-CI 51-71) and 42% (95%-CI 33-54), respectively.

severe AA compared to patients with non-severe AA (HR for severe AA 0.4 [95%-CI 0.3-0.7] and for very severe AA 0.5 [95%-CI 0.3-0.8]). None of these risk factors had a significant effect on the likelihood of stopping all non-transplant treatments after having become transfusion-independent.

Patients aged 60 or older were more likely to die (without alloSCT or treatment for AML or MDS) compared to patients younger than 60 (HR 7.3, 95%-CI 1.5-34.3) (Table 2). To show the impact of age, presence of a PNH-clone and disease severity on the different outcomes, we calculated model- based prognoses for reference patients with different baseline characteristics. Figure 4 shows the model-based outcomes after IST for a reference patient with severe aplastic anemia who is <40 years old and has no PNH-clone, a patient of 60 years or older with the same characteristics, and a patient of 60 years or older with non-severe AA and a PNH-clone of ≥1.0% at baseline. Five years after the start of IST, the model-based TT-DFS was 47% (95%-CI 33-68) for the young patient and 24% (95%-CI 14-41) for the older patient with bad characteristics, and 41% (95%-CI 26-65) for the older patient with NSAA and a PNH clone. Model-based outcomes for all 18 possible reference patients (based on age, disease severity and presence of PNH-clone at diagnosis) are shown in Supplemental Figure 4 and Supplemental Table 3.

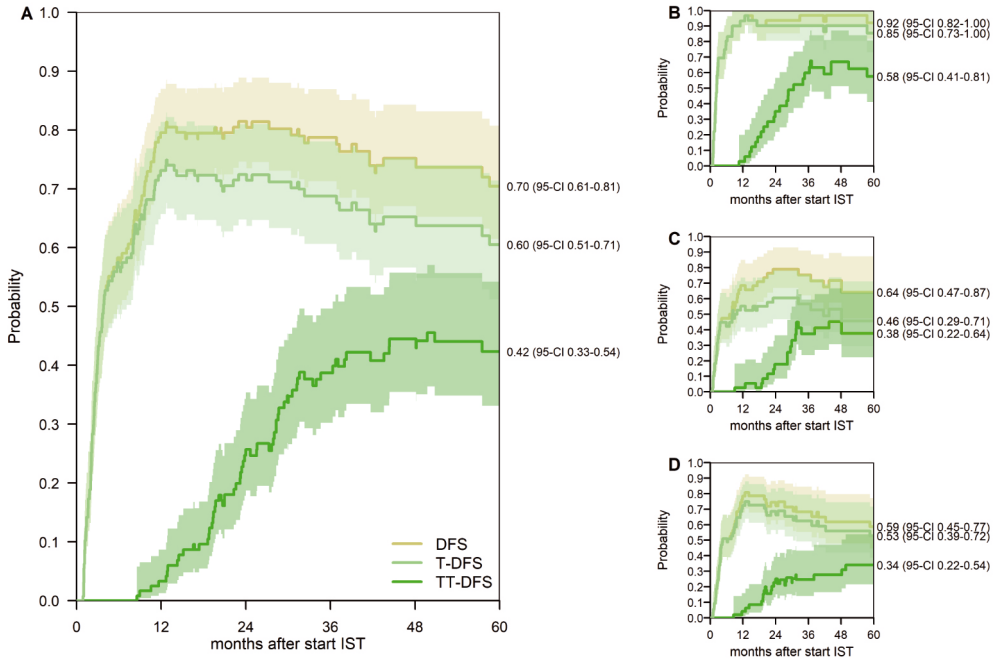


Figure 3. DFS, T-DFS and TT-DFS. Probabilities of Disease-free survival, Transplantation- and Disease-free survival and Transplantation-, Treatment- and Disease-free survival with associated 95% confidence intervals for the total cohort (A) and stratified by age group: 18-39 years (B), 40-59 years (C) and 60 years or older (D). The 5-year probabilities are stated next to the panels.

DISCUSSION

This study is based on real-world data from a unique cohort of unselected adult AA patients who were treated with standard first-line IST in the Netherlands between 2012

Transition	Factor	HR (95%-CI)	P-value
Transfusion-dependent → Free of transfusion with therapy	PNH-clone: ≥1% vs <1%	2.19 (1.42-3.40)	<0.001
	Age: 40-59 vs 18-39 years	0.41 (0.24-0.69)	<0.001
	Age: 60+ vs 18-39 years	0.43 (0.26-0.71)	<0.001
	Severity: SAA vs NSAA	0.42 (0.27-0.67)	<0.001
	Severity: VSAA vs NSAA	0.46 (0.27-0.80)	0.005
Free of transfusion with non-transplant therapy → Free of transfusion and therapy	PNH-clone: ≥1% vs <1%	1.12 (0.65-1.93)	0.69
	Age: 40-59 vs 18-39 years	0.85 (0.43-1.70)	0.65
	Age: 60+ vs 18-39 years	0.84 (0.45-1.58)	0.59
	Severity: SAA vs NSAA	0.83 (0.43-1.59)	0.57
	Severity: VSAA vs NSAA	0.99 (0.48-2.02)	0.98
Transfusion-dependent, Free of transfusion with therapy or Free of transfusion and therapy → Death without alloSCT/AML/MDS	Age: 60+ vs 18-59 years	7.28 (1.55-34.32)	0.01

Table 2. Prognostic factors for becoming transfusion-independent with only non-transplant therapy, becoming free of transfusion and therapy, or dying without alloSCT, AML or MDS. Cox proportional hazards models for the transition from 'transfusion-dependent' to 'transfusion-independent with therapy', for the transition from 'transfusion-independent with therapy' to 'free of transfusion and therapy' and for the combined transitions to 'death without alloSCT/AML/MDS'. Based on complete case analysis (n=112, n=91 and n = 127 for the three models, respectively). Age is determined at time of start IST. HR, hazard ratio; 95%-CI, 95% confidence interval; SAA, severe aplastic anemia; NSAA, non-severe aplastic anemia; VSAA, very severe aplastic anemia

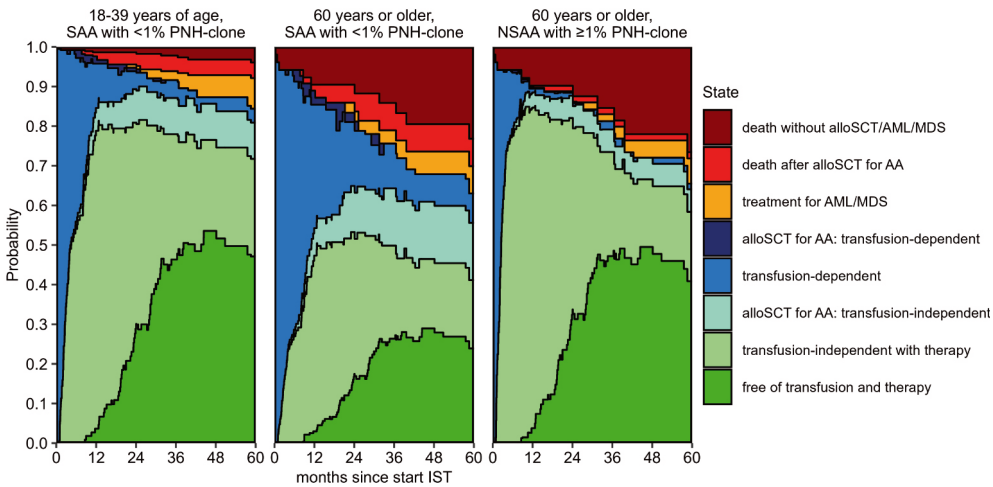


Figure 4. Model-based outcomes for three reference patients with different characteristics. These are based on the multi-state model in Figure 1 and the transition-specific Cox models in Table 2. The difference between two adjacent curves represents the probability of being in the corresponding state. Supplemental Figure 4 and Supplemental Table 3 show the model-based outcomes for all possible combinations of characteristics.

and 2021. These patients had a 5-year overall survival of 79% (95%-CI 70-87), which is comparable to outcomes reported in Phase II studies[19,20]. We showed that after 5 years 60% (95%-CI 51-71) of the patients were alive and transfusion-independent, without MDS or AML and without having received an alloSCT for AA. Only 42% (95%-CI 33-54) of the patients were alive and transfusion-independent without needing medication to treat AA, without MDS or AML and without having received an alloSCT for AA, which we consider as ultimate treatment success. We showed that age, disease severity and the presence of a PNH-clone are associated with treatment success.

Based on data from two prospective clinical trials from the NIH, Prabahran concluded that ATG-based IST is safe and effective in older patients with AA and that this should be the preferred first-line treatment in these patients.[5] We confirm these findings in a real-world cohort showing an OS at 5 years of 65% in patients aged 60 years and older. In the NIH analysis, the overall response (OR) rate at 6 months in these older patients was comparable to the OR in younger patients (72% and 73%, respectively). For older patients, the relapse rate was 71% and the cumulative incidence of high-risk clonal evolution was 17%. This raises the question whether the hematological responses in older patients are durable. If a majority of the older patients with a response to ATGAM-based IST suffer a relapse, develop high-risk clonal evolution or need an alloSCT as second-line treatment, the long-term success of this treatment would not be very satisfying for this patient group. The use of cumulative incidences to quantify response and different causes of failure to IST does not give information about the percentage of patients who have a long-term response after first-line IST without encountering one of these events. We advocate the use of multi-state modeling to deal with recovery after relapse and to show changes in patients' states over time. This makes it possible to analyze dynamic outcomes like DFS including transfusion dependency, DFS without the need for an alloSCT (T-DFS) and T-DFS without need of ongoing medication (TT-DFS). Figure 1 shows that transitions between these dynamic states (transfusion-dependent, transfusion-independent with therapy and transfusion-independent without therapy) occur frequently and Figure 2 shows that only a small minority of patients is transfusion-dependent at 5 years after IST but that 18% still uses medication for the treatment of AA. In those aged 60 years or older, the probability of being transfusion-independent without alloSCT (T-DFS) was 53% at 5 years, while the probability of being transfusion-independent without any treatment for AA (TT-DFS) was only 34%. These measures are more relevant in clinical practice than separate cumulative incidences of relapse and other failures. Although T-DFS and TT-DFS are lower for older patients compared to younger patients, we do not think that older age should be an absolute contra-indication for ATGAM-based IST in patients with aplastic anemia as suggested by others[6]. In our cohort this treatment is effective and safe as indicated by the occurrence of only one death potentially associated with ATGAM treatment. A possible explanation for this difference between our study and the study of Fattizzo is that in that study patients were treated with different types of ATG including rabbit-derived ATG which is associated with deeper T-cell depletion and potentially more toxicity than the currently used horse-derived ATGAM[21,22], and that patients were included who were treated decades ago when the supportive therapy including the possibility for anti-fungal prophylaxis was less developed than nowadays. As therapy with ATG is associated with non-lethal side effects and necessitates hospitalization, it can be argued that old and frail patients are preferentially treated with less intensive regimens

like ciclosporin with or without eltrombopag with no need for hospitalization.[6,23] The long-term treatment success of these therapies should however be compared to the outcomes of ATGAM-based IST.

How the addition of eltrombopag to first-line IST will improve the enduring treatment success is not yet known. All patients in this analysis received first-line treatment with ATGAM and CsA without the addition of eltrombopag as this had not been registered for first-line treatment in the Netherlands during the study period. The prospective randomized RACE study showed that the addition of eltrombopag to this combination leads to an increased complete response rate at 3 months (22% versus 10%). However, this did not translate into a statistical difference in OS at 2 years (90% versus 85%). Furthermore the number of patients that underwent alloSCT as second-line treatment was similar in both treatment arms.[9] Longer follow-up of the RACE study will show whether the addition of eltrombopag to first-line IST for AA will improve the long-term outcome, preferably by showing a better TT-DFS. In our opinion, in addition to OS, analysis of long-term treatment success of IST should include comprehensive endpoints incorporating the development of MDS or AML and the need for chronic treatment for AA.

In order to predict which adult patients with AA potentially benefit most from first-line treatment with ATGAM-based IST, it is important to identify which baseline factors are associated with long-term treatment success. The detailed clinical data from the Dutch AA registry formed the basis for our multi-state model, which was used to calculate model-based outcomes for different reference patients. We showed that at the start of treatment, lower age, less severe AA and the presence of a PNH-clone of ≥1% all increased the probability of reaching the TT-DFS state. For example, patients of 60 years or older with severe AA and no PNH-clone have a poor model-based 5-year TT-DFS (24%, 95%-CI 14-41). On the other hand, patients of 60 years or older with non-severe AA and a PNH-clone of ≥1% have a model-based 5-year TT-DFS of 41% (95%-CI 26-65). For both reference patients the model-based 5-year T-DFS, which can still be considered a satisfying outcome, is 17 percentage points higher. Predictions of treatment success may be relevant in the shared decision making whether to start ATG-based IST.

In summary, we introduced Transplantation-, Treatment- and Disease-free survival, as the ultimate goal for patients with AA who are treated with ATGAM-based IST. Based on patients' age, disease severity and presence of PNH-clone, the probability of reaching this positive outcome can be estimated. If confirmed in another cohort, this could be used as input for personalized treatment decisions.

