Model-based therapy in allogeneic hematopoietic cell transplantation

Jurgen Bastiaan Langenhorst

Thesis University Medical Center Utrecht, Utrecht University, Infection & Immunity, Laboratory of Translational Immunology, Group Boelens-Nierkens

ISBN:	978-94-6380-316-8
Print:	Proefschriftmaken —— www.proefschriftmaken.nl
Cover design:	Rick van Drie design —— www.rickvandrie.nl
Copyright:	©2019, Jurgen Langenhorst

Model-based therapy in allogeneic hematopoietic cell transplantation
Therapie gebaseerd op wiskundige modellen in allogene stamcel transplantatie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 14 mei 2019 des middags te 2.30 uur

door

Jurgen Bastiaan Langenhorst

geboren op 30 april 1988 te Hoogeveen

Promotor:	Prof. dr. A.D.R. Huitema
Copromotoren:	Dr. J.J. Boelens
	Dr. C. van Kesteren

The work presented in this thesis was supported by Foundation Children Cancerfree (KiKa) under project number 190.

Beoordelingscommissie:	Prof. dr. A.C.G. Egberts Prof. dr. R. Pieters Prof. dr. J.H. Beijnen Dr. E. Svensson Prof. dr. H.J. Vormoor
Paranimfen:	Drs. K.J. Polman Dhr. M. Favier

Contents

General introduction

1.1 Allogeneic Hematopoietic cell transplantation

Allogeneic hematopoietic cell transplantation (HCT) is a high-risk procedure with a curative intention. Indications for HCT include a variety of malignant and benign disorders[1]. Malignancies treated with HCT are relapsed cancers with an origin in immune cells and/or the bone marrow, such as leukemia (acute/chronic myeloid/lymphocytic), lymphoma, plasma cell disorders and myelodysplastic syndrome (MDS). While these relapsed malignancies are prevalent in both children and adults, the benign disorders are mostly found in children and include bone marrow failures, immune deficiencies, metabolic-, and autoimmune diseases. Albeit the curative intention for otherwise often infaust diagnoses, the procedure comes at a high cost, with an overall mortality of around 40%. In all settings, a substantial amount of deaths are attributable to the treatment.
A schematic overview of the procedure is provided in figure 1. First, a suitable donor has to be found, which is often referred to as donor selection, and elaborated on in subsection 1.2. When a donor is available, the immune system of the host has to be depleted, which is called conditioning, and described further in subsection 1.3. Shortly after conditioning the graft can be infused or transplanted in the 'empty' patient. Donor stem cells of the graft have to home towards the empty bone marrow, a process called engraftment. From there on the reconstitution of a new immune system can take place, as summarized in subsection 1.4.

Between conditioning and immune reconstitution, a critical period exists where the body is relatively unprotected against pathogens[2]. Furthermore, in malignant settings, the graft-versus-leukemia effect (GVL) effect has to take place. This GVL is the immunotherapeutic component of the procedure and should commence as early as possible after transplantation, to prevent relapse[3]. As donor T-lymphocytes are crucial in exhibiting GVL[4–7], swift reconstitution of these specific cells is essential in maximizing disease control. A downside of the immunological effects of donor T-lymphocytes is the prevalence of graft-versus-host disease (GvHD) counterbalancing GVL[3]. GvHD arises when the immune cells from the infused graft recognize host-cells as hostile and start to form an immunological response. Especially the more severe cases of GvHD associate with a high mortality rate[8].
These opposing risks of GvHD versus GVL are also the hallmark of other components of HCT. For example, donor selection influences the GVL-GvHD ratio, where a larger amount or more active immune cells infused can possibly exhibit a stronger GVL effect, but also increase the risk of GvHD. In addition, conditioning needs to be sufficiently intense to assure engraftment

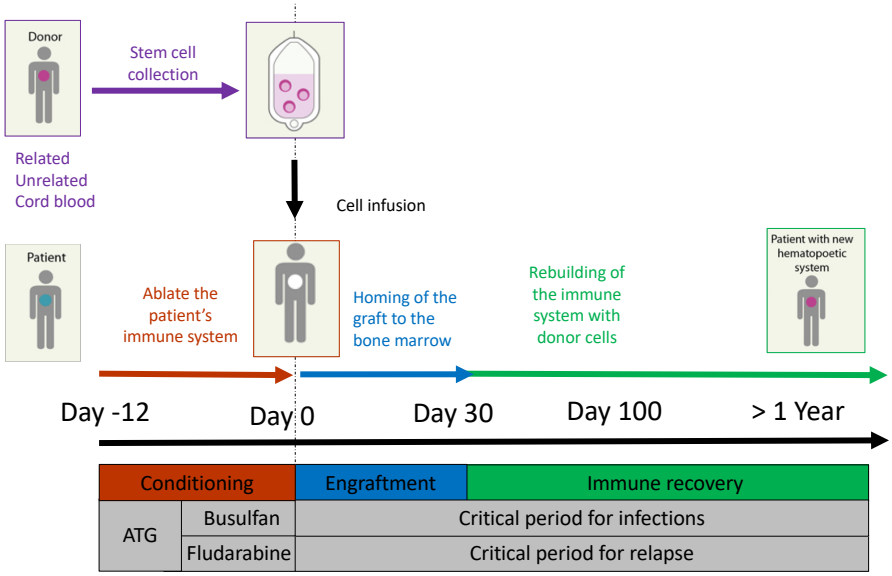

Figure 1: General scheme of UMC-U myeloablative hematopoietic cell transplantation setting. Different colours indicate the crucial parts of the procedure. ATG = Anti-thymocyte globulin.

and disease control, but not so intense that it induces irreversible toxicity or even death. Maintaining a delicate balance between such opposing effects, calls for sophisticated analysis techniques. Choices made during treatment, regarding the directly modifiable components such as donor selection and conditioning, need to be quantitatively associated with not directly modifiable outcome of the procedure, such as the engraftment, immune reconstitution, GvHD, relapse, and eventually survival. Throughout this thesis, mathematical models are constructed to find associations between the different HCT components that can be influenced and the resulting outcomes, in order to fine-tune them for optimal use.

1.2 Donor selection

Historically, the preferred donor source is a sibling that is matched (matched sibling donor (MSD)) to the patient via human leukocyte antigen (HLA).

However, such a donor is not always available, necessitating the search for alternative donors. In early HCT days, the next step has been to search for a HLA-matched unrelated donor (MUD) via international donor-banks. Nowadays, alternatives such as cord blood (CB) or haplo-identical donors have been proven equally well or better than MUD or even MSD in several indications[1]. In chapter 2 the outcome of HCT is compared between cell sources in children and young adults transplanted for acute myeloid leukemia (AML).

1.3 Conditioning

The conditioning regimen is administered to ablate the bone marrow and the immune system of the host. It consists of high-dose radiation and/or chemotherapy. The intensity ranges from non-myeloablative, to reduced intensity onto myeloablative regimens. The less intensive regimens rely more on GVL and are generally less toxic, while conditioning with higher intensity has more inherent anti-leukemic activity, but can also lead to higher toxicity. Throughout the second section, the frequently used myeloablative busulfan (Bu) combined with fludarabine (Flu) with or without the addition of clofarabine or rabbit anti-thymocyte globulin (rATG) is investigated. Firstly, in chapter 4 the current knowledge of why and how to use BuFlu is described. Next, further individualization of the BuFlu regimen is pursued. As optimal exposures have been found for rATG[5,9,10] and busulfan[11] previously, the focus in this section is to expand on these findings as follows. In order to quantify pharmacokinetics for both research and future clinical implementation, a method to quantify both busulfan, fludarabine and clofarabine in one plasma sample is developed in chapter 5. In chapter 6 a semi-mechanistic pharmacokinetic model is constructed to further elucidate a previously identified decrease in busulfan clearance over the course of conditioning[12], ultimately aiming at further optimizing dosing of busulfan. The population pharmacokinetics of fludarabine in an HCT population of all ages is described in chapter 7. Subsequently, the relationship between fludarabine exposure and treatment outcome is described in chapter 8, eventually leading to the definition of a target fludarabine exposure. Finally, a new HCT clinical trial simulation platform is developed and applied to the fludarabine exposure response relationship in chapter 9. Herein, the goal is to find optimal design and the expected results of a fludarabine individualized dosing trial.

1.4 Immune reconstitution

In the period between ablation of the host immune system and engraftment of the donor cells, patients lack immunity as a whole. Shortly after engraftment the first innate immune cells are observed: within a few weeks cells appear and in a few months full recovery is expected[7]. Adaptive immunity however, takes longer to develop (1 month to 1 year) and variability herein is high: some patients have adequate T-cell levels within 3 months after HCT, while for others this takes more than 6 months. As innate recovery generally precedes adaptive immune recovery[7], it might influence and/or predict CD4+ T-cell reconstitution and thereby (indirectly) predict outcome after HCT. In chapter 3 the relationship between adaptive and innate immune recovery within individuals and factors influencing innate reconstitution are explored.

1.5 Aim

In general, this thesis aims to further optimize HCT in children and adults. Mathematical models will be used to find sources that predict the current heterogeneity in HCT outcome, while quantifying the unaccounted variability. The same models will be used to simulate effects of proposed treatment alterations, thereby allowing to provide rational suggestions for optimizing current treatment strategies. This framework of quantifying, predicting and simulating HCT outcome can be summarized in the term 'Model-based therapy in allogeneic hematopoietic cell transplantation'.