REFERENCES

1. Young, N. S. Aplastic Anemia. N Engl J Med 379, 1643-1656 (2018). doi: 10.1056/NEJMra1413485
2. Pool, E. S., Luk, S. J., IJsselsteijn, M. E. et al. Imaging Mass Cytometry Reveals the Order of Events in the Pathogenesis of Immune-Mediated Aplastic Anemia. Blood (2025). doi: 10.1182/blood.2025028723
3. Kulasekararaj, A., Cavenagh, J., Dokal, I. et al. Guidelines for the diagnosis and management of adult aplastic anaemia: A British Society for Haematology Guideline. Br J Haematol 204, 784-804 (2024). doi: 10.1111/bjh.19236

4. Tjon, J. M., de Groot, M. R., Sypkens Smit, S. M. A. et al. Short-term efficacy and safety of antithymocyte globulin treatment in elderly patients with acquired aplastic anaemia. Br J Haematol 180, 459-462 (2018). doi: 10.1111/bjh.14372
5. Prabahran, A., Durrani, J., Coelho-Da Silva, J. et al. Safety and efficacy of immunosuppressive therapy for elderly patients with severe aplastic anaemia. Br J Haematol 205, 1170-1179 (2024). doi: 10.1111/bjh.19648
6. Fattizzo, B., Gurnari, C., Giammarco, S. et al. Elderly Patients With Aplastic Anemia: Treatment Patterns and Outcomes in the Real World. Am J Hematol 100, 584-591 (2025). doi: 10.1002/ajh.27611
7. Sugimori, C., Chuhjo, T., Feng, X. et al. Minor population of CD55-CD59- blood cells predicts response to immunosuppressive therapy and prognosis in patients with aplastic anemia. Blood 107, 1308-1314 (2006). doi: 10.1182/blood-2005-06-2485
8. Kulagin, A., Lisukov, I., Ivanova, M. et al. Prognostic value of paroxysmal nocturnal haemoglobinuria clone presence in aplastic anaemia patients treated with combined immunosuppression: results of two-centre prospective study. Br J Haematol 164, 546-554 (2014). doi: 10.1111/bjh.12661
9. Peffault de Latour, R., Kulasekararaj, A., Iacobelli, S. et al. Eltrombopag Added to Immunosuppression in Severe Aplastic Anemia. N Engl J Med 386, 11-23 (2022). doi: 10.1056/NEJMoa2109965
10. Huisman, C., de Groot, M. R., Broers, A. E. C. et al. Richtlijn 'Diagnostiek en behandeling van verworven aplastische anemie bij volwassenen'. Ned Tijdschr Hematol 6, 202-213 (2013).
11. Camitta, B. M. Criteria for severe aplastic anaemia. Lancet 1, 303-304 (1988). doi: 10.1016/s0140-6736(88)90388-1
12. Schemper, M. & Smith, T. L. A note on quantifying follow-up in studies of failure time. Control Clin Trials 17, 343-346 (1996). doi: 10.1016/0197-2456(96)00075-x
13. Putter, H., van der Hage, J., de Bock, G. H., Elgalta, R. & van de Velde, C. J. Estimation and prediction in a multi-state model for breast cancer. Biometrical journal. Biometrische Zeitschrift 48, 366-380 (2006). doi: 10.1002/bimj.200510218
14. de Wreede, L. C., Fiocco, M. & Putter, H. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. Comput Methods Programs Biomed 99, 261-274 (2010). doi: 10.1016/j.cmpb.2010.01.001
15. Therneau, T. M. & Grambsch, P. M. Modeling Survival Data: Extending the Cox Model. (Springer, 2000).
16. Gerds, T. A. prodlim: Product-Limit Estimation for Censored Event History Analysis, <https://cran.r-project.org/package=prodlim> (2019).
17. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. (Springer-Verlag New York, 2016).
18. ggh4x: Hacks for 'ggplot2' (2022).
19. Tichelli, A., de Latour, R. P., Passweg, J. et al. Long-term outcome of a randomized controlled study in patients with newly diagnosed severe aplastic anemia treated with antithymocyte globulin and cyclosporine, with or without granulocyte colony-stimulating factor: a Severe Aplastic Anemia Working Party Trial from the European Group of Blood and Marrow Transplantation. Haematologica 105, 1223-1231 (2020). doi: 10.3324/haematol.2019.222562
20. Patel, B. A., Groarke, E. M., Lotter, J. et al. Long-term outcomes in patients with severe aplastic anemia treated with immunosuppression and eltrombopag: a phase 2 study. Blood 139, 34-43 (2022). doi: 10.1182/blood.2021012130
21. Marsh, J. C., Bacigalupo, A., Schrezenmeier, H. et al. Prospective study of rabbit antithymocyte globulin and cyclosporine for aplastic anemia from the EBMT Severe Aplastic Anaemia Working Party. Blood 119, 5391-5396 (2012). doi: 10.1182/blood-2012-02-407684
22. Feng, X., Scheinberg, P., Biancotto, A. et al. In vivo effects of horse and rabbit antithymocyte globulin in patients with severe aplastic anemia. Haematologica 99, 1433-1440 (2014). doi: 10.3324/haematol.2014.106542
23. Scheinberg, P., Finelli, C., Montaňo-Figueroa, E. H. et al. Activity and safety of eltrombopag in combination with cyclosporin A as first-line treatment of adults with severe aplastic anaemia (SOAR): a phase 2, single-arm study. Lancet Haematol 11, e206-e215 (2024). doi: 10.1016/s2352-3026(23)00395-2

SUPPLEMENTAL METHODS

Guideline and treatment schedule

Dutch Guidelines for the diagnosis and management of adults with acquired aplastic anemia 2013. Published in Dutch in: Nederlands Tijdschrift voor Hematologie 2013 (6),202-213.

Guideline for treatment

	Young patients	Elderly patients
First line	AlloSCT if a HLA identical sibling is available IST if no HLA identical sibling is available	IST
Second line	AlloSCT with HLA identical unrelated donor	AlloSCT with an HLA identical sibling or unrelated donor
	Second-line IST if no HLA identical unrelated donor is available	Second-line IST if no HLA identical unrelated donor is available
Third line	Alternative IST AlloSCT with alternative donor Supportive care only	Alternative IST AlloSCT with alternative donor Supportive care only

Young patients: below 40 years but in case of severe AA without comorbidities, upper age limit of 50 years can be considered.

For second-line IST rabbit-derived Thymoglobulin is suggested.

Treatment schedule with IST

Day	1	2	3	4	5	6
ATGAM 40 mg/kg						
ciclosporin 2.5 mg/kg daily						

Systemic corticosteroids should be given for 1 3 to 4 week period after start of ATGAM to avoid serum sickness.

Aim for ciclosporin levels between 200 and 300 ug/l, in elderly patients or in case of toxicity levels between 150 and 200 ug/l. In case of a hematological response it is advised to taper ciclosporin starting at 6 to 12 months after start with 5-10% per month.

Semi-parametric analyses

To estimate the effects of patient age, presence of a PNH-clone and severity of the AA on the hazards of becoming transfusion-independent with and without non-transplant-treatment, we used two multivariable Cox proportional hazards regression models for the transition from 'transfusion-dependent' to 'transfusion-independent with therapy' and the transition from 'transfusion-independent with therapy' to 'free of transfusion and therapy', and one univariable model for the combined transitions from 'transfusion-dependent', 'transfusion-independent with therapy' or 'free of transfusion and therapy' to 'death without alloSCT/AML/MDS'. The first and second model included patient age (18-39, 40-59 and 60-80 years), the presence of a PNH-clone (<1% and ≥1% PNH-clone), and AA severity (NSAA, SAA, and VSAA) because of their potential relevance shown in previous studies. The third model included only patient age (18-59 and 60-80 years). The transitions for the third model were combined by assuming a shared baseline hazard because of the low number of events for the separate transitions to death without alloSCT/AML/MDS (n = 1 to 7). Patients for whom data about PNH-clone at baseline was not known were excluded from the semi-parametric analyses. Two-sided p-values <0.05 were considered statistically significant.

SUPPLEMENTAL TABLES

Hospital	Year
Leiden University Medical Centre	2012
Amsterdam Medical Centre	2012
University Medical Centre Groningen	2012
Antonius Hospital Nieuwegein	2012*
Medical Centre Twente	2013
Erasmus Medical Centre Rotterdam	2014
Radboud Medical Centre Nijmegen	2014
University Medical Centre Utrecht	2014
VU Medical Centre	2014
OLVG Hospital	2014**
Maastricht University Medical Centre	2017

Supplemental Table 1. Participating hospitals and year in which inclusion of patients in the registry started. *Hospital stopped treating AA patients with ATG in 2015 **Hospital stopped treating AA patients with ATG in 2014

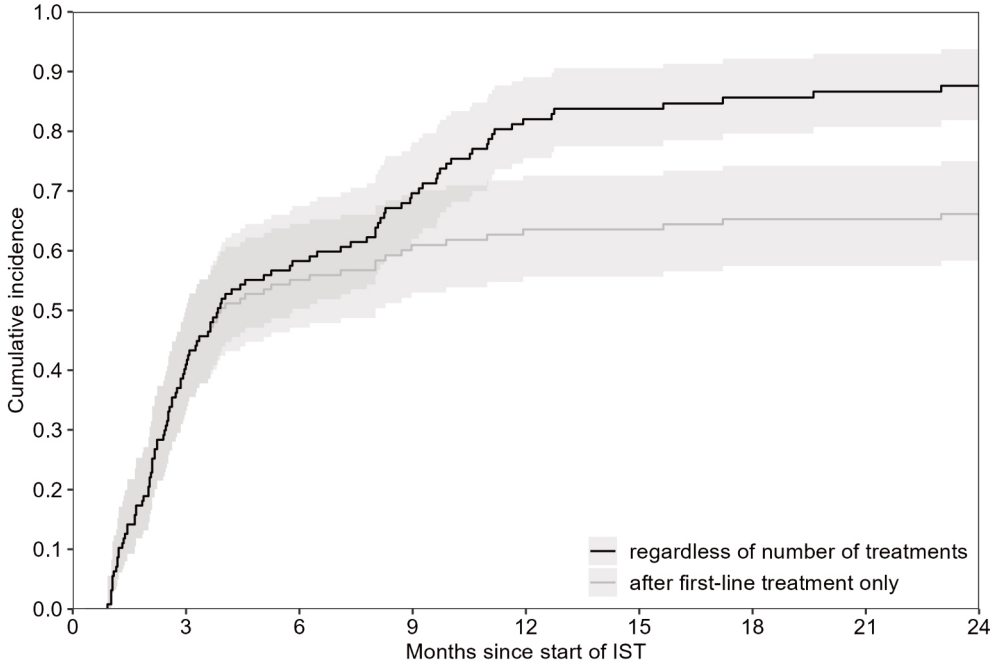
Cause of death (n)	Time after start of IST in months
Complications of alloSCT as second-line treatment (5)	8, 10, 12, 16, 25
MDS/AML (3)	8, 45, 50
Pneumonia (3)	11, 20, 32
Other malignancy (3)	1, 3, 24
Hemorrhage (2)	1, 37
Refractory aplastic anemia (2)	32, 59
Cardiac failure (1)	0.3
Abdominal aneurysm (1)	9
Unknown (1)	40

Supplemental Table 2. Causes of death during 5 years after start of IST.

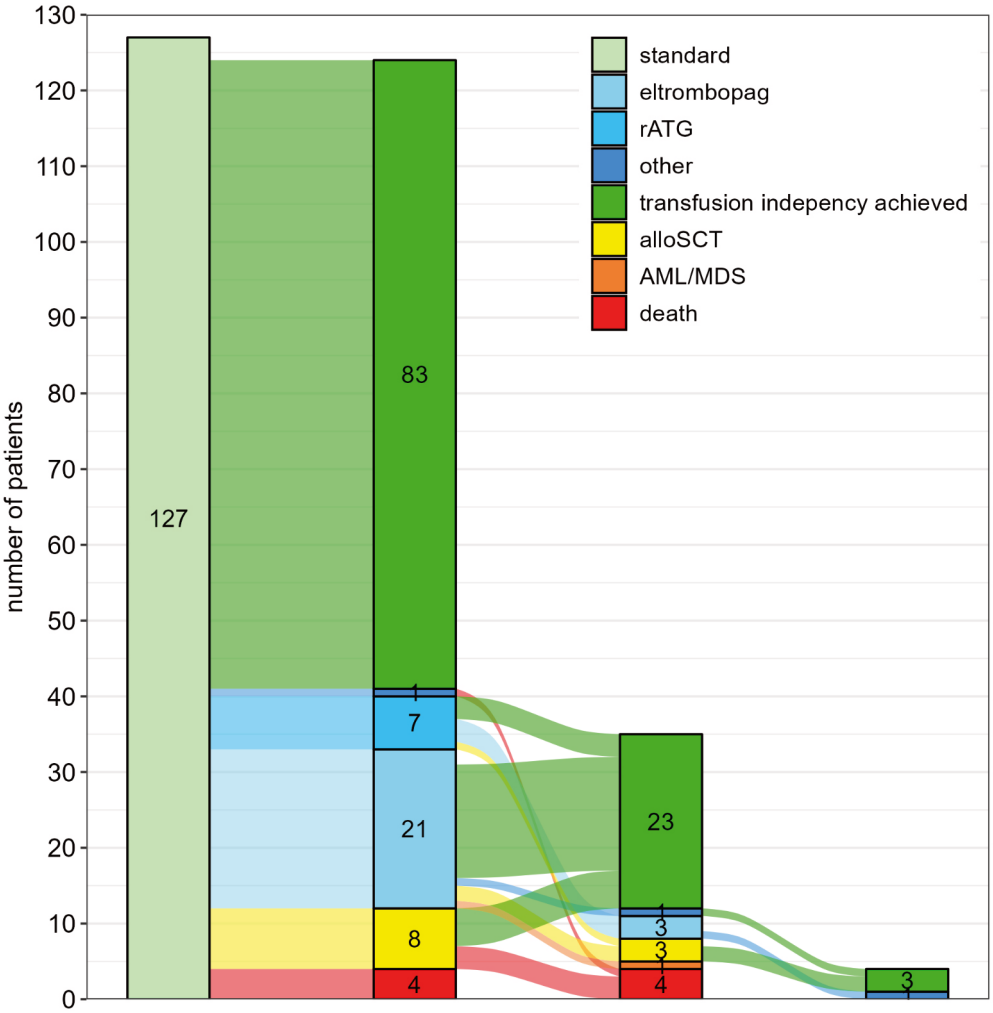
Reference patient*	T-DFS	TT-DFS
18-39, ≥1% PNH-clone, NSAA	0.77 (0.40-1.00)	0.60 (0.41-0.88)
40-59, ≥1% PNH-clone, NSAA	0.78 (0.68-0.90)	0.55 (0.38-0.79)
60-80, ≥1% PNH-clone, NSAA	0.58 (0.44-0.77)	0.41 (0.26-0.65)
18-39, <1% PNH-clone, NSAA	0.79 (0.69-0.90)	0.57 (0.40-0.79)
40-59, <1% PNH-clone, NSAA	0.71 (0.61-0.83)	0.47 (0.31-0.71)
60-80, <1% PNH-clone, NSAA	0.54 (0.40-0.71)	0.35 (0.22-0.56)
18-39, ≥1% PNH-clone, SAA	0.78 (0.69-0.90)	0.55 (0.40-0.75)
40-59, ≥1% PNH clone, SAA	0.70 (0.60-0.83)	0.45 (0.29-0.69)
60-80, ≥1% PNH-clone, SAA	0.53 (0.39-0.71)	0.34 (0.21-0.54)
18-39, <1% PNH-clone, SAA	0.72 (0.61-0.84)	0.47 (0.33-0.68)
40-59, <1% PNH-clone, SAA	0.54 (0.42-0.70)	0.31 (0.18-0.56)
60-80, <1% PNH-clone, SAA	0.41 (0.29-0.59)	0.24 (0.14-0.41)
18-39, ≥1% PNH-clone, VSAA	0.78 (0.68-0.90)	0.59 (0.41-0.84)
40-59, ≥1% PNH-clone, VSAA	0.71 (0.61-0.84)	0.50 (0.33-0.75)
60-80, ≥1% PNH-clone, VSAA	0.54 (0.40-0.72)	0.37 (0.23-0.59)
18-39, <1% PNH-clone, VSAA	0.73 (0.62-0.85)	0.52 (0.35-0.77)
40-59, <1% PNH-clone, VSAA	0.56 (0.43-0.74)	0.36 (0.21-0.62)
60-80, <1% PNH-clone, VSAA	0.43 (0.30-0.61)	0.27 (0.16-0.46)

Supplemental Table 3. Model-based 5-year T-DFS and TT-DFS with associated 95%-CI for all possible reference patients. *Some combinations of characteristics are scarcely present in the data.

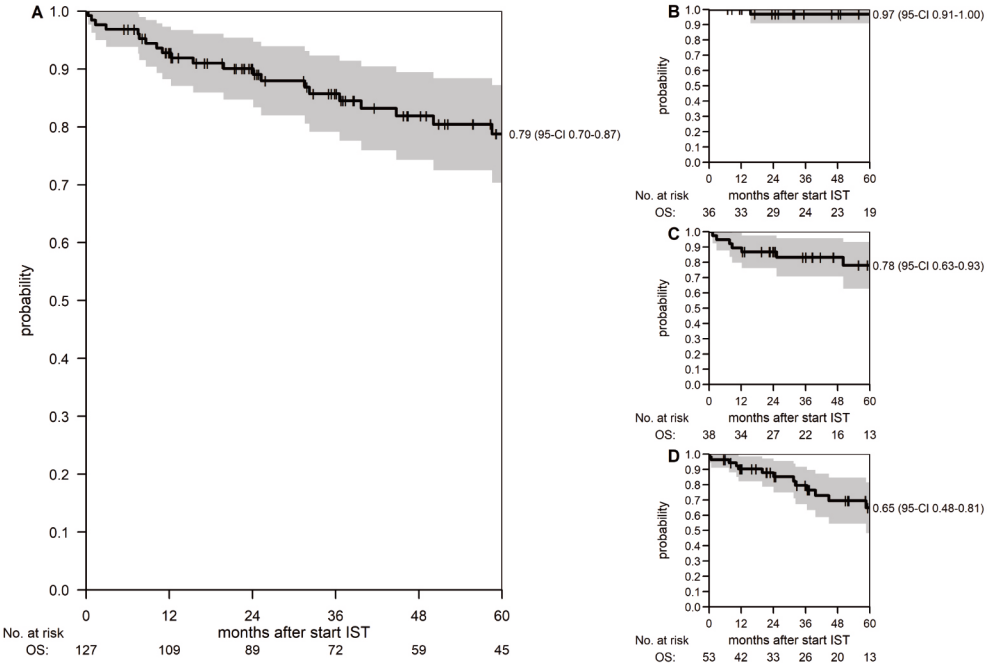
SUPPLEMENTAL FIGURES



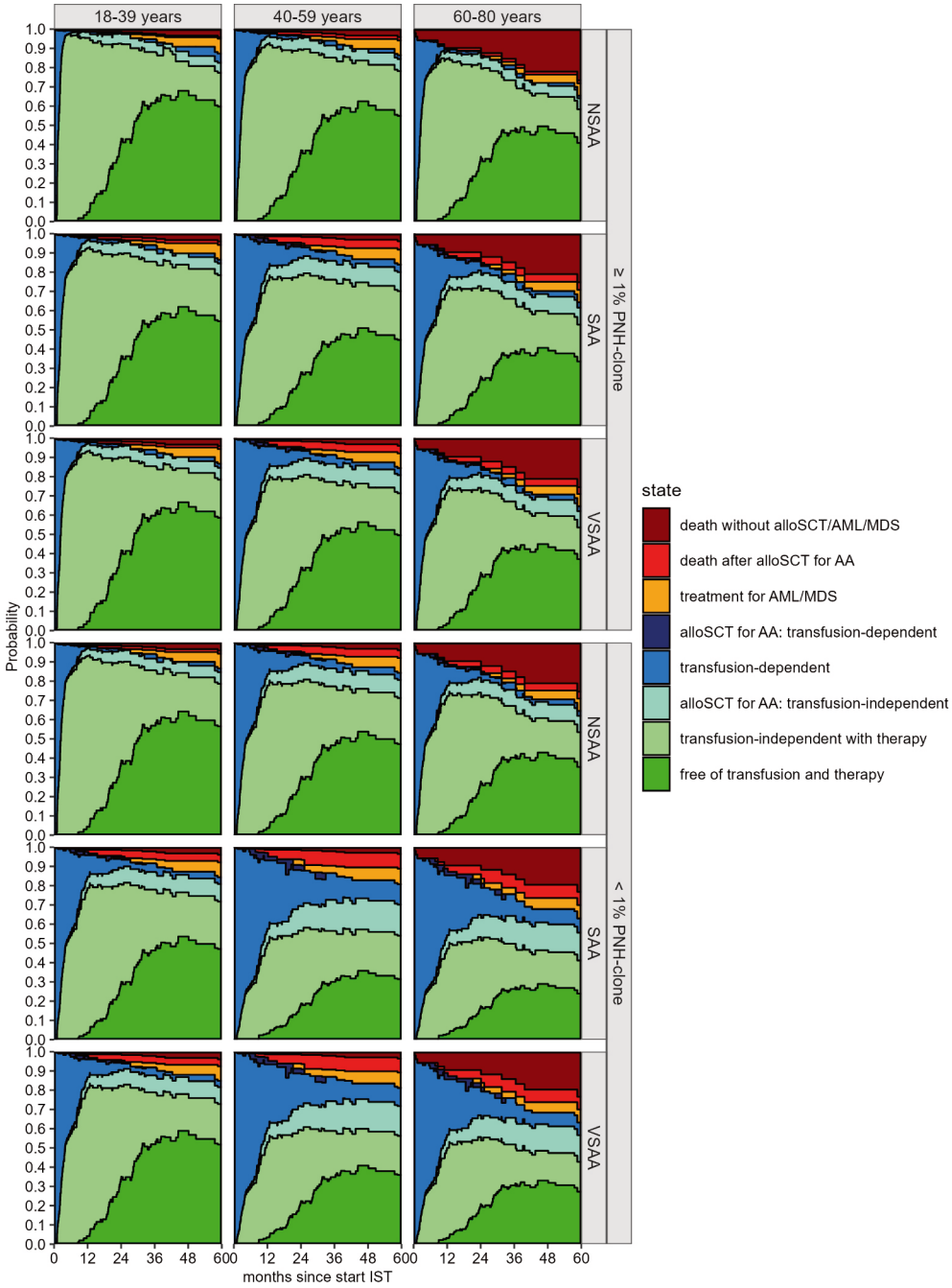
Supplemental Figure 1. Cumulative incidence of achievement of transfusion independency during the first 2 years. The two curves with associated 95% confidence intervals were estimated in separate competing risks models. The first model considered death and treatment for AML or MDS as competing events (black curve), the second model also considered start of any second-line treatment as competing event (grey curve).



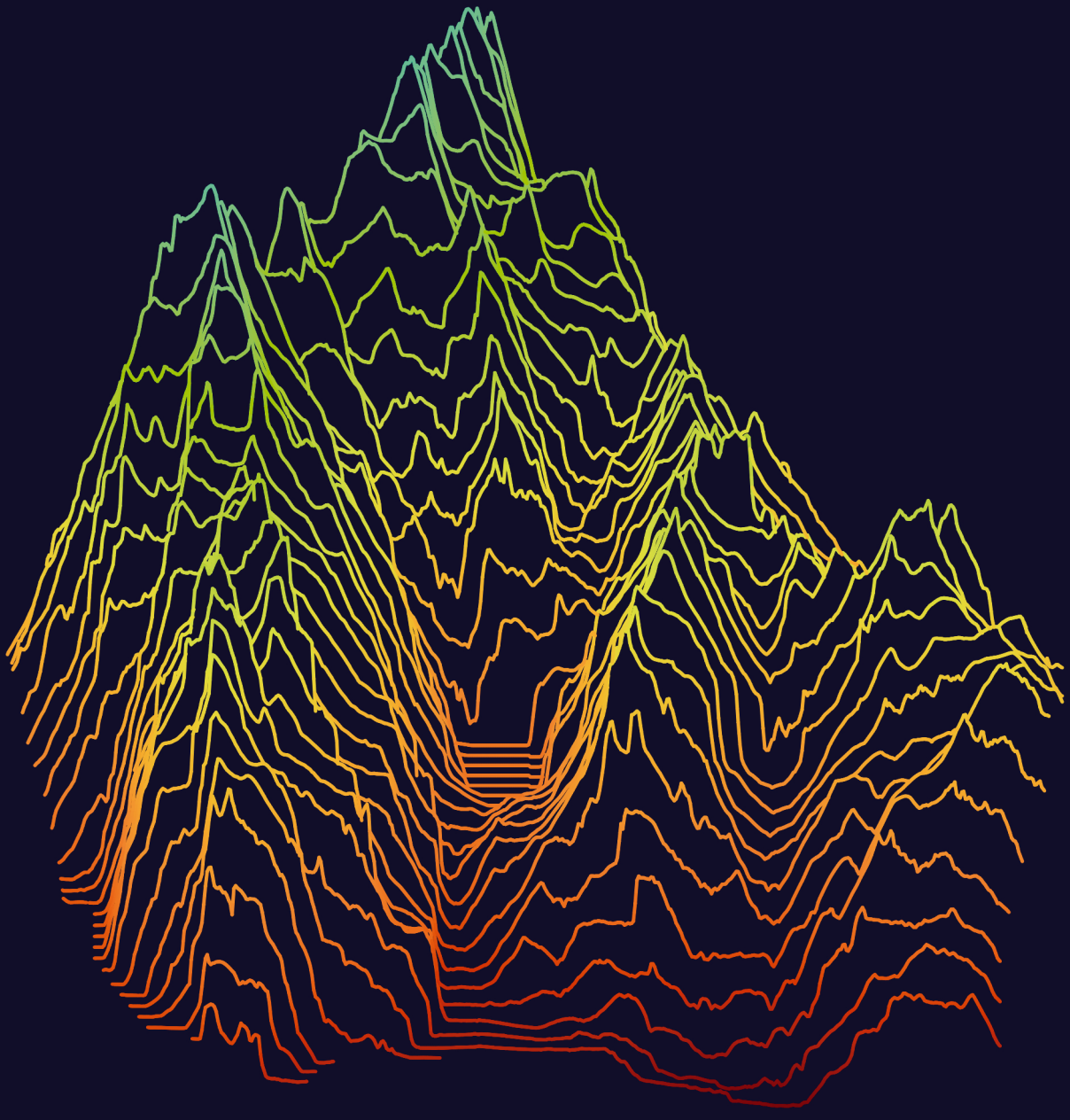
Supplemental Figure 2. Overview of the different lines of treatment until the first achievement of transfusion independency during the first 2 years. All patients started with the standard therapy consisting of ATG and ciclosporin. Each new column represents addition of another line of treatment, achievement of transfusion independency or occurrence of a clinical event. For example, 83 patients became transfusion-independent without any other treatment. Eltrombopag was started in 21 transfusion-dependent patients, after which 15 became transfusion-independent. 8 patients received alloSCT as second-line treatment, of whom 3 died. Death following Acute Myeloid Leukemia or Myelodysplastic Syndrome is ignored. After alloSCT, only death and achievement of transfusion independency are considered.



Supplemental Figure 3. Overall survival. Probabilities with associated 95% confidence intervals for the total cohort (A) and stratified by age group: 18-39 years (B), 40-59 years (C) and 60 years or older (D). | indicates censoring. The 5-year probabilities are stated next to the panels.



Supplemental Figure 4. Model-based outcomes for reference patients with all possible combinations of characteristics. Based on the multi-state model in Figure 1 and the transition-specific Cox models in Table 2. The difference between two adjacent curves represents the probability of being in the corresponding state. See Supplemental Table 3 for the model-based T-DFS and TT-DFS at 5 years after start of IST. Note that some combinations of characteristics are scarcely present in the data.



7

Summary and general discussion

SUMMARY

The aim of this thesis was to investigate how analysis of specific clinical settings and application of advanced statistical methodology on high-quality observational data can be used to investigate complex mechanisms and research questions in the field of hematology.

In **Chapter 2**, we aimed to disentangle the effects of competitive repopulation and allo-immunological pressure on the patient- and donor-derived lymphohematopoietic recovery after allogeneic hematopoietic stem cell transplantation (alloSCT). For this, we selected a cohort of 281 patients with acute leukemia receiving alemtuzumab-based T-cell depleted (TCD) alloSCT after myeloablative (MA) or nonmyeloablative (NMA) conditioning. Part of this cohort received a prophylactic donor lymphocyte infusion (DLI) at 3 months after alloSCT because of an anticipated high relapse risk, while the rest of the cohort could receive DLI from 6 months after alloSCT. This setting provided us a natural control and intervention group for the 3-month DLI. We first investigated the recovery before any DLI. Without DLI, the allo-immunological pressure was low: the 3-month cumulative risk of clinically relevant Graft-versus-Host-Disease (GVHD), i.e., GvHD requiring therapeutic systemic immunosuppression (tIS), was 13% (95% confidence interval [95%-CI] 9-17) in the total cohort and only 2% (95%-CI 0-5) after NMA conditioning. Despite the low allo-immunological pressure, 99% of the patients engrafted, showing that primary engraftment did not depend on MA conditioning or the presence of evident allo-immunological pressure. However, the establishment of complete donor-derived hematopoiesis depended on both the conditioning intensity and the presence of allo-immunological pressure: at 3 months, 32% of the NMA-conditioned patients without any GvHD showed full-donor bone marrow (BM) chimerism compared to 71% of the MA-conditioned patients without any GvHD and 88% of the MA-conditioned patients with GvHD. Granulocytes, monocytes, natural killer (NK) cells and B cells closely followed the BM repopulation status. In contrast, even in patients with complete donor-derived hematopoiesis, circulating CD4+ and CD8+ T cells could be predominantly of patient origin. The 3-month level of donor chimerism in these cell populations depended on the conditioning intensity: of the NMA-conditioned patients 7% and 12% had full-donor chimerism (FDC) in CD4+ and CD8+ T cells, respectively, compared to 33% and 41% of the MA-conditioned patients. We did not observe a significant difference between MA-conditioned patients with and without GvHD, which may be explained by the systemic immunosuppressive treatment that the patients with GvHD still received at the time of this measurement. To assess the impact of the introduction of allo-immunological pressure by DLI, we first compared the 3- and 6-month measurements between patients with a 3-month DLI who did not develop GvHD and patients who did not receive a 3-month DLI and did not develop GvHD. The latter group showed stable BM chimerism kinetics in this period (66% FDC at 3 months compared to 61% at 6 months). In contrast, patients of the DLI group often showed conversion to FDC: 38% had FDC at 3 months compared to 63% at 6 months, demonstrating that the 3-month DLI could induce chimerism conversion even in the absence of GvHD. CD4+ and CD8+ T-cell chimerism kinetics showed similar patterns with increasing levels of donor chimerism in the patients with DLI but not in those without DLI, suggesting that for the establishment of a completely donor-derived T-cell compartment, some allo-immunological pressure is needed. Finally, we investigated the

allo-immunological effects of the total DLI strategy in all patients with mixed hematopoiesis at the time of their first DLI. Of the 65 patients, 72% converted to FDC, of whom only 34% developed clinically relevant GvHD. These results illustrate that the Graft-versus-Leukemia (GvL) effect can be separated from GvHD.