References

1. Majhail, N. S., Farnia, S. H., Carpenter, P. A., Champlin, R. E., Crawford, S., Marks, D. I., *et al.* Indications for Autologous and Allogeneic Hematopoietic Cell Transplantation: Guidelines from the American Society for Blood and Marrow Transplantation. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **21,** 1863–1869 (11 Nov. 2015).
2. Admiraal, R., de Koning, C., Lindemans, C. A., Bierings, M. B., Wensing, A. M., Versluys, A. B., *et al.* Viral Reactivations and Associated Outcomes in Context of Immune Reconstitution after Pediatric Hematopoietic Cell Transplantation. *J Allergy Clin Immunol* (2017).
3. Negrin, R. S. Graft-versus-host disease versus graft-versus-leukemia. *Hematology. American Society of Hematology. Education Program* **2015,** 225–230 (2015).
4. Admiraal, R., Chiesa, R., Lindemans, C. A., Nierkens, S., Bierings, M. B., Versluijs, A. B., *et al.* Leukemia-free survival in myeloid leukemia, but not in lymphoid leukemia, is predicted by early CD4+ reconstitution following unrelated cord blood transplantation in children: a multicenter retrospective cohort analysis. *Bone Marrow Transplant* **51,** 1376–1378 (2016).
5. Admiraal, R., van Kesteren, C., Jol-van der Zijde, C. M., Lankester, A. C., Bierings, M. B., Egberts, T. C., *et al.* Association between anti-thymocyte globulin exposure and CD4+ immune reconstitution in paediatric haemopoietic cell transplantation: a multicentre, retrospective pharmacodynamic cohort analysis. *Lancet Haematol* **2,** e194–203 (2015).
6. Bartelink, I. H., Belitser, S. V., Knibbe, C. A., Danhof, M., de Pagter, A. J., Egberts, T. C., *et al.* Immune reconstitution kinetics as an early predictor for mortality using various hematopoietic stem cell sources in children. *Biol Blood Marrow Transplant* **19,** 305–13 (2013).
7. De Koning, C., Plantinga, M., Besseling, P., Boelens, J. J. & Nierkens, S. Immune Reconstitution after Allogeneic Hematopoietic Cell Transplantation in Children. *Biol Blood Marrow Transplant* **22,** 195–206 (2016).

8. Castilla-Llorente, C., Martin, P. J., McDonald, G. B., Storer, B. E., Appelbaum, F. R., Deeg, H. J., *et al.* Prognostic factors and outcomes of severe gastrointestinal GVHD after allogeneic hematopoietic cell transplantation. *Bone marrow transplantation* **49,** 966–971 (7 July 2014).

9. Admiraal, R. & Boelens, J. J. Individualized conditioning regimes in cord blood transplantation: Towards improved and predictable safety and efficacy. *Expert Opinion on Biological Therapy* **16,** 801–813 (2016).

10. Admiraal, R., Nierkens, S., de Witte, M. A., Petersen, E. J., Fleurke, G. J., Verrest, L., *et al.* Association between anti-thymocyte globulin exposure and survival outcomes in adult unrelated haemopoietic cell transplantation: a multicentre, retrospective, pharmacodynamic cohort analysis. *Lancet Haematol* **4,** e183–e191 (2017).

11. Bartelink, I. H., Lalmohamed, A., van Reij, E. M., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *Lancet Haematol* **3,** e526–e536 (2016).

12. Bartelink, I. H., Boelens, J. J., Bredius, R. G., Egberts, A. C., Wang, C., Bierings, M. B., *et al.* Body weight-dependent pharmacokinetics of busulfan in paediatric haematopoietic stem cell transplantation patients: towards individualized dosing. *Clin Pharmacokinet* **51,** 331–45 (2012).

Donor selection & Immune reconstitution: Monitor and regulate immunogenicity after transplantation

2 The influence of stem cell source on transplant outcomes for pediatric patients with acute myeloid leukemia

Jurgen B Langenhorst*, Amy K Keating*, John E Wagner, Kristin M Page, Paul Veys, Robert F Wynn, Heather Stefanski, Reem Elfeky, Roger Giller, Richard Mitchell, Filippo Milano, Tracey A O'Brien, Ann Dahlberg, Colleen Delaney, Joanne Kurtzberg, Michael R Verneris* and Jaap Jan Boelens*

*contributed equally

Blood Advances. Accepted (February 2019)

2.1 Abstract

When allogeneic hematopoietic cell transplantation (HCT) is necessary for children with acute myeloid leukemia (AML) there remains debate about the best stem cell source. Post-HCT relapse is a common cause of mortality and complications such as chronic graft-versus-host disease (cGvHD) are debilitating and also life-threatening. To compare post-HCT outcomes of different donor sources, we retrospectively analyzed consecutive transplants performed in several international centers from 2005-2015. A total of 317 patients were studied, 19% matched sibling donor (MSD), 23% matched unrelated donor (MUD), 39% cord blood (CB) and 19% double cord blood (dCB) recipients. The median age at transplant was 10 years (range 0.42-21) and median follow-up was 4.74 years (range 4.02-5.39). Comparisons were made while controlling for patient, transplant, and disease characteristics. There were no differences in relapse, leukemia-free survival, or non-relapse mortality. dCB recipients had inferior survival compared to MSD, but all other comparisons showed similar overall survival. Despite the majority of CB transplants being human leukocyte antigen (HLA) mismatched, the rates of cGvHD were low, especially compared to the well matched MUD recipients (HR= 0.3; 0.14-0.67; $p = 0.02$). The composite measure of cGvHD and leukemia free survival cGvHD-leukemia free survival (LFS)), which represents both the quality of life and risk of mortality, was significantly better in the CB compared to the MUD recipients (HR=0.56; 0.34-1; $p = 0.03$). In summary, the use of CB is an excellent donor choice for pediatric AML patients when a matched sibling cannot be identified.

2.2 Introduction

Matched sibling donor (MSD) allogeneic hematopoietic cell transplantation (HCT) remains the standard approach for high risk or relapsed acute myeloid leukemia (AML). In approximately 70% of cases a matched sibling is not available[1] and there is much debate as to the best alternative donor source. How alternative donor sources compare in current treatment eras has not been studied in the pediatric population. In many centers, matched unrelated donor (MUD) are the alternative donor of choice, however MUD transplantation requires the identification of a donor and planning of collection. As well, MUD transplantation requires a high degree of human leukocyte antigen (HLA) matching, limiting donor options. As well, MUD recipients have high rates of chronic graft-versus-host disease (cGvHD), which can have lifelong and debilitating consequences on pediatric patients[2–5]. However, these risks are potentially counterbalanced by rapid donor engraftment and acceptable rates of relapse. Over the past 25 years, cord blood (CB) has proven to be an acceptable alternative stem cell donor source and has been increasingly used for HCT of leukemia patients[1,6–10]. CB as an alternative donor source has many logistical advantages. First, the cells are already procured, infectious disease tested, and HLA typed so they are typically quick to obtain. In CB transplantation there is a greater allowance for HLA disparity between donor and recipient, substantially increasing the donor pool for harder to match patients[9,11]. Moreover, CB transplantation has published historical rates of cGvHD that are lower than MUD[12,13]. However, CB transplantation has historically been associated with delayed neutrophil and platelet recovery, as well as higher rates of infectious complications and treatment related mortality, when compared to MUD recipients[14,15]. These historical disadvantages have been overcome with the recognition of the importance of HLA matching at 8 loci, improving the transplanted cell dose and advances in supportive care. Whether outcomes differ between these two alternative stem cell sources in a more contemporary time frame is unknown for pediatric patients needing transplantation[5]. In adult leukemia patients, the composite outcome of relapse-free survival and cGvHD has been established as an important endpoint[16,17]. The integration of cGvHD into a combined outcome is particularly important in the pediatric population, where cGvHD and its therapies can affect individuals potentially for 60-70 years after treatment. In large adult and pediatric studies evaluating risk, cGvHD is specifically associated with poor prognosis and worse long-term survival[18,19]. Here we present a large multi-center retrospective study of 316 pediatric AML patients evaluating post-HCT outcomes based

on donor source with a particular emphasis on the impact of stem cell source on the composite outcome that includes both leukemia free survival (LFS) and cGvHD (cGvHD-LFS).

2.3 Methods

Study population and stem cell source

Data on patients aged 0 to 21 years old, with AML undergoing allogeneic HCT in a complete remission, was collected retrospectively from eight international institutions with accredited pediatric bone marrow transplant programs. Data collected included patients that were consecutively transplanted between the years 2005 and 2015 following a myeloablative conditioning regimen, which contained either total body irradiation (TBI) (>7Gy single dose, 1200 or 1320cGy fractionated), busulfan (>9mg/kg), or treosulfan ($> 10 g/m^2$ (depending on age)). All patients received a stem cell source that was chosen to be the best available at that time by the treating physician and institution, and MSD, MUD, CB, or double cord blood (dCB). All stem cell sources were T-replete and otherwise unmanipulated, no ex vivo expanded stem cell sources were included in this analysis. High risk was defined as a leukemic clone with monosomy 7, monosomy 5, deletion of 5q, high FLT3 allelic ratio of the internal tandem duplication, or persistent leukemia after chemotherapy. All patients received GvHD prophylaxis per the institutional standard.
MSD and MUD donors were evaluated for match with the recipient at the allele level for HLA-A, HLA-B, HLA-C, and HLA-DRB1, with fully matched being defined as 8 of 8. CB and dCB products were evaluated for match with the recipient at the antigen level for HLA-A, HLA-B, and at the allele level for HLA-DRB1, with fully matched being defined as 6 of 6[20,21]. For dCB recipients matching was determined by the matching between the engrafting unit and the recipient as previously described[22].

Outcomes and Statistical Analysis

The main outcome of interest was cGvHD-LFS. Events considered for this composite end-point were moderate and/or severe cGvHD as graded at the time of cGvHD diagnosis, graft failure, leukemia relapse, and non-relapse mortality (NRM). Other outcomes of interest were overall survival (OS), LFS, leukemia relapse, NRM, grades II-IV acute graft-versus-host disease (aGvHD), and cGvHD. Relapse was defined as disease recurrence, NRM as death while in complete remission. Both graft rejection and non-engraftment

were considered graft failure, where in case of non-engraftment, the time was set at day +50 day or time of follow-up/death, whichever occurred first. aGvHD and cGvHD were classified according to the Glucksberg[23] and Shulman[24] criteria.

Duration of follow-up was defined as the time from HCT to last contact or death. Patients were censored at the date of last contact. Median time to follow-up was calculated using the reverse Kaplan-Meier method[25]. Factors, other than stem cell source, considered to influence outcome included patient variables (age at transplantation), treatment variables (conditioning backbone: TBI or chemotherapy, conditioning regimen: BuCy-like, BuFlu, BuFluClo, BuMel-like, TBICy-based, other, GvHD prophylaxis: CSA/MMF, CSA/Mtx, CSA/steroids), donor variables (HLA-disparity) and disease variables (baseline risk: High/Standard, history of central nervous system disease (CNS) leukemia: yes/no, complete remission (CR) status: CR1/CR2). Baseline characteristics of patients for different stem cell sources were compared using the Kruskall-Wallis test for categorical covariates and one-way analysis of variance for continuous variables. Unadjusted probability of cGvHD-LFS, LFS, and OS were computed with the use of the Kaplan–Meier method and p-values were calculated using a 2-sided log-rank test. Unadjusted probability of events subject to competing risk was estimated using cumulative incidence curves and p-values were calculated using Gray's test. Adjusted estimates for cGvHD-LFS, LFS, and OS were computed using Cox-regression models. The adjusted incidence of events subject to competing risk was calculated using Fine-Gray models. Covariates that were significant in the univariate setting, were included in the multivariate models. Model-adjusted estimates correspond to the probability given an equal distribution of model-included-covariates in all groups. P-values for categorical covariates in the regression models were calculated using Wald's test and for continuous covariates using the likelihood-ratio test. P-values for cell source comparisons in the secondary outcomes are adjusted for multiple testing using Bonferroni's method. Furthermore, a sensitivity analysis was performed for the primary outcome (cGvHD-LFS) to assess center effect. Here, the cox regression model was refitted, excluding data from each center separately, and hazard-ratio (HR) was recalculated. Statistical analyses were performed using R version 3.2.4 with packages, cmprsk, survival and rms. De-identified data can be obtained by contacting the corresponding author.

2.4 Results

Patient, Disease, and Transplant Characteristics

A total of 317 patients were included in the study (table 1), and of this group 61 (19%) patients had an MSD available. Amongst the alternative donor sources, 73 (23%) patients received a MUD donor, 122 (39%) patients received a single CB unit, and 61 (19%) received a dCB.Within the MSD group 58 patients received bone marrow stem cells and 3 received peripheral blood stem cells, while in the MUD group 51 patients received bone marrow and 22 received peripheral blood stem cells. The median follow-up at the time of the study was 1730 days, and was similar across donor types (MSD, MUD, CB, dCB). The median age of patients studied was 10 years and ranged from 0.42 to 21 years; age was comparable amongst the groups, except for the dCB, which had a slightly increased age. All patients were in a morphological remission at the time of transplant, although it did vary whether they were in their first or second remission based on their donor source, with 69% of MSD recipients being in a CR1 at time of transplant, compared to 45%, 45%, and 41% of the MUD, CB, and dCB recipients respectively ($p = 0.006$). At diagnosis, 75% of patients in the study cohort were considered high risk. While the alternative donor (MUD, CB, and dCB) recipients were more likely to be high risk (range of 75% to 84%), only 51% of the MSD recipients were high risk ($p < 0.001$). Few patients had a history of CNS leukemia, and this was equally distributed amongst the donor types. Minimal residual disease (MRD) testing was not routinely collected at most centers during the era of this investigation, however on those it was reported (51%), most patients were MRD negative at the time of transplant; MRD negative status was equally distributed amongst the various groups. As expected, degree of HLA match differed between the donor sources ($p = 0.001$). All MSD recipients were fully HLA matched (8/8) to their donors and 92% of the MUD donors were matched at 8/8 at HLA-A, -B, -C and DRB1, with the remainder mismatched at a single locus (7/8 HLA match). In the single CB recipients only 29% were fully HLA matched (6/6 at HLA-A, -B and DRB1) while 28% and 43% were HLA 5/6 and 4/6 loci matched. In the dCB recipients only 18% were fully HLA matched (6/6),while 33% and 49% were matched at 5/6 and 4/6 HLA loci. As shown in table 2, myeloablative regimens also differed amongst the groups ($p < 0.001$); chemotherapy-only based regimens were used in 75% of the MSD, 62% of the MUD, and 57% of the CB transplants. Only 31% of the dCB received only chemotherapy, with the majority (69%) receiving

TBI-based preparation. Nearly all of the MSD transplant preparative regimens (97%) contained no serotherapy, while in contrast, 61% of the MUD recipients, 40% of the single CB recipients, and 16% of the dCB received serotherapy of some type ($p < 0.001$). For those who received serotherapy, the type differed based on stem cell source. Most of the MUD recipients received alemtuzumab and cord blood (CB or dCB) recipients were treated with rabbit anti-thymocyte globulin (rATG). Finally, GvHD prophylaxis also varied by HCT source. The predominant regimen used in both MSD and MUD recipients was a calcineurin inhibitor and methotrexate, whereas CB recipients most commonly received a calcineurin inhibitor and mycophenolate mofetil ($p < 0.001$).

Engraftment

The median time to neutrophil engraftment was 17 days for MSD recipients and 19 days for both MUD ($p = 0.31$) and CB ($p = 0.03$), while the dCB recipients engrafted at 22 days post HCT ($p < 0.001$) (figure 1A). The median time to platelet engraftment was 28 days for MSD and MUD, with 43 and 45 days for CB and dCB respectively (figure 1B). At 60 days, the neutrophil engraftment was similar among all cell sources ($\sim$95%). At 180 days post-HCT, 95% of MSD recipients had engrafted platelets compared to 85% for all other cell sources.

Relapse and Survival

The likelihood of leukemic relapse in the entire group was 22% (17-27%) and the adjusted incidence of relapse was not different amongst the various stem cell sources (figure 2). For all patients, disease status prior to transplant (CR1 vs. CR2) did not affect the cumulative incidence of relapse following HCT (data not shown). The adjusted LFS was 57% (50-66%) for all patients and was similar among stem cell sources (figure 3A). Adjusted NRM at 2 years was 16% (11-20%) and again there were no differences amongst the stem cell sources (figure 3B). The overall survival of the entire group was 63% (57-70%). MSD recipients fared equally well when compared to MUD and CB, but had superior OS compared to dCB ($p = 0.02$, figure 3C). A history of CNS leukemia did not impact relapse or survival (data not shown). The most common causes of death amongst all donor sources was leukemia relapse (57%) followed by GvHD (13%) and infection (11%); less common causes included non-infectious pulmonary disease (8.2%) and multi-organ failure (3.6%).

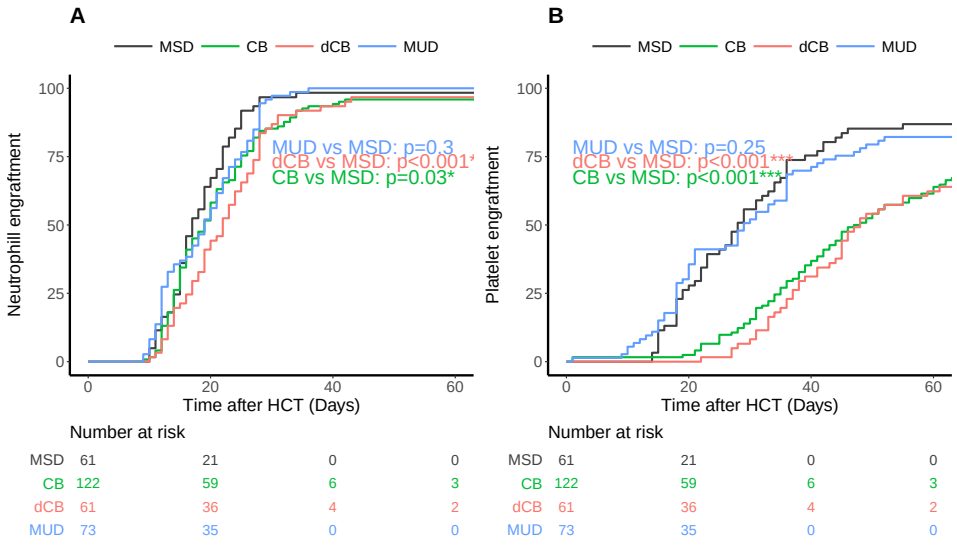

Figure 1: Hematopoietic engraftment. Probability and timeline of (A) neutrophil and (B) platelet recovery by stem cell source.