In **Chapter 3**, we investigated the complex associations between immune cell kinetics and alloreactivity by joint modeling. We selected the same clinical setting as in **Chapter 2**, except that we only included 166 NMA-conditioned patients in order to have a single conditioning intensity without any post-alloSCT GvHD prophylaxis that might have influenced the immune cell kinetics. First, we investigated the effect of the 3-month DLI on the kinetics of T-cell and NK-cell counts after TCD alloSCT. For this, we constructed a joint model that considered the first 6 months after alloSCT and compared two groups in an intention-to-treat approach: those scheduled for a 3-month DLI because of an anticipated high risk of relapse (the 'high risk' group) and those who were not (the 'non-high risk' group). The model was run separately for the counts of total (CD3+) T cells, CD4+ T cells, CD8+ T cells and NK cells. The clinical events of interest were start of tIS for GvHD, relapse and other failure (i.e., death, graft failure, start of systemic immunosuppression for a non-GvHD indication and virus-specific T-cell infusion for a severe viral infection). Aside from disease risk group the model also considered donor type (related donor [RD] versus unrelated donor [UD] with anti-thymocyte globulin (ATG) additionally to the alemtuzumab) and patient/donor CMV status (both seronegative or not). Compared to patients with a RD, patients with an UD receiving an ATG-containing conditioning regimen had lower T-cell counts during the first 3 months after alloSCT, illustrating the enduring effect of ATG. However, for those with an UD, starting from 3 months the T-cell trajectories started to diverge between the high and non-high risk groups, resulting in higher T-cell counts in those intended to receive a 3-month DLI. As the only plausible explanation for this increase is the 3-month DLI, these data show that DLI can lead to detectable T-cell expansion. Notably, we did not see a divergence between the risk groups with a RD. We observed significantly more GvHD in the high risk group (hazard ratios [HRs] ranging between 6.3 [CD8 model] and 7.3 [CD4 model]). Also higher CD3 and CD4 counts were associated with a higher risk of GvHD (HR per unit log count increase: 2.4 [95%-CI 1.4-4.1] and HR 1.5 [95%-CI 1.0-2.3], respectively). Higher CD4 counts decreased the risk of relapse (HR 0.6, 95%-CI 0.5-0.9) and other failure (HR 0.7, 95%-CI 0.6-1.0). NK cell counts were associated with a higher risk of GvHD and a lower risk of relapse. However, when including both CD4 and NK cell counts in an exploratory time-dependent cause-specific Cox model for GvHD, the effect of NK cell counts disappeared, suggesting that the observed association between NK cell counts and GvHD merely reflected the high correlation between the counts of NK cells and CD4+ T cells and not a direct effect of NK cell counts on the risk of GvHD. To further investigate the T-cell kinetics after the 3-month DLI, we constructed a second joint model starting from this DLI, only including those who actually received this DLI. Having an UD (HRs ranging between 7.0 [CD8 model] and 22.5 [CD4 model]) and higher CD3, CD4 and CD8 counts (HRs ranging between 1.6 [CD8 model] and 6.7 [CD4 model]) were all associated with a higher risk of GvHD during the first 3 months after DLI.

In **Chapter 4**, we aimed to identify other risk factors that influence the alloreactivity of DLI, considering a cohort of patients with acute leukemia receiving their first DLI at 3

(n = 88) or 6 (n = 76) months after alemtuzumab-based TCD alloSCT. First, we assessed the relationship between the timing and dose of DLI and the risk of clinically relevant GvHD in relation to donor and conditioning type. The tenfold dose difference between the 3- and 6-month DLI resulted in similar risks of GvHD: 28% (95%-CI 20-40) and 30% (95%-CI 22-43) at 3 months after the 3- and 6-month DLI, respectively. For both DLIs, the 50% dose reduction in case of an UD sufficed for equalizing the GvHD risks between patients with RD and UD after MA conditioning. In contrast, NMA-conditioned patients with an UD still had a higher risk of GvHD than NMA-conditioned patients with a RD. Then, we focused on three conditions at the time of DLI that could promote T-cell activation: the presence of patient-derived antigen-presenting cells (APCs) as estimated by the BM chimerism status, lymphopenia and the presence of a viral infection close to DLI. As we wanted to estimate the effects of these risk factors on the development of GvHD, the risk of death during GvHD and on the total clinical outcome after DLI, we constructed a time-inhomogeneous Markov multi-state model starting at the time of first DLI and considering the following events: start of tIS for GvHD, stop of tIS, relapse, death and second DLI. The model was run three times for the 3- and 6-month DLI separately, each time including only one of the factors of interest and donor/conditioning type. For the 3-month DLI, viral infections close to DLI increased the risk of GvHD (HR 3.7, 95%-CI 1.7-7.9), while we observed no significant associations with BM chimerism or lymphopenia. At the time of the 6-month DLI, viral infections were uncommon and played no important role in the development of GvHD. Instead, the presence of ≥5% mixed chimerism (MC) in the BM significantly increased the risk of GvHD (HR 3.6, 95%-CI 1.2-11.3) while the presence of 1-5% MC in the BM and lymphopenia showed a trend of increasing the risk of GvHD. We did not observe significant associations between the risk of death during tIS and the main risk factors of GvHD, i.e., viral infections for the 3-month DLI and BM chimerism for the 6-month DLI. To demonstrate the impact of viral infections on current GvHD-relapse-free survival (cGRFS) after the 3-month DLI, we extended the multi-state model and compared the 6-month cGRFS from different starting states: 61% (95%-CI 50-73) from the state 'DLI without viral infection' versus 31% (95%-CI 19-52) from the state 'DLI with viral infection'. For the 6-month DLI, we integrated the two transition-specific Cox models with BM chimerism as components in the multi-state model to predict the outcome for two reference patients, a MA-conditioned patient with a RD and FDC at time of DLI and a MA-conditioned patient with a RD and ≥5% MC. The 6-month cGRFS for these reference patients was 77% (95%-CI 60-98) and 44% (95%-CI 19-100), respectively. The strong impact of viral infections and BM chimerism on cGRFS underline the clinical relevance of these findings.

In **Chapter 5**, we investigated how the transplantation strategy affects the alloreactivity of DLI by considering a different clinical setting, DLI following alloSCT with posttransplant cyclophosphamide (PTCY). Like TCD, PTCY can be applied as partial *in vivo* T-cell depletion early post-transplant to reduce the risk of severe GvHD after alloSCT, but it leads to faster immunological recovery and more FDC early after alloSCT compared to TCD. In this setting, the low-dose DLI was given at 4 months after alloSCT instead of the 3 months in the case of alemtuzumab-based TCD alloSCT. First, we examined the risk factors we had identified in **Chapter 4**. All risk factors were uncommon: of the 83 patients receiving a 4- or 6-month DLI, only 5% had a viral infection close to DLI, 6% had ≥5% mixed BM chimerism and 17% had lymphopenia,

far less that what we had observed in the alemtuzumab setting (19%, 27% and 47%, respectively). We then investigated the development of clinically relevant GvHD after DLI. In line with the low presence of these risk factors, the risk of GvHD was very low: 4% (95%-CI 0-8) at 3 months after DLI. Only one patient died of GvHD, after receiving a 6-month DLI while having 14% patient material in the BM chimerism sample. The combined results of **Chapter 4** and **Chapter 5** indicate that transplantation strategies have a profound impact on the conditions at time of DLI, which in turn influence DLI alloreactivity. To investigate whether DLI after PTCY alloSCT could still induce chimerism conversion from MC to FDC, we examined the BM chimerism kinetics of the 28 patients with MC at time of their first DLI: 79% converted to FDC, of whom only 9% developed clinically relevant GvHD. This conversion rate was similar to what we had observed in **Chapter 2**, while the risk of GvHD was lower. None of the responders relapsed, indicating achievement of a meaningful GvL effect.

In **Chapter 6**, we aimed to investigate whether and how the multi-state framework could be used to develop a comprehensive measure of "treatment success" that can capture the complex clinical recovery and failure patterns of patients with aplastic anemia (AA) receiving immunosuppressive therapy (IST). We defined three levels of treatment success. The broad aim of IST for AA is to achieve and maintain transfusion independency without the development of secondary BM diseases like acute myeloid leukemia (AML) or myelodysplastic syndrome (MDS). We captured this by the endpoint Disease-free survival (DFS). Of note, this endpoint does not consider blood cell counts: it is possible for a patient to be in a DFS state while still having granulopenia. However, transfusion independency is likely more indicative for quality of life than exact blood cell counts. DFS should preferably be achieved without requiring an alloSCT: Transplantation- and Disease-free survival, T-DFS. The ultimate aim is to stop all AA therapy after a response is achieved, which is captured by the endpoint Transplantation-, Treatment- and Disease-free survival, TT-DFS. We estimated these endpoints in a cohort of 127 transfusion-dependent patients with AA using a time-inhomogeneous Markov multi-state model to allow for gain, loss and recovery of response. The 5-year probabilities of DFS, T-DFS and TT-DFS were 70% (95%-CI 61-81), 60% (95%-CI 51-71) and 42% (95%-CI 33-54), respectively. In the next step, we investigated the effects of age, AA severity and the presence of a GPI-deficient cell clone on different transitions. As GPI-deficient cells are considered to be less sensitive to an immune attack and can emerge as a detectable cell population after the other HSCs have been depleted by autoimmune cells, the presence of such a clone can indicate that the AA was caused by autoimmunity. Since only those with AA caused by an autoimmune response are likely to respond to IST, the presence of a GPI-deficient cell clone may be associated with achieving a response. Indeed, having a GPI-deficient cell clone of ≥1% increased the hazard of becoming transfusion-independent (HR 2.2, 95%-CI 1.4-3.4), while age of 40 or above and having severe or very severe AA decreased the hazard with HRs ranging between 0.4 and 0.5. We did not observe significant effects of these risk factors on the likelihood of being able to stop all non-transplant therapy after having become transfusion-independent. As expected, age of 60 or above was a strong predictor for the risk of death (HR 7.3, 95%-CI 1.5-34.3). To demonstrate the impact of these risk factors on the outcomes, we calculated model-based prognoses for reference patients with different baseline characteristics. For example, the model-based 5-year probability of TT-DFS for a patient of 40 years or younger with severe aplastic anemia and no GPI-

deficient cell clone was 47% (95%-CI 33-68), compared to 24% (95%-CI 14-41) for a patient of at least 60 years old with the same characteristics, and 41% (95%-CI 26-65) for a patient of at least 60 years old with a GPI-deficient cell clone and non-severe AA. These results indicate that the three risk factors have a strong impact on the probability of long-term treatment success.

GENERAL DISCUSSION

In this thesis, we demonstrated how detailed observational data and advanced statistical methods can be used to answer clinical and immunological research questions in the field of hematology by capturing complex recovery and failure patterns and their underlying mechanisms. The use of cohorts of patients with or without DLI after different transplantation strategies allowed us to investigate the impact of the transplantation strategy and DLI on the lymphohematopoietic and clinical recovery and to disentangle the roles of competitive repopulation and allo-immunological pressure. Via joint modelling we could quantify the DLI effect on immune cell counts and the associations between the immune cell counts and different clinical events. With the Markov multi-state framework we could investigate the effects of risk factors on different components of the recovery and failure process and translate these to clinically relevant outcome measures. Taking advantage of the versatility of the multi-state framework, we constructed a single comprehensive multi-state model that could estimate the probabilities of three different levels of treatment success after IST for AA over time.

Vital importance of setting of DLI

In **Chapters 4 and 5**, we aimed to investigate the effects of factors that might increase the risk of DLI-induced GvHD: the presence of patient-derived APCs, a proinflammatory environment and cytokines that can amplify GvHD. However, none of these factors were measured standardly. Instead, for each of the actual factors of interest we considered a measure that was available as a proxy: the presence of MC in the BM, viral infection (one of the most common causes of inflammation) and lymphopenia (during which the concentrations of cytokines that promote T-cell proliferation are higher), respectively. Because these had been measured as part of the standard clinical care, we could use the full cohorts for our analyses. We demonstrated that the presence of mixed chimerism and viral infections can have a strong impact on the risk of DLI-induced GvHD, and that their presence depends on the transplantation strategy and timing of DLI.

Current DLI recommendations do not take conditions at the time of DLI into account aside from strongly advising against the use of DLI to boost the GvL effect in the presence of GvHD and uncontrolled infections.[1,2] This can partly explain the wide range of reported risks of GvHD[3]. More personalized DLI protocols that consider more conditions at the time of DLI would likely improve the balance between efficacy (GvL effect) and toxicity (GvHD) of prophylactic and preemptive DLI. Firstly, the recommendations for DLI in case of infection could be extended. Postponing a DLI until an infection has been cleared is not always feasible, for instance when MRD is increasing rapidly. As an alternative, the DLI dose could be reduced. If an infection occurs just after

DLI, one could more aggressively start tIS upon signs of GvHD. However, the benefit of reducing the risk of severe GvHD should be weighed against the disadvantage of suppressing the immune system while fighting an infection. Secondly, the presence of lymphopenia or MC in the BM could be considered for the dose determination in the preemptive and prophylactic setting. Reducing the dose in case of MC may feel counterintuitive, since patients with MC are more likely to develop a relapse. However, as long as there are no signs of relapsing disease, the presence of MC merely reflects the absence of a GvL effect; the magnitude of MC should not have an effect on the relapse risk. In contrast, a patient with 5% MC likely has more patient-derived APCs and thereby a larger GvHD risk than a patient with 1% MC. Following this reasoning, one would recommend a lower dose for the former patient to minimize the risk of severe GvHD. If no response occurs, further DLIs can be given following a dose-escalating approach. Conversely, for patients with FDC and therefore likely having fewer patient-derived professional APCs, one could increase the DLI dose if they are considered to have a high risk of relapse. None of the FDC patients in our study died of GvHD, suggesting that there is room for increasing the dose. Thirdly, for patients who need a stronger alloimmune response one could *initiate* a proinflammatory condition by administering lymphodepleting chemotherapy before DLI. Miller et al. showed that this is an effective method to increase the alloreactivity of DLI, but had to stop the trial because it caused too much GvHD-related toxicity.[4] Guillaume et al. applied lower doses of lymphodepleting chemotherapy and observed no toxicity in patients receiving DLI after 3 days of 25mg/m^2 fludarabine.[5] While these studies took place in the setting of therapeutic DLI with small patient cohorts, they suggest that low-dose lymphodepletion can be an effective tool to increase DLI efficacy.

Before any of these suggestions can be implemented, they need to be tested and validated in other cohorts and settings. Two types of studies are needed: observational studies to validate the estimated effects of risk factors for GvHD after DLI and intervention studies to investigate dose adjustments based on the conditions at the time of DLI and the use of lymphodepletion in situations where a prompt alloimmune response is needed.