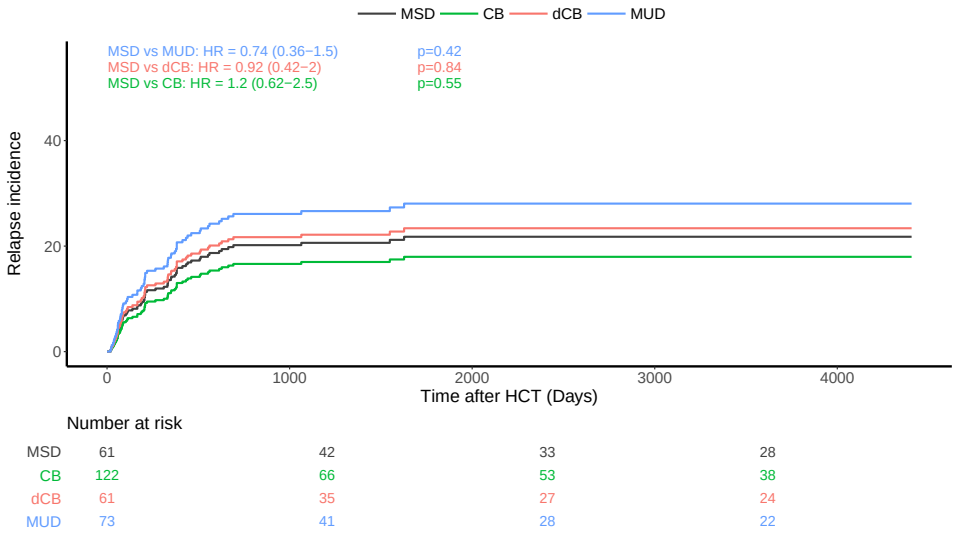

Figure 2: Kaplan-Meier estimates of post-transplant relapse. Incidence of relapse by (A) stem cell source overall, (B) or based on number of remissions.

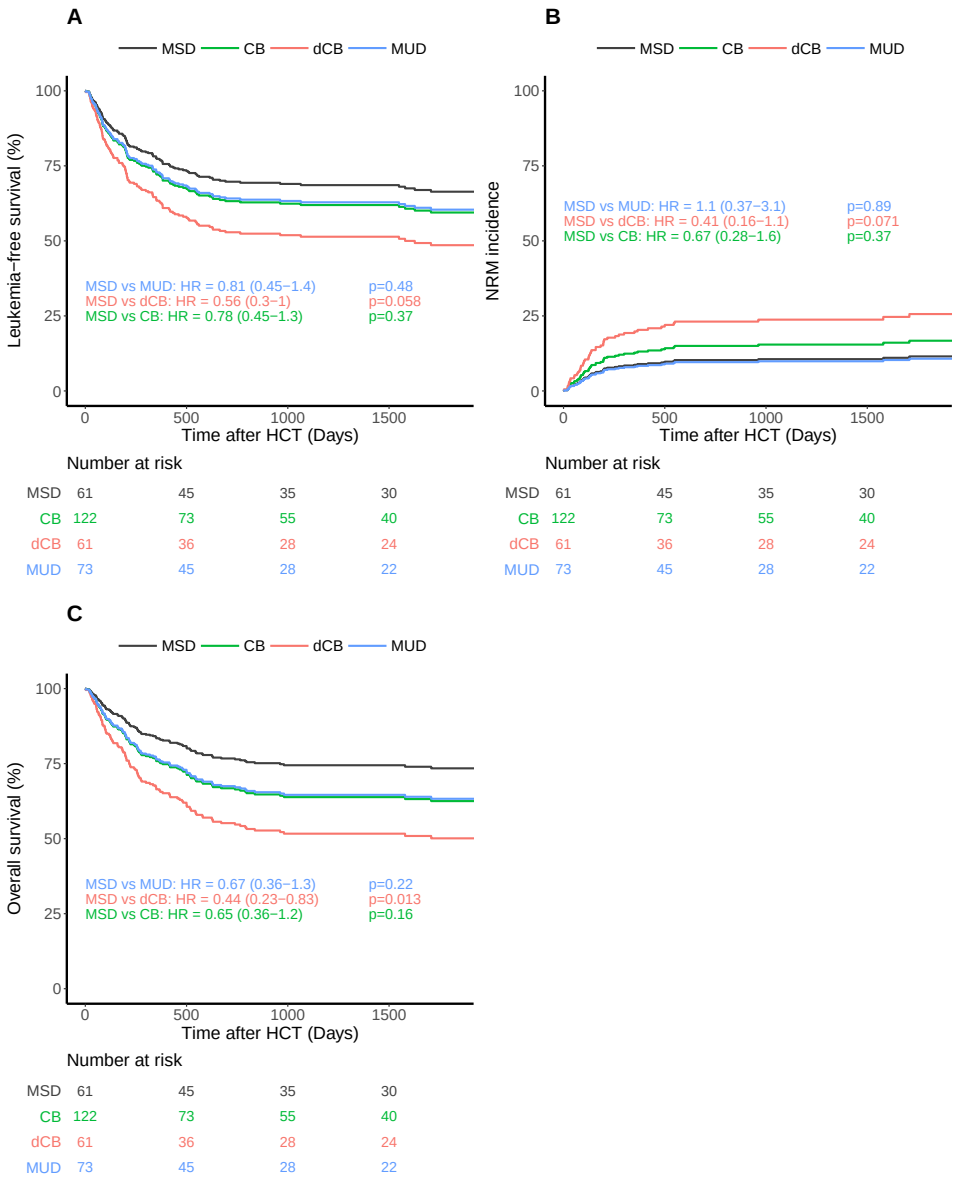

Figure 3: Kaplan-Meier estimates of post-transplant outcomes by stem cell source. (A) leukemia-free survival, (B) non-relapse mortality, and (C) overall survival compared by stem cell source.

Table 1: Characteristics of Patient by Donor Type

	All	MSD	MUD	CB	dCB	p-value[1]
Patients by stem cell source N (% of Total)	317 (100%)	61 (19%)	73 (23%)	122 (38%)	61 (19%)	
Median follow-up	1730	1969	1468	1631	1877	
Median age at HSCT (range)	10 (0.42-21)	8.7 (0.63-21)	11 (1-20)	7.9 (0.42-20)	13 (1.1-20)	p<0.001***
Remission status at HSCT						p=0.0063**
CR1	155 (49%)	42 (69%)	33 (45%)	55 (45%)	25 (41%)	
CR2	162 (51%)	19 (31%)	40 (55%)	67 (55%)	36 (59%)	
Baseline risk stratification						p<0.001***
High	237 (75%)	31 (51%)	55 (75%)	102 (84%)	49 (80%)	
CNS status at HSCT						p=0.6
CNS-disease	86 (27%)	20 (33%)	16 (22%)	33 (27%)	17 (28%)	
Centre of HSCT						p<0.001***
Australia	24 (7.6%)	8 (13%)	1 (1.4%)	11 (9%)	4 (6.6%)	
Denver	48 (15%)	14 (23%)	11 (15%)	18 (15%)	5 (8.2%)	
Duke	58 (18%)	16 (26%)	3 (4.1%)	30 (25%)	9 (15%)	
Gosh	16 (5%)	4 (6.6%)	6 (8.2%)	6 (4.9%)	0 (0%)	
Manchester	39 (12%)	6 (9.8%)	23 (32%)	7 (5.7%)	3 (4.9%)	
Minneapolis	52 (16%)	0 (0%)	10 (14%)	13 (11%)	29 (48%)	
Seattle	37 (12%)	6 (9.8%)	14 (19%)	8 (6.6%)	9 (15%)	
Utrecht	43 (14%)	7 (11%)	5 (6.8%)	29 (24%)	2 (3.3%)	
Primary cause of death						p=0.14
GvHD	14 (13%)	1 (6.2%)	1 (4.2%)	7 (17%)	5 (18%)	
Infection	12 (11%)	2 (12%)	0 (0%)	5 (12%)	5 (18%)	
MOF	4 (3.6%)	1 (6.2%)	2 (8.3%)	1 (2.4%)	0 (0%)	
Non-infecteous lung	9 (8.2%)	1 (6.2%)	2 (8.3%)	3 (7.1%)	3 (11%)	
Other	8 (7.3%)	2 (12%)	2 (8.3%)	3 (7.1%)	1 (3.6%)	
Relapse	63 (57%)	9 (56%)	17 (71%)	23 (55%)	14 (50%)	

[1] p-values were calculated using one-way analysis of variance for continuous variables and Kruskall-Wallis for categorical variables. MSD, matched sibling donor; MUD, matched unrelated donor; CB, umbilical cord blood; dCB, double umbilical cord blood; NA, not applicable; CR, complete remission; CNS, central nervous system;

Graft-versus-Host Disease

The unadjusted incidence of aGvHD at 180 days was lowest in recipients of MSD grafts; 24% (13-36%) and increased to 43% (31-55%, $p = 0.07$) for MUD, 52% for CB (43-62%, $p < 001$), and 56% for dCB (43-69%, $p < 0.001$). However, the increased aGvHD incidence in MUD and dCB was found to be mainly attributable to different GvHD prophylaxis and more TBI-based