Added value of complex statistical analyses

Two advanced statistical methods were applied in this thesis: Markov multi-state modeling for investigating complex sequences of events and joint modelling for analyzing the trajectories of biomarkers and their effects on survival outcomes. The multi-state models constructed in this thesis showcase several advantages of multi-state modelling. Firstly, because multi-state models can consider sequences of events and analyze events in continuous time without assuming a constant hazard, they can capture dynamic measures of treatment success such as cGRFS and TT-DFS. cGRFS differs from GvHD-relapse-free survival (GRFS) by considering recovery after GvHD: patients in whom GvHD does not resolve or who die of GvHD remain in a failure state, while those who recover move on to a non-failure state. This better reflects the clinical situation during follow-up, since patients with resolved GvHD can have comparable quality of life compared to those who never developed GvHD. [6] Moreover, patients with resolved GvHD may benefit from the concomitantly established GvL effect reducing their risk of relapse, as shown in **Chapter 4** (none of the patients who started tIS for GvHD relapsed) and by others[7]. In the setting of IST for AA, we defined three dynamic

endpoints, the DFS, T-DFS and TT-DFS, to evaluate different levels of treatment success (**Chapter 6**). In contrast to the commonly used 6-month response rate and overall survival, these measures shows the loss and gain of different levels of response over time and give information on the different treatments of AA during follow-up and the risks of death and development of secondary BM diseases. Secondly, the approach of including transition-specific risk factors focuses on the underlying biological processes. This is more logical than estimating the effect of risk factors directly on composite endpoints such as GRFS. For instance, Tan et al. [8] constructed an extensive multivariable Cox model for GRFS. While such an approach may be sufficient if one is only interested in the prediction of GRFS, it does not help to understand why certain risk factors are important: donor type was not significant, but does this mean that it was irrelevant or that the opposing effects of higher genetic disparity on relapse and GvHD cancelled each other out? Several studies have shown the added value of multi-state modelling by reanalyzing trials.[9-11] For instance, Bakunina et al. reanalyzed a trial[12] where patients with AML were randomized to receive remission induction therapy with or without clofarabine, showing that clofarabine reduced the risk of relapse but did not improve survival. Their multi-state approach enabled them to consider intermediate events such as consolidation by alloSCT and achievement of MRD negativity, and showed that the addition of clofarabine reduced the risk of relapse irrespective of MRD status or alloSCT, but increased the risk of non-relapse mortality before alloSCT.[10] In **Chapter 6**, the transition-specific age effects in the multi-state model on patients with AA receiving IST showed that the relatively poor TT-DFS of patients aged 60 years or older can be explained by both a lower hazard of achieving a response and a higher hazard of death, of which a descriptive analysis indicated that the majority was not related to treatment toxicity but to pancytopenia. The possibility to use the estimated transition-specific effects to calculate model-based prognoses and thereby show the total impact on the clinical outcome is the third advantage of multi-state modelling. Of note, predictions can be given for each state separately or for combinations of states, allowing to show the impact on several clinical outcomes of interest (**Chapters 4 and 6**). For treatment decisions and prognosis, outcome predictions are often more relevant than HRs, as they can take the full recovery process and opposing effects of the same risk factor on different transitions into account as well as the baseline hazards.

There are also some limitations to the multi-state models used in this thesis. They depend on the Markovian assumption, meaning that the risk to make a certain transition only depends on the state, the time since start of the analysis and, if estimated, the transition-specific covariate effects. This is a simplification, as for instance the risk of death after relapse likely also depends on the timing of the relapse (early relapses usually have a worse prognosis than later relapses) and the time since relapse (e.g., those who are still alive a year after relapse likely have a lower mortality risk than those who just developed a relapse). Multiple timescales have only been implemented in parametric models[13], which require more assumptions than non- and semi-parametric models. As an alternative, states can be split into multiple states to include some of the information of the time since the start of a clinical event. For instance, continuing with the relapse example, relapse could be split into early and late relapse and additional states could be added such as '1 year after relapse'. However, this approach can quickly lead to very extensive multi-state models requiring more data in order to have sufficient events for the transitions. Moreover, interpreting covariate effects would be more complex, as separate

coefficients are calculated for each transition. Because of these reasons, we did not apply this approach in the studies in this thesis, even though for example in **Chapter 6**, the likelihood of stopping all non-transplant therapy after having become transfusion-independent likely depends on both the time since start (as according to the Dutch AA guidelines IST should be given for at least 6 months) and on the time since achieving transfusion independency.

The other advanced model applied in this thesis is the joint model. In **Chapter 3**, the raw data (see Supplemental Figure 3) did not hint at a difference in the T-cell count trajectories between those who eventually did or did not develop GvHD, but the association could be made visible and quantified using joint modelling. This method is far more efficient than landmarking such as performed by Podgorny et al.[14], as landmarking required them to exclude patients who had developed GvHD before the landmark time and to consider the immune cell counts at the landmark time as fixed baseline covariates for their Cox proportional hazards model. As outlined in **Chapter 1**, joint models can model the trajectories of biomarkers over time without assuming constant values between measurements or absence of measurement error. Moreover, joint models can yield individualized predictions to visualize and quantify how changes in the biomarker values affect the risk of clinical events, as illustrated by Baart et al.[15] As an example they predicted the risk of neo-aortic valve regurgitation for two patients who underwent surgery shortly after birth because of transposition of the great arteries, updating their risk each time the neo-aortic root diameter was measured (i.e., dynamic prediction). While at the start both patients had the same risk of this event, their risks diverged considerably over time as in one patient the root diameter increased more slowly than in the other patient. Dynamic prediction tools like this can have a great value in the clinical follow-up of patients.

Barriers for widespread application of complex statistical methods

During the studies explored in this thesis, we encountered several limitations due to the relatively low numbers of patients and events compared to the complexity of our models. The joint model in **Chapter 3** was based on 166 patients and considered four immune cell populations and three endpoints of interest, GvHD, relapse and a composite of all other failures. As GvHD and relapse depend on the presence and absence of allo-immunological pressure, respectively, we expected opposite effects of the immune cell counts on these events and included them as separate endpoints. The third endpoint was needed to stop the follow-up as soon as an event occurred that could influence the immune cell counts or the risks of relapse and GvHD (aside from the 3-month DLI, which was the intervention of interest). As incorporation of all four immune cell populations in a single model would require far more assumptions regarding the association structure, we had to investigate the immune cell populations of interest in separate models, which was a lot considering the size of our dataset. In **Chapter 4**, we had 88 patients with a 3-month DLI and 76 with a 6-month DLI, which were analyzed in two separate multi-state models with 14 states. In both models, about 30 patients made the transition from DLI1 to start of tIS for GvHD. The three risk factors of interest had to be analyzed in separate Cox models as we needed to include conditioning/donor combination, which was associated with both the presence of the risk factors and the risk of GvHD after DLI. Having more events would have allowed us to include all our risk

factors of interest in one comprehensive model. The numbers of patients who died during tIS were even lower (16 and 12 for the 3- and 6-month DLI, respectively), which likely explained why we did not detect significant effects of high MC or viral infections on the risk of death during tIS for GvHD. Moreover, none of the patients with FDC died during tIS for GvHD, necessitating us to use the presence of either FDC or low MC as the reference category for high MC. Therefore, we could not quantify the increased hazard of death during tIS for patients with high MC compared to patients with FDC. In **Chapter 6** we had a larger cohort, but the low numbers of events for the transitions to death required us to assume a shared baseline hazard in order to assess the effect of age on the risk of death. Thus, while our analyses yielded valuable results, they were less precise and required more assumptions than if we would have had more events. In order to construct even more complex models or get more precise estimates, larger datasets are needed. However, multi-state and joint models require high-quality detailed observational data, which is often only collected in relatively small cohorts. Registries have large cohorts, but will most likely need to improve their data collection in order to have data of sufficient quality and detail for these types of models. This requires more commitment of the registries and their participating centers.

Another barrier for the use of complex statistical methods such as multi-state and joint models is the required level of statistical knowledge to construct these models and to correctly interpret the results. **Chapter 3** was a joint project of the department of Hematology and the department of Biomedical Data Sciences to ensure sufficient knowledge both on the clinical and immunological processes and on the modelling techniques. Assuming a shared baseline hazard for different transitions as we did in **Chapter 6**, could only be done after carefully considering the medical/biological implications of this statistical assumption. Moreover, having a correct model does not guarantee correct interpretation. An audience with less experience in multi-state methodology may find it difficult to for instance understand why risk factors can have opposing effects on different transitions in a multi-state model or to understand the implications of all model assumptions.

Outlook

Our studies show that the combination of detailed data and advanced statistical methods can be used to answer complex research questions using real-world clinical data. Standard collection of detailed observational data has several advantages over data collected for a specific study. Firstly, standardly collected data that are stored in a single place can be easily used for multiple studies. It also provides a reliable data source for all consecutive patients that can be accessed rapidly in the case of fast-changing developments. For instance, the unexpected change in DLI-induced GvHD risk after switching from alemtuzumab-based TCD to PTCY alloSCT could be investigated early after implementation of PTCY, because in both settings detailed information was collected in a standardized way. Ethically, data may only be collected for research if they are needed to answer relevant research questions. To this end, it is essential that both before the start of and during the data collection, researchers, data managers and methodologists discuss which research questions may be of interest, what measurements could be useful and analyzable and how they can be most efficiently collected. Also for

data being collected for clinical care, it is valuable to discuss with a statistician whether the data are usable for analysis.

If more data with sufficient detail and quality become available, it will be possible to extend the models used in this thesis or apply them in other settings. For the AA model, more covariates could be included, and we could also include covariates on other transitions, such as relapse after having achieved transfusion independency. The observed differences between the alemtuzumab-based TCD and PTCY setting suggest that it would be very valuable to repeat the analyses in different alloSCT settings to validate the results. For all multi-state models, it would also be possible to incorporate biomarkers, for instance T-cell counts in the cGRFS model and blood counts in the AA model. This approach has already been implemented by Ferrer et al.[16], but has not yet been applied in the field of hematology. As mentioned before, large datasets are mainly available from registries. It is unlikely that very large registries such as the European Society for Blood and Marrow Transplantation will collect very detailed clinical data or biomarker data. Collaboration projects between centers and regional or national registries are more likely to obtain sufficient data of the required quality, since a lower number of participating centers usually means higher commitment for precise and detailed data collection.

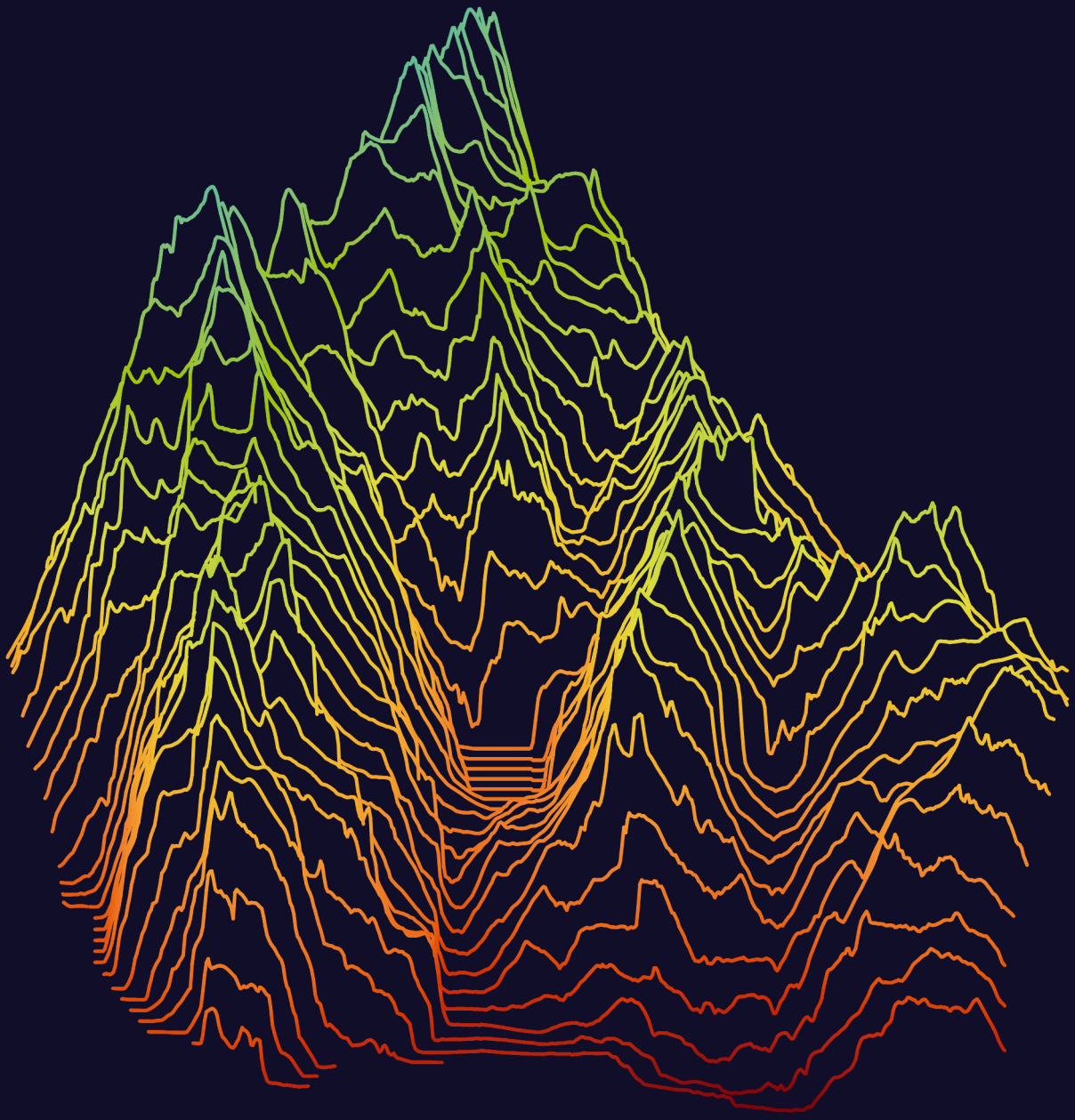
After the models have been validated, they may be of value in clinical practice. Model predictions can help with treatment decisions, shared decision making and counseling of the patients. A major advantage of both the multi-state model and the joint model is that they can incorporate new information such as biomarker trajectories and the occurrences of clinical events during follow-up. This allows to update the prognosis of patients based on their trajectory so far, which may help in the decision for further treatment lines. The models in this thesis were constructed to better understand underlying mechanisms behind recovery and failure patterns and not to serve as a prediction tool. For that, the results of the semi-parametric analyses need to be validated in large cohorts, that are preferably treated according to the current treatment guidelines. For instance, the recent addition of eltrombopag to the first-line IST treatment protocol for patients with AA may affect both the prognosis and the effect sizes of the risk factors: the estimated values in **Chapter 6** might not hold for patients that are treated according to the new guideline.