Table 2: Characteristics of Transplant by Donor Type

	All	MSD	MUD	CB	dCB	p-value[1]
CD34+ cells ($*10^6$)	1 (0.01-620)	9.6 (0.24-620)	6.4 (0.66-380)	0.21 (0.02-3)	0.43 (0.01-11)	p<0.001***
Nucleated cells ($*10^8$)	1.7 (0.1-460)	4.9 (0.16-320)	5 (0.92-460)	0.62 (0.18-30)	0.47 (0.1-32)	p<0.001***
HLA-matching						p<0.001***
1 mismatch	89 (28%)	0 (0%)	6 (8.2%)	53 (43%)	30 (49%)	
>1 mismatch	54 (17%)	0 (0%)	0 (0%)	34 (28%)	20 (33%)	
full match	174 (55%)	61 (100%)	67 (92%)	35 (29%)	11 (18%)	
Myeloablation						p<0.001***
Chemo-based*	181 (57%)	46 (75%)	46 (63%)	70 (57%)	19 (31%)	
TBI-based	136 (43%)	15 (25%)	27 (37%)	52 (43%)	42 (69%)	
Serotherapy						p<0.001***
No Serotherapy	212 (67%)	59 (97%)	28 (38%)	74 (61%)	51 (84%)	
ATG	67 (21%)	1 (1.6%)	9 (12%)	47 (39%)	10 (16%)	
Campath	38 (12%)	1 (1.6%)	36 (49%)	1 (0.82%)	0 (0%)	
GvHD prophylaxis						p<0.001***
CSA/MMF	178 (56%)	21 (34%)	26 (36%)	75 (61%)	56 (92%)	
CSA/Mtx	87 (27%)	40 (66%)	47 (64%)	0 (0%)	0 (0%)	
CSA/steroids	52 (16%)	0 (0%)	0 (0%)	47 (39%)	5 (8.2%)	

* Busulfan/Cyclophosphamide, Busulfan/Fludarabine, Busulfan/Fludarabine/Clofarabine, Busulfan/Melphalan

conditioning in multivariate analyses. The higher aGvHD rates in (d)CB did not translate into higher rates of cGvHD. Conversely, the adjusted rates of cGvHD after single CB (21%) and dCB (22%) were lower than MUD recipients (48%, figure 4B).

Composite Measurement: cGvHD-LFS

When considering the principal composite endpoint which incorporates adjusted LFS (LFS here including graft failure, leukemia relapse and NRM) and the incidence of moderate and/or severe cGvHD, we found a mean adjusted probability of 37% (27-51%). Patients who received a standard MSD HCT had an adjusted cGvHD-free LFS at 5 years of 44%, while CB (49%) and dCB (44%) achieved very similar rates. However, MUD recipients had a cGvHD-free LFS of only 29%. Using multiple variate analysis, recipients of MUD transplants had a significantly lower probability of cGvHD-LFS as compared to single CB recipients (CB vs MUD HR=0.56 (0.34-1.0, $p = 0.03$), figure 5, table 3). This effect was not transplant center specific as the individual center sensitivity analyses had similar results with the HR

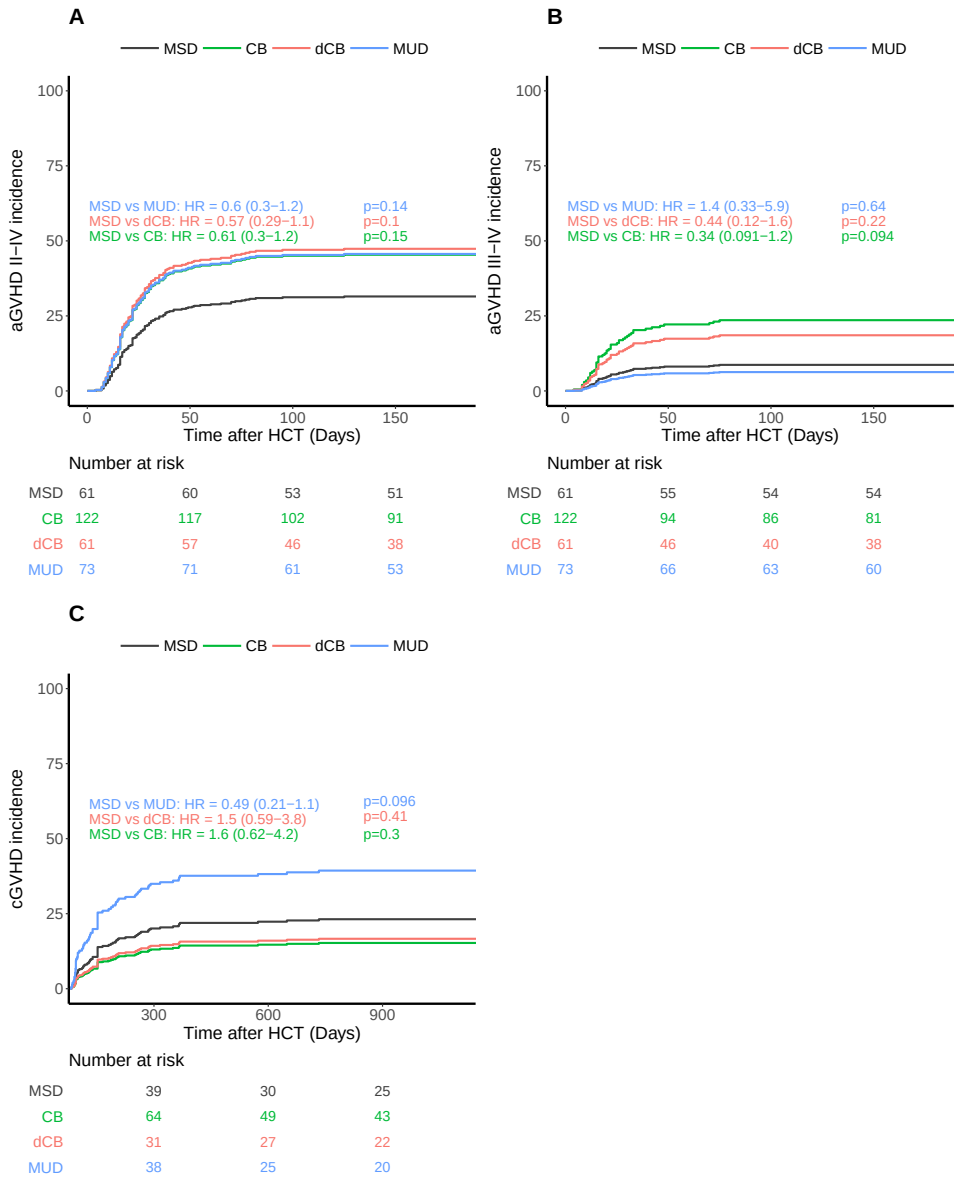

Figure 4: Incidence of Graft versus Host Disease (GvHD). (A) The incidence of grades II-IV (A) and III-IV (B) acute GvHD and (C) chronic GvHD compared by stem cell source.

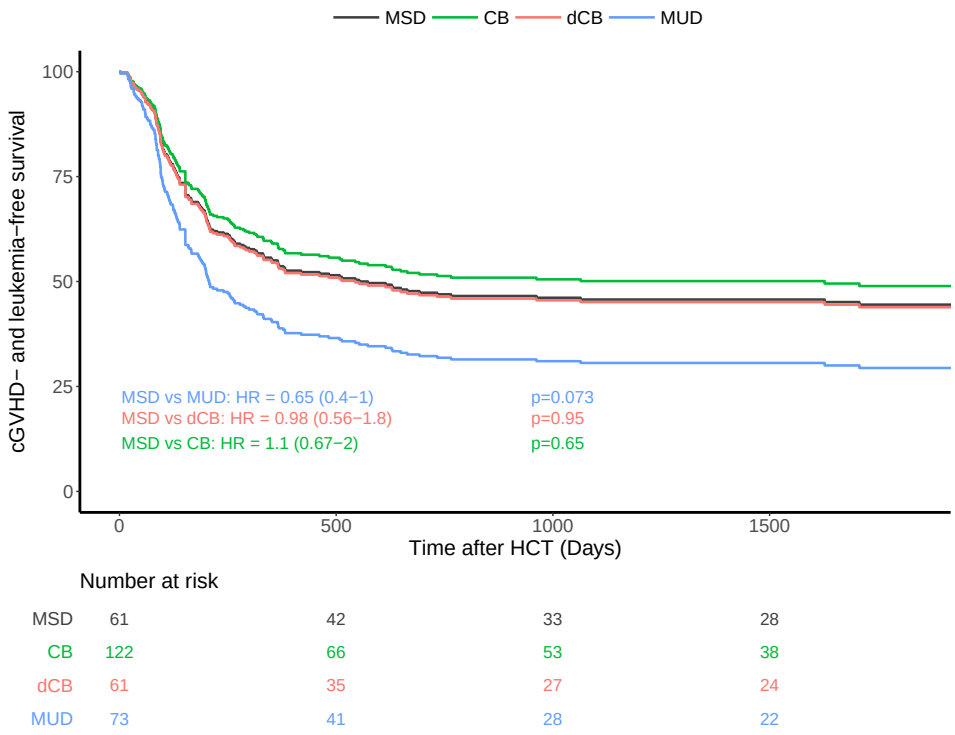

Figure 5: Composite Endpoint cGvHD-LFS. The composite endpoint of chronic graft versus host disease and leukemia-free survival (cGvHD-LFS).

ranging from 0.56 to 0.71.