Finally, to narrow the gap between clinical researchers and methodologists, papers applying advanced statistical methods should become more accessible to a less methodological audience. To serve both audiences, it is important to explain the methods both on a more general level and a more technical level. In the studies of this thesis that report multi-state models, we discussed the model results in steps: first the non-parametric analyses to show how the transition probabilities changed over time, then the Cox models per transition and finally predictions for reference patients. By analyzing clinically relevant questions with real-world data, publishing in clinical or immunological journals and explaining the methods comprehensively, not only methodologists but also clinical researchers can be reached. Hopefully, by showing by example the added value of more advanced statistical models in answering important questions, the collaboration and cross-talk between the different fields will increase, leading to more opportunities for methodologists to develop or apply methods on clinically relevant applications and for clinical researchers to answer complex research questions.

REFERENCES

1. Falkenburg, J. H. F., Schmid, C., Kolb, H. J. & Kuball, J. in The EBMT Handbook: Hematopoietic Cell Transplantation and Cellular Therapies (eds Anna Sureda et al.) 531-539 (Springer International Publishing, 2024).
2. Pagliuca, S., Schmid, C., Santoro, N. et al. Donor lymphocyte infusion after allogeneic haematopoietic cell transplantation for haematological malignancies: basic considerations and best practice recommendations from the EBMT. The Lancet Haematology 11, e448-e458 (2024). doi: 10.1016/S2352-3026(24)00098-X
3. Biederstädt, A. & Rezvani, K. How I treat high-risk acute myeloid leukemia using preemptive adoptive cellular immunotherapy. Blood 141, 22-38 (2023). doi: 10.1182/blood.2021012411
4. Miller, J. S., Weisdorf, D. J., Burns, L. J. et al. Lymphodepletion followed by donor lymphocyte infusion (DLI) causes significantly more acute graft-versus-host disease than DLI alone. Blood 110, 2761-2763 (2007). doi: 10.1182/blood-2007-05-090340
5. Guillaume, T., Gaugler, B., Chevallier, P. et al. Escalated lymphodepletion followed by donor lymphocyte infusion can induce a graft-versus-host response without overwhelming toxicity. Bone Marrow Transplantation 47, 1112-1117 (2012). doi: 10.1038/bmt.2011.231
6. Kurosawa, S., Yamaguchi, T., Oshima, K. et al. Resolved versus Active Chronic Graft-versus-Host Disease: Impact on Post-Transplantation Quality of Life. Biol Blood Marrow Transplant 25, 1851-1858 (2019). doi: 10.1016/j.bbmt.2019.05.016
7. Stern, M., de Wreede, L. C., Brand, R. et al. Sensitivity of hematological malignancies to graft-versus-host effects: an EBMT megafile analysis. Leukemia 28, 2235-2240 (2014). doi: 10.1038/leu.2014.145
8. Tan, J., Wang, Y., Yu, S. J. et al. Prognostic factors on graft-versus-host disease-free and relapse-free survival after allogeneic hematopoietic stem cell transplantation for adults with acute leukemia. Leuk Res 59, 1-7 (2017). doi: 10.1016/j.leukres.2017.05.011
9. Putter, H., van der Hage, J., de Bock, G. H., Elgalta, R. & van de Velde, C. J. Estimation and prediction in a multi-state model for breast cancer. Biometrical journal. Biometrische Zeitschrift 48, 366-380 (2006). doi: 10.1002/bimj.200510218
10. Bakunina, K., Putter, H., Versluis, J. et al. The added value of multi-state modelling in a randomized controlled trial: The HOVON 102 study re-analyzed. Cancer medicine 11, 630-640 (2022). doi: 10.1002/cam4.4392
11. Hermans, S. J. F., van Norden, Y., Versluis, J. et al. Benefits and risks of clofarabine in adult acute lymphoblastic leukemia investigated in depth by multi-state modeling. Cancer medicine 13, e6756 (2024). doi: 10.1002/cam4.6756
12. Löwenberg, B., Pabst, T., Maertens, J. et al. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. Blood 129, 1636-1645 (2017). doi: 10.1182/blood-2016-10-740613
13. Iacobelli, S. & Carstensen, B. Multiple time scales in multi-state models. Statistics in Medicine 32, 5315-5327 (2013). doi: 10.1002/sim.5976
14. Podgorny, P. J., Liu, Y., Dharmani-Khan, P. et al. Immune cell subset counts associated with graft-versus-host disease. Biol Blood Marrow Transplant 20, 450-462 (2014). doi: 10.1016/j.bbmt.2014.01.002
15. Baart, S. J., van der Palen, R. L. F., Putter, H. et al. Joint Modeling of Longitudinal Markers and Time-to-Event Outcomes: An Application and Tutorial in Patients After Surgical Repair of Transposition of the Great Arteries. Circ Cardiovasc Qual Outcomes 14, e007593 (2021). doi: 10.1161/circoutcomes.120.007593
16. Ferrer, L., Rondeau, V., Dignam, J. et al. Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer. Stat Med 35, 3933-3948 (2016). doi: 10.1002/sim.6972

7



&

Appendices

Nederlandse samenvatting
List of publications
Curriculum Vitae
Dankwoord

NEDERLANDSE SAMENVATTING

Introductie

Observationele studies op basis van informatie die al in het zorgproces is verzameld kunnen veel inzicht geven in ziektepatronen, behandelingen en risicofactoren, maar hebben als nadeel dat onderzoekers geen controle hebben over wanneer welke metingen worden gedaan. Het analyseren van uitkomsten en effecten van risicofactoren wordt nog ingewikkelder als de behandeling uit verschillende stappen bestaat of als complexe uitkomstmaten worden onderzocht. Hiervoor zijn vaak geavanceerdere statistische methoden nodig. In dit proefschrift worden twee complexe klinische situaties onderzocht, allogene hematopoëtische stamceltransplantatie (alloSCT) voor patiënten met acute leukemie en immuunsuppressieve therapie (IST) voor patiënten met verworven aplastische anemie (AA), waarbij met name twee geavanceerde statistische methoden worden toegepast: het multi-state model en het joint model.

AlloSCT voor patiënten met acute leukemie

In het beenmerg worden continu bloedcellen en bloedplaatjes aangemaakt vanuit hematopoëtische stamcellen (HSC's). Bij acute leukemie is een HSC zich ongeremd gaan delen tot ongedifferentieerde cellen. Voor genezing is vaak een alloSCT nodig. Hierbij worden gezonde HSC's en immuuncellen (waaronder T-cellen) van een donor aan de patiënt gegeven, meestal na enkele chemokuren om zoveel mogelijk leukemiecellen te doden. De doelen van alloSCT zijn vervanging van de hematopoëse van de patiënt door die van de donor en introductie van alloreactieve T-cellen van de donor die de overgebleven leukemiecellen van de patiënt kunnen elimineren: het Graft-versus-Leukemia (GvL) effect. Alloreactieve T-cellen van de donor herkennen deze leukemiecellen door genetische verschillen tussen patiënt en donor. Het is echter ook mogelijk dat de alloreactieve T-cellen van de donor zich richten tegen niet-hematopoëtische weefsels zoals de huid, lever of darm van de patiënt, wat kan leiden tot Graft-versus-Host-Disease (GvHD). Het succes van alloSCT als behandeling tegen acute leukemie hangt af van het ontstaan van voldoende GvL zonder ernstige GvHD. Het GvL-effect zelf is lastig te meten. Als een patiënt een recidief krijgt, was er blijkbaar te weinig GvL, maar bij patiënten die (nog) geen recidief hebben, is het onbekend of dit komt doordat er een voldoende krachtig GvL-effect is.

AlloSCT bestaat uit een aantal stappen: conditionering van de patiënt door middel van chemotherapie, bestraling en/of antilichamen tegen immuuncellen, infusie van het transplantaat, interventies ter voorkoming van ernstige GvHD en, bij sommige strategieën, interventies ter verbetering van het GvL-effect. Al deze stappen beïnvloeden het lymfohematopoëtische herstel van de patiënt en de risico's op GvHD, recidief van de ziekte, infecties en sterfte.
De conditionering kan myeloablatief (MA) of nonmyeloablatief (NMA) zijn, waarbij de tweede vorm minder toxisch is maar ook minder van de gezonde en maligne HSC's van de patiënt doodt: na infusie van het transplantaat moeten de donor-HSC's meer concurreren met de overgebleven HSC's van de patiënt tijdens de competitieve repopulatie van het beenmerg en is het nog meer van belang dat er een GvL-effect ontstaat om de overgebleven leukemiecellen te elimineren. Zolang in een patiënt zowel donor- als patiënt-HSC's gevonden worden, heeft een patiënt gemengd chimerisme

(mixed chimerism, MC). Wanneer alleen nog maar donor-HSC's worden gemeten, heeft een patiënt volledig-donor chimerisme (full donor chimerism, FDC). Als een patiënt een tijd na alloSCT van MC naar FDC converteert, komt dit waarschijnlijk door een alloimmuunrespons van donor-T-cellen tegen de overgebleven hematopoëtische cellen van de patiënt. Deze conversie wordt vaak gebruikt als een indicator voor het bestaan van een GvL-effect. Omdat leukemiecellen hematopoëtische cellen zijn, zullen bij een alloimmuunrespons tegen patiënt-HSC's namelijk ook de leukemiecellen aangevallen worden.
Chimerisme kan ook in andere celpopulaties zoals T-cellen gemeten worden. Gedurende de eerste maanden na alloSCT heeft een patiënt maar weinig T-cellen, omdat de meeste patiënt-T-cellen de conditionering niet overleven en er in het transplantaat onvoldoende T-cellen zitten om deze te vervangen. Als reactie op de lymfopenie delen de patiënt- en donor-T-cellen die er wél zijn zich meer. Dit heet homeostatische proliferatie.
Om ernstige GvHD te voorkomen krijgen patiënten meestal gedurende de eerste maanden na alloSCT profylactische immuunsuppressie. Om het risico op ernstige GvHD verder te verlagen wordt in sommige strategieën zwaardere immuunsuppressie gegeven door middel van bijvoorbeeld T-celdepletie (TCD) op basis van alemtuzumab of toediening van cyclofosfamide vlak na alloSCT (post-transplantatie cyclofosfamide, PTCY). Het nadeel van TCD en PTCY is dat ook de alloreactieve T-cellen die verantwoordelijk zijn voor het GvL-effect, en de niet-alloreactieve T-cellen die verantwoordelijk zijn voor de bescherming tegen virussen, in meer of mindere mate kunnen worden geraakt. Hierdoor hebben patiënten een hoger risico op een recidief en infecties.
Om het GvL-effect na alloSCT te versterken kunnen extra alloreactieve donor-T-cellen aan de patiënt worden toegediend. Dit kan door middel van ongemodificeerde donorlymfocyteninfusies (DLI), die alloreactieve en niet-alloreactieve T-cellen en andere immuuncellen bevatten. DLI's kunnen therapeutisch gegeven worden aan patiënten met een recidief, pre-emptief aan patiënten met MC of minimale restziekte of profylactisch aan alle patiënten zonder GvHD (teken van sterke alloreactiviteit). Hoe hoger de T-celdosis, hoe effectiever en potentieel toxischer de DLI, d.w.z. hoe sterker het GvL-effect en hoe hoger het risico op ernstige GvHD. Volgens de huidige richtlijnen hangt de DLI-dosis af van de indicatie (bij de behandeling van een recidief wordt meer toxiciteit geaccepteerd dan bij DLI ter voorkoming van een recidief), de donor (onverwante donoren hebben meer genetische verschillen met de patiënt dan verwante donoren wat de kans op GvHD vergroot) en de tijd sinds alloSCT (DLI's vroeg na alloSCT geven meer GvHD dan DLI's die later worden gegeven). Ondanks deze aanpassingen is de DLI-alloreactiviteit erg variabel: sommige patiënten overlijden aan ernstige GvHD na DLI, terwijl anderen geen tekenen van GvHD en GvL vertonen en een recidief krijgen.
Sommige strategieën combineren TCD of PTCY alloSCT met profylactische DLI om de balans tussen GvHD en GvL te verbeteren. Het idee hierachter is om de alloSCT in twee stappen uit te voeren: eerst wordt donorhematopoëse geïntroduceerd met een minimaal risico op ernstige GvHD en pas daarna wordt allo-immuniteit van de donor geïntroduceerd om voldoende GvL te bewerkstelligen. Omdat de tweede stap plaatsvindt nadat het initiële herstel heeft plaatsgevonden, komen de alloreactieve T-cellen in een minder pro-inflammatoire omgeving terecht, wat leidt tot een lager risico op GvHD.

Immunosuppressieve behandeling voor patiënten met verworven aplastische anemie

AA is een hematologische ziekte die wordt gekenmerkt door een hypocellulair beenmerg en falen van hematopoëse, wat leidt tot pancytopenie. Hierdoor kunnen patiënten overlijden aan anemie, bloedingen of infecties. In de meeste gevallen lijkt niet-aangeboren AA te worden veroorzaakt door een auto-immuunreactie tegen hematopoëtische cellen. Een van de aanwijzingen hiervoor is dat patiënten met AA vaak een grote (≥1%) populatie van GPI-deficiënte HSC's hebben. Deze cellen hebben een bepaald eiwit veel minder of zelfs helemaal niet op hun celwand waardoor ze mogelijk minder gevoelig zijn voor auto-immuunaanvallen. GPI-deficiënte HSC's kunnen ook in heel lage hoeveelheden bij gezonde mensen voorkomen, maar kunnen bij patiënten met AA uitgroeien tot grotere populaties doordat ze een overlevingsvoordeel hebben ten opzichte van 'normale' HSC's.

De meeste volwassen patiënten met AA krijgen een immuunsuppressieve behandeling op basis van antithymocytenglobuline (ATG, antilichaamtherapie gericht tegen T-cellen) en ciclosporine. Twee derde van de patiënten heeft een respons, maar vaak slechts gedeeltelijk: deze patiënten worden transfusie-onafhankelijk, maar hun bloedwaarden blijven laag. Het kan zes maanden of zelfs langer duren voordat een respons zichtbaar is, omdat de auto-immuunrespons voldoende onderdrukt moet worden, voordat de weinige overgebleven HSC's kunnen uitgroeien. Na het bereiken van een respons wordt de IST afgebouwd met als doel om volledig te stoppen. Bij 30% van de patiënten die reageren, treedt echter een recidief van de ziekte op, waardoor de IST herstart of opgehoogd moeten worden, of zelfs een alloSCT gegeven moet worden. Daarnaast ontwikkelen patiënten met AA die behandeld worden met IST soms klonale evolutie van hematopoëtische cellen, wat uiteindelijk kan leiden tot andere beenmergziekten zoals acute myeloïde leukemie.