2.5 Discussion

The goal of hematopoietic cell transplantation for children with AML is to use the immunological graft versus leukemia reactions to eradicate the disease, providing long term remission and cure. However, this frequently comes with collateral damage in the form of acute and chronic GvHD. The impact of cGvHD is especially detrimental to young children who are frequently still developing. Because of this and prior data suggesting that different stem cell sources might be associated with varying rates of cGvHD[5,26,27], we compared outcomes of patients receiving different stem cell sources for the composite endpoint of cGvHD-LFS, which is arguably the best endpoint to infer satisfactory outcomes, including quality of life. In this unique and large multi-institutional comparison we found that MSD and CB recip-

ients experienced similar outcomes for the primary composite endpoint of cGvHD-LFS, while MUD recipients had a significantly lower cGvHD-LFS. No differences in OS, LFS and relapse rates were noted.

A previous large retrospective CIBMTR analysis of 1525 adult patients with acute leukemia noted equivalent LFS in CB recipients as compared to MUD recipients[12]. Similarly, a multicenter, retrospective analyses of pediatric leukemia patients receiving either a well matched MUD (n=262) or HLA-disparate CB (n=99), the LFS and OS was similar, however CB recipients had delayed neutrophil engraftment and a greater incidence of day 100 NRM[14]. Also, in a larger pediatric retrospective study comparing MUD and CB recipient outcomes (n=785) from 1995-2003, LFS was similar amongst the groups, but the CB recipients had longer median times to neutrophil and platelet engraftment and higher rates of NRM, but lower rates of relapse[15]. Our multi-center retrospective analysis also showed no difference in relapse rates amongst the various donor sources. This may reflect transplant in an era where donor source has less of an impact due to either changing indications for allogeneic transplant in AML[28], or more recently because MRD-based timing of transplantation likely impacts relapse rates. The characteristics of the patients and their grafts in this present analysis were similar to the previous study but reflect contemporary practices, with higher average cell doses and more robust supportive care. This has likely resulted in comparable NRM, relapse rates, and hematopoietic (neutrophil and platelet) recovery among the groups. During the era being investigated many clinical trials assigned patients to undergo HCT in first remission when a MSD was available. Our results reflect this trend with the majority of MSD recipients transplanted while in CR1 (68%) compared to only 41-45% of the alternative donor recipients in CR1 at time of transplant. Additionally, the presence of high risk AML features was often a criterion used to proceed into transplant. Not surprisingly, alternative donor recipients more likely to be high risk in our study cohort as compared MSD recipients. Despite this, we found that neither CR status (CR1 vs. CR2) nor high risk leukemia features impacted outcomes.

GvHD can be a severe and life-threatening complication of allogeneic HCT, although it is counterbalanced by the potential for beneficial graft-versus-leukemia reactions. Similar to prior studies[8,13,14,29,30], we found in multivariate analysis similar rates of aGvHD II-IV in CB and MUD recipients that was comparatively higher than MSDs. Likewise, the incidence of cGvHD also did not vary between the groups, although MUD recipients tended towards increased rates of cGvHD as has been previously reported by others[12,29]. Importantly, survival alone does not completely describe a

successful post-HCT outcome. cGvHD is a debilitating and life-threatening transplant complication that significantly impacts the quality and quantity of life. In a large study of 1800 adult HCT survivors, ongoing cGvHD was associated with increased medication needs, decreased employment, and overall decreased resiliency and quality of life[31]. Similarly, in a large outcomes study of nearly 2000 pediatric patients, children transplanted for leukemia had decreased survival if they suffered from cGvHD after HCT. For pediatric AML, the HR for mortality was 1.7 to 2.34 higher for recipients who experienced any cGvHD[18].

Because individual survival and GvHD analyses do not adequately describe the long-term wellbeing of the recipient after HCT, Holtan and coworkers developed the novel composite outcome GvHD-free/relapse-free survival (GRFS) and presented it as 'ideal HCT recovery' that could be used to compare therapies. In their analysis of over 900 patients, the GRFS at 1 year did not demonstrate any difference between CB and MUD recipients (31% and 32%, respectively)[16]. Ruggeri et al more recently published a very large retrospective analysis of adult AML patients (n=20,937). The GRFS at 3 years, which better represents long-term well-being, was ~40% for all sources[17]. In our study we applied this combined outcome analysis to the pediatric AML population, including moderate or extensive cGvHD and excluding the incidence of aGvHD in the combinatorial analysis. We surmised if there is no conversion into cGvHD, the inclusion of aGvHD was unnecessary in the analysis as this does not impact long-term morbidity and mortality.

Retrospective analysis of institutional data from a group of international centers comes with the inherent limitations. Specifically, the potential for bias based on center practices such as which patients receive transplantation, choice of stem cells, preparative regimen, GvHD prophylaxis, treatment and reporting of outcomes. Furthermore, this study does not draw any comparisons to the more contemporary approach of haploidentical transplant (alpha/beta depletion or post-HCT cyclophosphamide) which are becoming more common in the pediatric hematologic malignancy setting[32,33]. The numbers of haploidentical transplants in this timeframe were unfortunately too small, and the manner by which these transplants were done is to disparate from the current approach to allow for inclusion. Ideally, this could be further studied in prospective trials to better understand how haploidentical recipient outcomes compare to CB and MUDs.

In this large multicenter analysis, we were able to demonstrate for the first time in a pediatric cohort that CB recipients experienced improved cGvHD-LFS as compared to MUD recipients. This important finding is further am-

Table 3: Multivariate analysis

		HR (95% CI)	p-value	p-value (unadjusted)
Relapse incidence	**CB vs MUD**	**0.59 (0.31-1.1)**	**0.84**	**0.12**
	MSD vs CB	**1.2 (0.59-2.5)**	**1**	**0.59**
	MSD vs MUD	**0.71 (0.34-1.5)**	**1**	**0.4**
	MSD vs dCB	**0.91 (0.4-2)**	**1**	**0.8**
	Chemo-based vs TBI-based	0.77 (0.45-1.3)	0.36	0.36
	Age at HCT (spline 1)	0.86 (0.76-0.97)	0.015	0.015
	Age at HCT (spline 2)	1.2 (1-1.4)	0.03	0.03
Leukemia free survival	**CB vs MUD**	**1 (0.62-1.7)**	**1**	**0.9**
	MSD vs CB	**0.77 (0.45-1.3)**	**1**	**0.37**
	MSD vs MUD	**0.83 (0.45-1.4)**	**1**	**0.48**
	MSD vs dCB	**0.56 (0.3-1)**	**0.406**	**0.058**
	Age at HCT (spline 1)	0.88 (0.8-0.96)	0.0054	0.0054
	Age at HCT (spline 2)	1.1 (1-1.3)	0.012	0.012
	BuCy-like vs BuFlu	0.43 (0.17-1.1)	0.077	0.077
	BuCy-like vs BuMel-like	0.91 (0.5-1.7)	0.77	0.77
	BuCy-like vs Other	0.77 (0.4-1.4)	0.39	0.39
	BuCy-like vs TBICy-based*	1 (0.56-1.8)	0.99	0.99
Non-relapse mortality	**CB vs MUD**	**1.6 (0.67-3.8)**	**1**	**0.28**
	MSD vs CB	**0.67 (0.28-1.6)**	**1**	**0.37**
	MSD vs MUD	**1.1 (0.37-3.1)**	**1**	**0.89**
	MSD vs dCB	**0.42 (0.16-1.1)**	**0.497**	**0.071**
	Chemo-based vs TBI-based	1.2 (0.62-2.2)	0.59	0.59
	Age at HCT (spline 1)	0.94 (0.81-1.1)	0.37	0.37
	Age at HCT (spline 2)	1.1 (0.9-1.3)	0.49	0.49
Overall survival	**CB vs MUD**	**1 (0.59-1.8)**	**1**	**0.92**
	MSD vs CB	**0.67 (0.36-1.2)**	**1**	**0.16**
	MSD vs MUD	**0.67 (0.36-1.3)**	**1**	**0.22**
	MSD vs dCB	**0.43 (0.23-0.83)**	**0.091**	**0.013**
	Age at HCT (spline 1)	0.91 (0.83-1)	0.053	0.053
	Age at HCT (spline 2)	1.1 (1-1.2)	0.061	0.061
	BuCy-like vs BuFlu	0.34 (0.13-0.91)	0.029	0.029
	BuCy-like vs BuMel-like	0.77 (0.38-1.4)	0.38	0.38
	BuCy-like vs Other	0.59 (0.3-1.2)	0.13	0.13
	BuCy-like vs TBICy-based*	0.91 (0.5-1.8)	0.82	0.82
Incidence of aGVHD grades II-IV	**CB vs MUD**	**1 (0.56-1.8)**	**1**	**0.98**
	MSD vs CB	**0.62 (0.3-1.2)**	**1**	**0.15**
	MSD vs MUD	**0.59 (0.3-1.2)**	**0.98**	**0.14**
	MSD vs dCB	**0.56 (0.29-1.1)**	**0.7**	**0.1**
	Chemo-based vs TBI-based	0.56 (0.36-0.83)	0.0046	0.0046
	Age at HCT (spline 1)	0.97 (0.89-1.1)	0.53	0.53