Het multi-state model om complexe gebeurtenisreeksen te analyseren

Beide klinische situaties betreffen een complexe behandeling waarbij patiënten vaak meerdere interventies krijgen en meerdere klinische gebeurtenissen meemaken. Om dit goed te analyseren kan een multi-state model gebruikt worden. In zo'n model verplaatsten patiënten zich tussen 'staten' wanneer klinische gebeurtenissen optreden of behandelingen worden gedaan. Een patiënt die bijvoorbeeld overlijdt na GvHD na DLI, gaat van de staat "DLI" via "GvHD" naar "Dood". De paden tussen de staten heten transities (bijvoorbeeld "DLI → GvHD"). Het multi-state model kan voor elke staat berekenen hoe groot de kans is dat een patiënt zich over de tijd hierin bevindt. Ook kan het multi-state model effecten van risicofactoren op elk van de transities schatten. Dit gebeurt meestal door middel van transitie-specifieke Cox proportional hazards modellen. Voor de transitie "DLI → GvHD" zou bijvoorbeeld donortype meegenomen kunnen worden om te berekenen hoeveel groter de kans op GvHD na DLI voor patiënten met een onverwante donor is ten opzichte van patiënten met een verwante donor. Het multi-state model kan de transitie-specifieke Cox modellen vervolgens meenemen in de berekening van het hele model om de prognose voor patiënten met bepaalde karakteristieken te schatten en zo de klinische impact van de risicofactoren te tonen op de kans om in een bepaalde staat te zitten.

Het joint model om effecten van biomarkers te onderzoeken
Biomarkers zijn meetbare indicatoren die informatie geven over wat er in het lichaam aan de gang is. In het voorbeeld van alloSCT kunnen bijvoorbeeld chimerisme en T-celaantallen als biomarker gezien worden: de waardes zeggen iets over de status van het lymfohematopoëtische herstel en kunnen over de tijd gemeten worden. Om de effecten van dit soort longitudinale metingen op de risico's van bepaalde gebeurtenissen zoals GvHD te schatten kan een joint model gebruikt worden. Dit model berekent twee submodellen, een voor de longitudinale metingen en een voor de klinische gebeurtenissen, en koppelt ze via een associatiestructuur. Hierdoor kan het model de trajecten van de longitudinale metingen correct modelleren en associaties tussen de longitudinale metingen en het ontstaan van klinische gebeurtenissen kwantificeren.

Dit proefschrift

In dit proefschrift werd onderzocht hoe het analyseren van specifieke klinische omstandigheden en het toepassen van geavanceerde statistische methodologie op gedetailleerde observationele data gebruikt kan worden om complexe onderzoeksvragen binnen de hematologie te beantwoorden.

In **hoofdstuk 2** probeerden we de effecten van competitieve repopulatie en allo-immunologische druk op het lymfohematopoëtische herstel na alloSCT van elkaar te onderscheiden. We onderzochten een cohort van 281 patiënten met acute leukemie die een TCD alloSCT met alemtuzumab ondergingen, waarna een deel van de patiënten een profylactische of pre-emptieve DLI op 3 maanden kreeg, terwijl de rest geen DLI kreeg of pas vanaf 6 maanden. Hierdoor hadden we een natuurlijke controle- en interventiegroep voor de DLI op 3 maanden. We begonnen met de eerste 3 maanden zónder DLI. In deze periode was de allo-immunologische druk laag door de TCD: het cumulatieve risico op klinisch relevante GvHD, d.w.z. GvHD waarvoor therapeutische systemische immuunsuppressie nodig is, was slechts 13%. Ondanks de lage allo-immunologische druk sloeg bij 99% het transplantaat aan, wat aantoont dat primaire engraftment niet afhankelijk was van MA conditionering of de aanwezigheid van evidente allo-immunologische druk. Het ontstaan van volledige donorhematopoëse was wél afhankelijk van zowel de intensiteit van de conditionering als de aanwezigheid van allo-immunologische druk. Om de impact van de introductie van allo-immunologische druk door DLI te analyseren vergeleken we eerst de metingen op 3 en 6 maanden na alloSCT tussen patiënten met een DLI op 3 maanden die daarna geen GvHD ontwikkelden en patiënten die geen DLI op 3 maanden kregen en geen GvHD ontwikkelden. De laatste groep vertoonde stabiel beenmergchimerisme in deze periode. Daarentegen converteerden de patiënten in de DLI-groep vaak naar FDC. Dit laat zien dat de 3-maanden-DLI chimerismeconversie kan induceren, zelfs bij afwezigheid van GvHD. De kinetiek van CD4+ en CD8+ T-celchimerisme vertoonde vergelijkbare patronen met toenemende niveaus van donorchimerisme bij de patiënten met DLI, maar niet bij degenen zonder DLI. Dit suggereert dat voor het tot stand brengen van een volledig donor-T-celcompartiment enige allo-immunologische druk nodig is. Ten slotte onderzochten we de allo-immunologische effecten van de totale DLI-strategie bij alle patiënten met gemengde hematopoëse ten tijde van hun eerste DLI. Van de 65 patiënten converteerde 72% naar FDC, van wie slechts 34% klinisch relevante GvHD ontwikkelde. Deze resultaten illustreren dat het GvL-effect kan worden gescheiden van GvHD.

In **hoofdstuk 3** onderzochten we de complexe associaties tussen de kinetiek van circulerende immuuncelaantallen en alloreactiviteit door middel van een joint model. We keken naar dezelfde klinische setting als in hoofdstuk 2, behalve dat we alleen NMA geconditioneerde patiënten (n=166) includeerden omdat zij geen post-alloSCT GvHD-profylaxe kregen die de immuuncelkinetiek zou kunnen beïnvloeden. We onderzochten eerst het effect van de 3-maanden-DLI op de kinetiek van T- en NK-celaantallen na TCD alloSCT. Hiervoor construeerden we een joint model voor de eerste 6 maanden vanaf alloSCT dat twee groepen met elkaar vergeleek in een intention-to-treat-benadering: patiënten die een indicatie voor een 3-maanden-DLI hadden vanwege een verwacht hoog risico op recidief (de 'hoog-risico'-groep) en patiënten die hier geen indicatie voor hadden (de 'niet-hoog-risico'-groep). Vergeleken met patiënten met een verwante donor hadden patiënten met een onverwante donor, die ook ATG bij de conditionering hadden gekregen, lagere T-celaantallen gedurende de eerste 3 maanden na alloSCT. Dit illustreert het lang aanhoudende effect van ATG. Bij patiënten met een onverwante donor begonnen de T-celtrajecten vanaf 3 maanden te divergeren tussen de groepen met een hoog en niet-hoog risico, waarbij degenen met indicatie voor een 3-maanden-DLI hogere T-celaantallen hadden. Aangezien de enige plausibele verklaring voor deze toename de 3-maanden-DLI is, toont dit aan dat DLI detecteerbare T-celexpansie kan geven. Opvallend genoeg zagen we geen divergentie tussen de risicogroepen met een verwante donor. Hierna keken we naar risicofactoren voor o.a. GvHD en recidief. Het behoren tot de hoog-risico-groep en het hebben van hogere aantallen van CD3+- en CD4+-T-cellen waren geassocieerd met een hoger risico op GvHD, terwijl een hoger aantal CD4+-T-cellen het risico op recidief verlaagde. Een hoger aantal NK-cellen was geassocieerd met een hoger risico op GvHD en een lager risico op recidief. De GvHD-associatie verdween echter toen zowel CD4+ T-cellen als NK-cellen werden meegenomen in een exploratieve analyse. Dit suggereert dat de waargenomen associatie tussen het aantal NK-cellen en GvHD slechts de hoge correlatie tussen het aantal NK-cellen en CD4+-T-cellen weerspiegelde, en niet een direct effect van het aantal NK-cellen op het risico op GvHD. Om de T-celkinetiek na de 3-maanden-DLI verder te onderzoeken construeerden we een tweede joint model vanaf de 3-maanden-DLI, waarbij we alleen de patiënten includeerden die daadwerkelijk deze DLI hadden gekregen. Het hebben van een onverwante donor en hogere aantallen van CD3+-, CD4+- en CD8+-T-cellen waren allemaal geassocieerd met een hoger risico op GvHD gedurende de eerste 3 maanden na deze DLI.

In **hoofdstuk 4** probeerden we factoren te identificeren die de alloreactiviteit van DLI beïnvloeden. We onderzochten hiervoor een cohort van patiënten met acute leukemie die hun eerste DLI op 3 (n = 88) of 6 (n = 76) maanden na TCD alloSCT met alemtuzumab hadden gekregen. We keken eerst naar de twee factoren die de DLI-dosis bepaalden: tijd sinds alloSCT en donortype. Het tienvoudige dosisverschil tussen de 3- en 6-maanden-DLI resulteerde in vergelijkbare risico's op GvHD. Na MA conditionering was voor beide DLI's de 50% dosisverlaging in geval van een onverwante donor voldoende om de GvHD-risico's tussen patiënten met verwante en onverwante donor gelijk te trekken. Daarentegen hadden NMA geconditioneerde patiënten met een onverwante donor nog steeds een hoger risico op GvHD dan NMA geconditioneerde patiënten met een verwante donor. Vervolgens richtten we ons op drie omstandigheden ten tijde van DLI die de T-celactivatie zouden kunnen bevorderen: de aanwezigheid van antigeenpresenterende cellen van de patiënt (geschat op basis van het

beenmergchimerisme), lymfopenie en de aanwezigheid van een virale infectie rond DLI. Om de effecten van deze risicofactoren op de ontwikkeling van klinisch relevante GvHD, het risico op overlijden tijdens GvHD en op de totale klinische uitkomst na DLI te schatten construeerden we een multi-state model vanaf de eerste DLI. Bij de 3-maanden-DLI verhoogden virale infecties rond DLI het risico op GvHD (hazard ratio [HR] 3.7). Ten tijde van de 6-maanden-DLI waren virale infecties zeldzaam. Bij deze DLI verhoogde de aanwezigheid van ≥5% MC het risico op GvHD significant (HR 3.6) en zagen we een trend van meer GvHD bij 1-5% MC en lymfopenie. We vonden geen significante effecten van virale infecties bij de 3-maanden-DLI of ≥5% MC bij de 6-maanden-DLI op het risico op overlijden tijdens GvHD, wat waarschijnlijk kwam door het lage aantal sterfgevallen. Tot slot keken we naar de kans om op 6 maanden na DLI zonder actieve GvHD en zonder recidief in leven te zijn (current GvHD-relapse-free survival, cGRFS). Bij de 3-maanden-DLI was de cGRFS 61% voor patiënten zonder een virale infectie in de laatste week voor DLI versus 31% voor patiënten met een virale infectie. Bij de 6-maanden-DLI keken we naar de modelgebaseerde kans op cGRFS voor twee referentiepatiënten met verschillende karakteristieken. Beide waren MA geconditioneerd met een verwante donor, maar de eerste had FDC ten tijde van de 6-maanden-DLI en zou 77% kans hebben om 6 maanden later zonder recidief of actieve GvHD in leven te zijn, terwijl de tweede ≥5% MC had en slechts 44% kans zou hebben. De sterke impact van virale infecties en beenmergchimerisme op de cGRFS onderstrepen de klinische relevantie van onze bevindingen.

In **hoofdstuk 5** onderzochten we hoe de transplantatiestrategie de alloreactiviteit van DLI beïnvloedt door een andere klinische setting te bekijken, DLI na alloSCT met PTCY, en de uitkomsten te vergelijken met onze studies in de hoofdstukken 2 en 4. In de PTCY-setting werd de lage-dosis-DLI op 4 maanden na alloSCT gegeven in plaats van 3 maanden. We onderzochten eerst de risicofactoren die we in hoofdstuk 4 hadden geïdentificeerd. Alle risicofactoren kwamen zelden voor: slechts 5% van de 83 patiënten die een DLI op 4 of 6 maanden kregen had een virale infectie rond DLI, 6% had ≥5% MC in het beenmerg en 17% had lymfopenie, veel minder dan wat we hadden waargenomen in de alemtuzumab-setting (respectievelijk 19%, 27% en 47%). Vervolgens keken we naar de ontwikkeling van klinisch relevante GvHD na DLI. Zoals verwacht op basis van de lage aanwezigheid van de risicofactoren was het risico op GvHD zeer laag: 4% op 3 maanden na DLI. Slechts één patiënt stierf aan GvHD na een 6-maanden-DLI te hebben ontvangen met op dat moment 14% MC in het beenmerg. De gecombineerde resultaten van hoofdstuk 4 en hoofdstuk 5 tonen aan dat de transplantatiestrategie een grote impact heeft op de omstandigheden ten tijde van DLI, die weer de DLI-alloreactiviteit beïnvloeden. Om te onderzoeken of DLI na PTCY alloSCT nog steeds conversie van MC naar FDC kon induceren, onderzochten we de chimerismekinetiek in het beenmerg van de 28 patiënten met MC ten tijde van hun eerste DLI: 79% converteerde naar FDC, van wie slechts 9% klinisch relevante GvHD ontwikkelde. Het conversiepercentage is vergelijkbaar met wat we in hoofdstuk 2 hadden gezien, terwijl het risico op GvHD lager was. Geen van de geconverteerde patiënten kreeg een recidief, wat duidt op een sterk genoeg GvL-effect.

In **hoofdstuk 6** wilden we een maat voor "behandelsucces" definiëren en meten die de complexe patronen van herstel en falen kan meenemen van patiënten met immunosuppressieve therapie tegen aplastische anemie. We definieerden drie niveaus van

behandelsucces. Het algemene doel van IST voor AA is het bereiken en behouden van transfusieonafhankelijkheid zonder de ontwikkeling van secundaire beenmergziekten. Dit is gevat in de uitkomst Disease-free survival (DFS). DFS zou bij voorkeur bereikt moeten worden zonder dat een alloSCT nodig is: Transplantation-free DFS (T-DFS). Het uiteindelijke doel is om alle therapie voor AA te stoppen nadat een respons is bereikt: Transplantation- and Treatment-free DFS (TT-DFS). We hebben deze uitkomsten gemeten door middel van een multi-state model op een cohort van 127 transfusie-afhankelijke patiënten met AA. De 5-jaarskansen op DFS, T-DFS en TT-DFS waren respectievelijk 70%, 60% en 42%. De kans om transfusie-onafhankelijk te worden was hoger indien een GPI-deficiënte celkloon van ≥1% aanwezig was (HR 2.2) en lager bij een leeftijd van 40 jaar of ouder en bij ernstige of zeer ernstige AA (HR's tussen 0.4 en 0.5). We zagen geen significante effecten van deze risicofactoren op de kans om alle therapie voor AA te kunnen stoppen na het bereiken van transfusie-onafhankelijkheid. Zoals verwacht was een leeftijd van 60 jaar of ouder een sterke voorspeller voor het risico op overlijden (HR 7.3). Om de impact van deze risicofactoren op de 5-jaarsuitkomsten te tonen berekenden we modelgebaseerde prognoses voor referentiepatiënten met verschillende kenmerken. Zo was de geschatte 5-jaarskans op TT-DFS voor een patiënt van 40 jaar of jonger met ernstige aplastische anemie zonder een GPI-deficiënte celkloon 47%, vergeleken met 24% voor een patiënt van ten minste 60 jaar met dezelfde kenmerken, en 41% voor een patiënt van ten minste 60 jaar met een GPI-deficiënte celkloon en niet-ernstige AA. Deze resultaten laten zien dat de drie risicofactoren de succeskans van IST voor AA op de lange termijn sterk beïnvloeden.