	Age at HCT (spline 2)	1 (0.9-1.1)	0.93	0.93
	CSA/MMF vs CSA/Mtx	2 (1.1-3.7)	0.017	0.017
	CSA/MMF vs CSA/steroids	1.4 (0.77-2.4)	0.31	0.31
	No Serotherapy vs ATG	1.3 (0.83-2)	0.3	0.3
	No Serotherapy vs Campath	0.91 (0.43-1.8)	0.73	0.73
Incidence of aGVHD grades III-IV	**CB vs MUD**	**4.2 (1-18)**	**0.357**	**0.051**
	MSD vs CB	**0.33 (0.091-1.2)**	**0.658**	**0.094**
	MSD vs MUD	**1.4 (0.33-5.9)**	**1**	**0.64**
	MSD vs dCB	**0.43 (0.12-1.6)**	**1**	**0.22**
	Chemo-based vs TBI-based	1 (0.45-2)	0.9	0.9
	Age at HCT (spline 1)	0.91 (0.79-1.1)	0.21	0.21
	Age at HCT (spline 2)	1.1 (0.93-1.3)	0.27	0.27
	CSA/MMF vs CSA/Mtx	1.1 (0.31-3.8)	0.88	0.88
	CSA/MMF vs CSA/steroids	2 (0.83-5)	0.13	0.13
	No Serotherapy vs ATG	1.1 (0.53-2.3)	0.8	0.8
	No Serotherapy vs Campath	0.91 (0.19-4.2)	0.88	0.88
Incidence of cGVHD	**CB vs MUD**	**0.3 (0.14-0.67)**	**0.0217**	**0.0031**
	MSD vs CB	**1.6 (0.62-4.2)**	**1**	**0.3**
	MSD vs MUD	**0.5 (0.21-1.1)**	**0.672**	**0.096**
	MSD vs dCB	**1.5 (0.59-3.8)**	**1**	**0.41**
	Chemo-based vs TBI-based	0.43 (0.24-0.83)	0.0076	0.0076
	Age at HCT (spline 1)	1.1 (0.93-1.2)	0.4	0.4
	Age at HCT (spline 2)	0.9 (0.78-1)	0.17	0.17
	CSA/MMF vs CSA/Mtx	2.1 (1.1-4.3)	0.032	0.032
	CSA/MMF vs CSA/steroids	3.6 (0.83-16)	0.093	0.093
	No Serotherapy vs ATG	1.1 (0.53-2.5)	0.75	0.75
	No Serotherapy vs Campath	1 (0.42-2.3)	0.94	0.94
cGVHD-LFS	**CB vs MUD**	**0.56 (0.34-1)**	**0.033**	**0.033**
	MSD vs CB	**1.1 (0.67-2)**	**0.65**	**0.65**
	MSD vs MUD	**0.67 (0.4-1)**	**0.073**	**0.073**
	MSD vs dCB	**1 (0.56-1.8)**	**0.95**	**0.95**
	Chemo-based vs TBI-based	0.83 (0.59-1.2)	0.37	0.37
	Age at HCT (spline 1)	0.95 (0.88-1)	0.16	0.16
	Age at HCT (spline 2)	1 (0.95-1.1)	0.41	0.41
	CSA/MMF vs CSA/Mtx	1.7 (1.1-2.8)	0.025	0.025
	CSA/MMF vs CSA/steroids	1.5 (0.91-2.6)	0.14	0.14

plified given that the majority of pediatric patients require an alternative donor for HCT. While other endpoints (LFS, OS, and engraftment rates) were equivalent among the various cell sources, the cGvHD-LFS endpoint better reflects the optimal post-HCT outcome. Based on these results, CB is an excellent alternative cell source if an MSD is lacking. Prospective validation is required.

References

1. Khandelwal, P., Millard, H. R., Thiel, E., Abdel-Azim, H., Abraham, A. A., Auletta, J. J., *et al.* Hematopoietic Stem Cell Transplantation Activity in Pediatric Cancer between 2008 and 2014 in the United States: A Center for International Blood and Marrow Transplant Research Report. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **23,** 1342–1349 (8 Aug. 2017).

2. Martin, P. J., Inamoto, Y., Carpenter, P. A., Lee, S. J. & Flowers, M. E. D. Treatment of chronic graft-versus-host disease: Past, present and future. *The Korean journal of hematology* **46,** 153–163 (3 Sept. 2011).

3. Vrooman, L. M., Millard, H. R., Brazauskas, R., Majhail, N. S., Battiwalla, M., Flowers, M. E., *et al.* Survival and Late Effects after Allogeneic Hematopoietic Cell Transplantation for Hematologic Malignancy at Less than Three Years of Age. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **23,** 1327–1334 (8 Aug. 2017).

4. Boyiadzis, M., Arora, M., Klein, J. P., Hassebroek, A., Hemmer, M., Urbano-Ispizua, A., *et al.* Impact of Chronic Graft-versus-Host Disease on Late Relapse and Survival on 7,489 Patients after Myeloablative Allogeneic Hematopoietic Cell Transplantation for Leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research* **21,** 2020–2028 (9 May 2015).

5. Jacobsohn, D. A., Arora, M., Klein, J. P., Hassebroek, A., Flowers, M. E., Cutler, C. S., *et al.* Risk factors associated with increased nonrelapse mortality and with poor overall survival in children with chronic graft-versus-host disease. *Blood* **118,** 4472–4479 (16 Oct. 2011).

6. Gluckman, E., Rocha, V., Boyer-Chammard, A., Locatelli, F., Arcese, W., Pasquini, R., *et al.* Outcome of cord-blood transplantation from related and unrelated donors. Eurocord Transplant Group and the European Blood and Marrow Transplantation Group. *The New England journal of medicine* **337,** 373–381 (6 Aug. 1997).

7. Gluckman, E. & Rocha, V. Cord blood transplantation: state of the art. *Haematologica* **94,** 451–454 (4 Apr. 2009).

8. Brunstein, C. G., Baker, K. S. & Wagner, J. E. Umbilical cord blood transplantation for myeloid malignancies. *Current opinion in hematology* **14,** 162–169 (2 Mar. 2007).

9. Ballen, K. K., Gluckman, E. & Broxmeyer, H. E. Umbilical cord blood transplantation: the first 25 years and beyond. *Blood* **122,** 491–498 (4 July 2013).

10. Ballen, K. K. & Barker, J. N. Has umbilical cord blood transplantation for AML become mainstream? *Current opinion in hematology* **20,** 144–149 (2 Mar. 2013).

11. Fuchs, E., O'Donnell, P. V. & Brunstein, C. G. Alternative transplant donor sources: is there any consensus? *Current opinion in oncology* **25,** 173–179 (2 Mar. 2013).

12. Eapen, M., Rocha, V., Sanz, G., Scaradavou, A., Zhang, M.-J., Arcese, W., *et al.* Effect of graft source on unrelated donor haemopoietic stem-cell transplantation in adults with acute leukaemia: a retrospective analysis. *The Lancet. Oncology* **11,** 653–660 (7 July 2010).

13. Weisdorf, D., Eapen, M., Ruggeri, A., Zhang, M.-J., Zhong, X., Brunstein, C., *et al.* Alternative donor transplantation for older patients with acute myeloid leukemia in first complete remission: a center for international blood and marrow transplant research-eurocord analysis. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **20,** 816–822 (6 June 2014).

14. Rocha, V., Cornish, J., Sievers, E. L., Filipovich, A., Locatelli, F., Peters, C., *et al.* Comparison of outcomes of unrelated bone marrow and umbilical cord blood transplants in children with acute leukemia. *Blood* **97,** 2962–2971 (10 May 2001).

15. Eapen, M., Rubinstein, P., Zhang, M.-J., Stevens, C., Kurtzberg, J., Scaradavou, A., *et al.* Outcomes of transplantation of unrelated donor umbilical cord blood and bone marrow in children with acute leukaemia: a comparison study. *Lancet (London, England)* **369,** 1947–1954 (9577 June 2007).

16. Holtan, S. G., DeFor, T. E., Lazaryan, A., Bejanyan, N., Arora, M., Brunstein, C. G., *et al.* Composite end point of graft-versus-host disease-free, relapse-free survival after allogeneic hematopoietic cell transplantation. *Blood* **125,** 1333–1338 (8 Feb. 2015).

17. Ruggeri, A., Labopin, M., Ciceri, F., Mohty, M. & Nagler, A. Definition of GvHD-free, relapse-free survival for registry-based studies: an ALWP-EBMT analysis on patients with AML in remission. *Bone marrow transplantation* **51,** 610–611 (4 Apr. 2016).

18. Bitan, M., Ahn, K. W., Millard, H. R., Pulsipher, M. A., Abdel-Azim, H., Auletta, J. J., *et al.* Personalized Prognostic Risk Score for Long-Term Survival for Children with Acute Leukemia after Allogeneic Transplantation. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **23,** 1523–1530 (9 Sept. 2017).

19. Arora, M., Hemmer, M. T., Ahn, K. W., Klein, J. P., Cutler, C. S., Urbano-Ispizua, A., *et al.* Center for International Blood and Marrow Transplant Research chronic graft-versus-host disease risk score predicts mortality in an independent validation cohort. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **21,** 640–645 (4 Apr. 2015).

20. Eapen, M., Klein, J. P., Ruggeri, A., Spellman, S., Lee, S. J., Anasetti, C., *et al.* Impact of allele-level HLA matching on outcomes after myeloablative single unit umbilical cord blood transplantation for hematologic malignancy. *Blood* **123,** 133–140 (1 Jan. 2014).

21. Eapen, M., Wang, T., Veys, P. A., Boelens, J. J., St Martin, A., Spellman, S., *et al.* Allele-level HLA matching for umbilical cord blood transplantation for non-malignant diseases in children: a retrospective analysis. *The Lancet. Haematology* **4,** e325–e333 (7 July 2017).

22. Brunstein, C. G., Petersdorf, E. W., DeFor, T. E., Noreen, H., Maurer, D., MacMillan, M. L., *et al.* Impact of Allele-Level HLA Mismatch on Outcomes in Recipients of Double Umbilical Cord Blood Transplantation. *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation* **22,** 487–492 (3 Mar. 2016).