Samenvattend laat het onderzoek in dit proefschrift zien hoe gedetailleerde observationele data en geavanceerde statistische methodes gebruikt kunnen worden om complexe klinische en immunologische onderzoeksvragen te beantwoorden. Multi-state en joint modellen vereisen data van hoge kwaliteit. Voor het analyseren van risicofactoren zijn bovendien voldoende patiënten nodig die het relevante eindpunt bereiken. Voldoende data van voldoende kwaliteit kan het beste verkregen worden door middel van gedetailleerde dataverzameling in samenwerkingsverbanden. Daarnaast is voor het correct toepassen en interpreteren van deze modellen zowel statistische als klinische kennis nodig. Meer samenwerking tussen klinische en methodologische onderzoekers is hiervoor essentieel.

LIST OF PUBLICATIONS

- McWilliams O, de Boer MGJ, **Koster EAS**[*], de Wreede LC[*]. Cause-specific excess mortality during the COVID-19 pandemic in the Netherlands. Submitted
- **Koster EAS**, Sluiskes MH, Putter H, Rosendaal FR, van Hylckama Vlieg A, de Boer MGJ, de Wreede LC. Dynamics of infection, vaccination and excess mortality during the COVID-19 pandemic among older individuals – a nationwide analysis. Submitted
- Halkes CJM[†], **Koster EAS**[†], Bogers EJM, Heubel-Moenen FCJI, Daenen LGM, Klein SK, Langemeijer SMC, Nur E, Raaijmakers MHG, Snijders TJF, Tjon JML, de Wreede LC. Anti-thymocyte globulin-based treatment frequently leads to enduring treatment success in both old and young adult patients with aplastic anaemia: a real-world analysis from the Dutch Aplastic Anaemia Registry. Submitted
- **Koster EAS**, von dem Borne PA, van der Hem JGK, van Balen P, Marijt EWA, Tjon JML, Snijders TJF, van Lammeren D, Veelken H, Falkenburg JHF, de Wreede LC, Halkes CJM. Transplantation strategy affects the risk of GvHD after prophylactic and preemptive donor lymphocyte infusion. Ann Hematol. 2025 Oct;104(10):5379-5387. doi: 10.1007/s00277-025-06662-x
- Sluiskes MH, **Koster EAS**, Goeman JJ, Rodríguez-Girondo M, Putter H, de Wreede LC. A new framework for disentangling different components of excess mortality applied to Dutch care home residents during Covid-19. BMC Med Res Methodol. 2025 May 10;25(1):126. doi: 10.1186/s12874-025-02579-1
- Fuchs KJ[†], Göransson M[†], Kester MGD, Ettienne NW, van de Meent M, de Jong RCM, **Koster EAS**, Halkes CJM, Scheeren F, Heemskerk MHM, van Balen P, Falkenburg JHF, Hadrup SR, Griffioen M. DNA barcoded peptide-MHC multimers to measure and monitor minor histocompatibility antigen-specific T cells after allogeneic stem cell transplantation. J Immunother Cancer. 2024 Dec 9;12(12): e009564. doi: 10.1136/jitc-2024-009564
- Caram-Deelder C, Vlieg AVH, Groenwold RHH, Chen Q, Mook-Kanamori DO, Dekkers OM, **Koster EAS**, de Wreede LC, van Nieuwkoop C, Toshkov DD, Rosendaal FR. Excess mortality during the first 2 years of the COVID-19 pandemic (2020-2021) in the Netherlands: Overall and across demographic subgroups. IJID Reg. 2024 Nov 27;14:100500. doi: 10.1016/j.ijregi.2024.100500
- **Koster EAS,** von dem Borne PA, van Balen P, Marijt EWA, Tjon JML, Snijders TJF, van Lammeren D, Veelken H, Falkenburg JHF, Halkes CJM, de Wreede LC. Risk factors for graft-versus-host-disease after donor lymphocyte infusion following T-cell depleted allogeneic stem cell transplantation. Front Immunol. 2024 Mar 13;15:1335341. doi: 10.3389/fimmu.2024.1335341
- Fuchs KJ, van de Meent M, Honders MW, Khatri I, Kester MGD, **Koster EAS**, Koutsoumpli G, de Ru AH, Van Bergen CAM, van Veelen PA, 't Hoen PAC, van Balen P, van den Akker EB, Veelken H, Halkes CJJ, Falkenburg JHF, Griffioen M. Expanding the repertoire reveals recurrent, cryptic and hematopoietic HLA class I minor histocompatibility antigens. Blood. 2024 Mar 1:blood.2023022343. doi: 10.1182/blood.2023022343

- **Koster EAS**†, Bonneville EF†, von dem Borne PA, van Balen P, Marijt EWA, Tjon JML, Snijders TJF, van Lammeren D, Veelken H, Putter H, Falkenburg JHF, Halkes CJM, de Wreede LC. Joint models quantify associations between immune cell kinetics and allo-immunological events after allogeneic stem cell transplantation and subsequent donor lymphocyte infusion. Front Immunol. 2023 Aug 1;14:1208814. doi: 10.3389/fimmu.2023.1208814
- van der Zouwen B, **Koster EAS**, von dem Borne PA, Oosten LEM, Roza-Scholten MWI, Snijders TJF, van Lammeren D, van Balen P, Marijt WAF, Veelken H, Falkenburg JHF, de Wreede LC, Halkes CJM. Feasibility, safety, and efficacy of early prophylactic donor lymphocyte infusion after T cell-depleted allogeneic stem cell transplantation in acute leukemia patients. Ann Hematol. 2023 May;102(5):1203-1213. doi: 10.1007/s00277-023-05145-1
- **Koster EAS**, von dem Borne PA, van Balen P, van Egmond EHM, Marijt EWA, Veld SAJ, Jedema I, Snijders TJF, van Lammeren D, Veelken H, Falkenburg JHF, de Wreede LC, Halkes CJM. Competitive Repopulation and Allo-Immunologic Pressure Determine Chimerism Kinetics after T Cell-Depleted Allogeneic Stem Cell Transplantation and Donor Lymphocyte Infusion. Transplant Cell Ther. 2023 Apr;29(4):268.e1-268.e10. doi: 10.1016/j.jtct.2022.12.022
- Huisman W, Roex MCJ, Hageman L, **Koster EAS**, Veld SAJ, Hoogstraten C, van Balen P, van Egmond HME, Van Bergen CAM, Einsele H, Germeroth L, Amsen D, Falkenburg JHF, Jedema I. Tracking the Progeny of Virus-specific T-cell Products in Patients Post-Transplant using TCR-Sequencing. Blood Adv. 2022 Sep 19:bloodadvances.2022007270. doi: 10.1182/bloodadvances.2022007270
- Laghmouchi A, Kester MGD, Hoogstraten C, Hageman L, de Klerk W, Huisman W, **Koster EAS**, de Ru AH, van Balen P, Klobuch S, van Veelen PA, Falkenburg JHF, Jedema I. Promiscuity of Peptides Presented in HLA-DP Molecules from Different Immunogenicity Groups Is Associated With T-Cell Cross-Reactivity. Front Immunol. 2022 Feb 16;13:831822. doi: 10.3389/fimmu.2022.831822
- Bakunina K, Putter H, Versluis J, **Koster EAS**, van der Holt B, Manz MG, Breems DA, Gjertsen BT, Cloos J, Valk PJM, Passweg J, Pabst T, Ossenkoppele GJ, Löwenberg B, Cornelissen JJ, de Wreede LC. The added value of multi-state modelling in a randomized controlled trial: The HOVON 102 study re-analyzed. Cancer Med. 2021 Dec 24. doi: 10.1002/cam4.4392
- Rienstra R, **Koster EAS**, Janssen CCAH. Methotrexate induced peritonitis: diagnosis per exclusionem. J Obstet Gynaecol. 2021 Jul;41(5):835-837. doi: 10.1080/01443615.2020.1756763
- Robbers Y, **Koster EA**, Krijbolder DI, Ruijs A, van Berloo S, Meijer JH. Temporal behaviour profiles of Mus musculus in nature are affected by population activity. Physiol Behav. 2015 Feb;139:351-60. doi: 10.1016/j.physbeh.2014.11.020

* Shared last authorship
† Shared first authorship

CURRICULUM VITAE

Eva Koster werd op 13 januari 1993 geboren in Leiden. Na het behalen van haar diploma aan het Stedelijk Gymnasium Leiden in 2011 begon ze aan de studie Geneeskunde in dezelfde stad. Tijdens de bachelor deed ze onderzoek naar de biologische klok op de sectie Neurofysiologie van de afdeling Cel en Chemische biologie (voorheen Moleculaire Celbiologie) van het LUMC onder prof. dr. J.H. Meijer, bij wie ze via het Pre-University College ook haar profielwerkstukonderzoek had gedaan. Daarnaast zat ze een jaar in het bestuur van haar studentendispuut en voltooide ze het Honours College van de faculteit Geneeskunde. Ze werd geselecteerd voor het MD-PhD traject, waarvoor ze onder prof. dr. J.H.F. Falkenburg, dr. C.J.M. Halkes en dr. I. Jedema op de afdeling Hematologie van het LUMC onderzoek deed naar het effect van conditionering en graftmanipulatie op de immuunreconstitutie na allogene stamceltransplantatie. Na haar coschappen ontwikkelde ze voor haar wetenschapsstage op dezelfde afdeling een multi-state model om het klinische beloop na T-cel-gedepleteerde allogene stamceltransplantatie en donorlymfocyteninfusie te evalueren onder leiding van dr. C.J.M. Halkes en dr. L.C. de Wreede. Eind 2017 behaalde ze haar artsentitel waarna ze aansluitend in 2018 haar observationele onderzoek onder prof. dr. J.H.F. Falkenburg met dr. C.J.M. Halkes en dr. L.C de Wreede als co-promoteren voortzette in het kader van een promotietraject. In 2023 werkte ze vier maanden als ANIOS op de afdeling Hematologie van het HagaZiekenhuis in Den Haag. Hierna stapte ze over naar de sectie Medische Statistiek van de afdeling Biomedical Data Sciences van het LUMC om onder dr. L.C. de Wreede onderzoek te doen naar oversterfte tijdens de coronapandemie in Nederland. Dit heeft geleid tot twee andere projecten die momenteel nog lopen: ontwikkeling van geavanceerde statistische methodes om de pandemische paraatheid te vergroten (een samenwerkingsproject met o.a. de sectie Medische Besliskunde van de afdeling Biomedical Data Sciences en de universiteit van Ljubljana) en een onderzoek naar het effect van orale anticoagulantia op oversterfte tijdens de coronapandemie in samenwerking met de afdeling Klinische Epidemiologie van het LUMC. Als hobby vertaalt ze sinds 2011 de Aeneis van Vergilius.

DANKWOORD

Een PhD-traject lijkt op een zware bergtocht: het is op sommige momenten lang en zwaar, maar al met al is het een prachtige reis. Ik wil graag iedereen bedanken die mij op deze reis geholpen heeft, zowel tijdens de zware stukken als op de momenten dat we even op een top stilstonden om te genieten van wat al is bereikt.

Mijn eerste ervaringen met onderzoek had ik bij de MCB Neurofysiologie onder prof. dr. J.H. Meijer. Joke, dank voor je enthousiasme en adviezen, ook in de jaren na mijn tijd op jouw afdeling.

Voor mijn PhD-traject ben ik naar de Hematologie gegaan. Ik wil mijn promotor prof. dr. J.H.F. Falkenburg en mijn co-promotoren dr. C.J.M. Halkes en dr. L.C. de Wreede (Biomedical Data Sciences) heel erg bedanken voor hun jarenlange begeleiding. Ze waren een veelzijdig team waarbij ik van ieder andere vaardigheden leerde om een betere onderzoeker te worden. Fred, dank voor jouw 'helikopter'-blik die me uitdaagde om steeds weer uit te zoomen en het grotere plaatje van mijn onderzoek te bekijken. Stijn, jouw aanstekelijke enthousiasme en nieuwsgierigheid zorgden ervoor dat ik elke keer weer gretig ging nadenken over mogelijke nieuwe analyses. Liesbeth, jouw scherpe blik op methodologisch gebied was essentieel om mijn studies goed uit te voeren. Je hebt je liefde voor de statistiek aan me overgebracht!

Ik wil ook mijn andere collega's op de afdeling Hematologie bedanken. Er was een mooie balans tussen kennisuitwisseling op de dinsdagochtendseminars of tussendoor en gezelligheid tijdens o.a. de lunchpauzes en de jaarlijkse Labdag. Ook de congresbezoeken waren prachtige ervaringen. Dank voor de goede werksfeer op C2 en C5! Ik voelde me op beide plekken heel erg thuis. Hoewel er hard gewerkt werd, was er altijd ruimte voor een goed gesprek. Veel dank voor jullie betrokkenheid en de leuke momenten! Lisette, Sylwia en Henk, dank jullie wel voor het verzamelen en verwerken van de data (wat een klus!). Boris, dank voor je steun bij het valideren. Joost, dank voor je hulp bij de dataverzameling. Gerrie en Esther, dank voor al het regel- en organiseerwerk.

Ik wil daarnaast mijn collega's op de sectie Medische Statistiek van de afdeling Biomedical Data Sciences bedanken. Ten eerste Ed, met wie ik een van mijn hoofdstukken heb geschreven. Ik heb erg genoten van onze samenwerking en veel van je geleerd! Ook dank aan alle PhD-kandidaten en post-docs. Jullie zijn een hartelijke groep die altijd voor elkaar klaar staat. Ik heb veel gehad aan jullie steun gedurende de laatste jaren van mijn PhD.

Natuurlijk ook veel dank aan Maartje en Anoeshka, mijn beste vriendinnen en paranimfen. Jullie waren er voor me op zowel de goede als de zware momenten. Ik ben heel blij dat jullie tijdens mijn verdediging naast me staan! Tot slot wil ik graag mijn familie en vrienden bedanken voor hun betrokkenheid, goede adviezen en fijne gesprekken. Een paar wil ik in het bijzonder noemen. Lieve pappa en mamma, dank voor de goede discussies, jullie bereidheid om af en toe mee te lezen, en bovenal jullie steun zowel binnen als buiten mijn PhD. Ype, dank voor je enthousiasme. Timo, Diede en Laurens, dank voor jullie luisterend oor en bemoedigende woorden!

&