23. Glucksberg, H., Storb, R., Fefer, A., Buckner, C. D., Neiman, P. E., Clift, R. A., *et al.* Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation* **18,** 295–304 (4 Oct. 1974).

24. Shulman, H. M., Sullivan, K. M., Weiden, P. L., McDonald, G. B., Striker, G. E., Sale, G. E., *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 (2 Aug. 1980).

25. Altman, D. G., De Stavola, B. L., Love, S. B. & Stepniewska, K. A. Review of survival analyses published in cancer journals. *Br J Cancer* **72,** 511–8 (1995).

26. Alsultan, A., Giller, R. H., Gao, D., Bathurst, J., Hild, E., Gore, L., *et al.* GVHD after unrelated cord blood transplant in children: characteristics, severity, risk factors and influence on outcome. *Bone marrow transplantation* **46,** 668–675 (5 May 2011).

27. Arora, M., Klein, J. P., Weisdorf, D. J., Hassebroek, A., Flowers, M. E. D., Cutler, C. S., *et al.* Chronic GVHD risk score: a Center for International Blood and Marrow Transplant Research analysis. *Blood* **117,** 6714–6720 (24 June 2011).

28. Horan, J. T., Logan, B. R., Agovi-Johnson, M.-A., Lazarus, H. M., Bacigalupo, A. A., Ballen, K. K., *et al.* Reducing the risk for transplantation-related mortality after allogeneic hematopoietic cell transplantation: how much progress has been made? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **29,** 805–813 (7 Mar. 2011).

29. Gutman, J. A., Ross, K., Smith, C., Myint, H., Lee, C.-K., Salit, R., *et al.* Chronic graft versus host disease burden and late transplant complications are lower following adult double cord blood versus matched unrelated donor peripheral blood transplantation. *Bone marrow transplantation* **51,** 1588–1593 (12 Dec. 2016).

30. Wagner, J. E., Eapen, M., Carter, S., Wang, Y., Schultz, K. R., Wall, D. A., *et al.* One-unit versus two-unit cord-blood transplantation for hematologic cancers. *The New England journal of medicine* **371,** 1685–1694 (18 Oct. 2014).

31. Rosenberg, A. R., Syrjala, K. L., Martin, P. J., Flowers, M. E., Carpenter, P. A., Salit, R. B., *et al.* Resilience, health, and quality of life among long-term survivors of hematopoietic cell transplantation. *Cancer* **121,** 4250–4257 (23 Dec. 2015).

32. Symons, H. J. & Fuchs, E. J. Hematopoietic SCT from partially HLA-mismatched (HLA-haploidentical) related donors. *Bone marrow transplantation* **42,** 365–377 (6 Sept. 2008).

33. Bashey, A., Zhang, X., Sizemore, C. A., Manion, K., Brown, S., Holland, H. K., *et al.* T-cell-replete HLA-haploidentical hematopoietic transplantation for hematologic malignancies using post-transplantation cyclophosphamide results in outcomes equivalent to those of contem-

poraneous HLA-matched related and unrelated donor transplantation. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **31,** 1310–1316 (10 Apr. 2013).

3 Innate immune recovery predicts CD4+ T-cell reconstitution after hematopoietic Cell transplantation

Jurgen B Langenhorst*, Coco de Koning*, Charlotte van Kesteren, Caroline A Lindemans, Alwin DR Huitema, Stefan Nierkens, Jaap Jan Boelens

*contributed equally

Biology of blood and marrow transplantation. 25, 819–826 (2019).

3.1 Abstract

Background

Innate immune cells are the first to recover after allogeneic hematopoietic cell transplantation (HCT). Nevertheless, reports of innate immune cell recovery and their relation to adaptive recovery after HCT, are largely lacking. Especially predicting CD4+ T-cell reconstitution is of clinical interest, as this parameter directly associates with survival chances after HCT.

Objectives

We evaluated whether innate recovery relates to CD4+ T-cell reconstitution probability, and investigated differences between innate recovery after cord blood transplantation (CBT) and bone marrow transplantation (BMT).

Methods

We developed a multivariate, combined non-linear mixed-effects model for monocytes, neutrophils and natural killer (NK)-cell recovery after transplantation. 205 Patients, undergoing a first HCT (76 BMT, 129 CBT) between 2007-2016, were included. The median age was 7.3 years (range 0.16-23).

Results

Innate recovery was highly associated with CD4+ T-cell reconstitution probability ($p < 0.001$) in multivariate analysis correcting for covariates. Monocyte ($p < 0.001$), neutrophil ($p < 0.001$), and NK-cell ($p < 0.001$) recovery reached higher levels during the first 200 days after CBT compared to BMT. The higher innate recovery after CBT may be explained by increased proliferation capacity (measured by Ki-67 expression) of innate cells in cord blood (CB)-grafts compared to BM-grafts ($p = 0.041$), and of innate cells in vivo after CBT compared to BMT ($p = 0.048$). At an individual level, patients with increased innate recovery after either CBT or BMT had received grafts with higher proliferating innate cells (CB; $p = 0.004$, BM; $p = 0.01$, respectively).

Conclusion

Our findings implicate the use of early innate immune monitoring to predict the chance of CD4+ T-cell reconstitution after HCT, with respect to higher innate recovery after CBT compared to BMT.

3.2 Introduction

Allogeneic hematopoietic cell transplantation (HCT) is a potentially curative treatment option for pediatric patients suffering from life-threatening refractory malignancies or non-malignant diseases. HCT in children is performed mostly using bone marrow (BM) or cord blood (CB) grafts. After HCT, the immune system needs to rebuild; a process referred to as immune reconstitution. In addition to early donor neutrophil engraftment, robust T-cell reconstitution is essential for disease control and infection control. Poor or delayed T-cell reconstitution relates to increased morbidity and mortality[1–8]. We previously reported that survival is significantly higher in patients with early CD4+ T-cell reconstitution (IR) compared to patients with delayed reconstitution[1,9,10]. It is, therefore, essential to have predictable IR, and thus to identify factors that influence T-cell reconstitution after HCT.

While recovery of T-cells to reference levels may take several months to even years, recovery of innate immune cells, such as neutrophils, monocytes, and natural killer (NK)-cells, only takes weeks to a few months[11]. Neutrophils are the first cells to arise directly from the graft after transplantation, and have, therefore, been adopted as an early measure for successful engraftment[2,11–13]. NK-cells and monocytes also recover within weeks after HCT[11]. As innate recovery generally precedes adaptive recovery[11], it might influence and/or predict IR and thereby (indirectly) affect outcome after HCT. Studies evaluating the dynamics of innate recovery in relation to adaptive immune recovery after HCT are, however, lacking.
The aim of this study was to evaluate innate immune cell recovery in relation to adaptive immune reconstitution after bone marrow transplantation (BMT) and cord blood transplantation (CBT) in a pediatric cohort. We applied an immune-modeling strategy using non-linear mixed effects modeling to evaluate whether innate immune cell recovery relates to IR, and to assess differences in innate immune cell recovery between CBT and BMT.

3.3 Patients and Methods

Study design and patients

In this study, we included pediatric patients receiving an allogeneic HCT between January 2007 and December 2016 at the University Medical Centre in Utrecht, The Netherlands. Only patients with successful engraftment of their first HCT, who did not need a second HCT, were included. Patients receiving serotherapy other than rabbit anti-thymocyte globulin (rATG) (brand: Thymoglobulin) were excluded. There were no limitations

regarding age, indication, or conditioning. Only patients receiving T-cell repleted donors were included. Patients were enrolled, and data were collected and registered prospectively, only after written informed consent had been given in accordance with the Helsinki Declaration. The study was approved by the local ethical committee (trial numbers 05-143 and 11-063k).

Procedures

Patients received either chemotherapy-based regimens (busulfan and fludarabine for benign disorders or clofarabine, fludarabine, and busulfan for malignant indications) or total body irradiation-based (busulfan intravenous), ATG, fludarabine, clofarabine, and total body irradiation, depending on diagnosis. Fludarabine phosphate was given daily from day -5 to -2 relative to transplantation, at a cumulative dose of 160 mg/m^2. Busulfan was given intravenously, targeted to a myeloablative cumulative exposure of 90 mg*h/L, or 30 mg*h/L for Fanconi anemia patients, (expressed as area under the plasma-concentration-time curve (AUC)). rATG was added in the unrelated donor HCT setting and given in 4 consecutive days from day -9 (10 mg/kg <30 kg, 7,5 mg/kg >30 kg). Graft-versus-host disease (GvHD) prophylaxis consisted of cyclosporine A (CsA) combined with methotrexate 10 mg/m^2 at day +1, +3 and +6 in bone marrow recipients, and CsA combined with prednisone 1 mg/kg/day in CB recipients until day +28, which was tapered in 2 weeks. Target trough levels for CsA were 200–350 mg/L. All patients were prophylactically treated with acyclovir, and received partial gut decontamination with ciprofloxacin and fluconazole and *Pneumocystis jiroveci* pneumonia prophylaxis using co-trimoxazol. All CBTpatients received 10 mg/kg granulocyte colony-stimulating factor (G-CSF; Neupogen®) from day +7 after HCT until neutrophils were above 2000 cells/μL. Patients were treated in high-efficiency, particle-free, air-filtered, positive-pressure isolation rooms.

Outcomes

In the first part of this study, we associated innate recovery with IR probability after HCT. Innate recovery was a model derived parameter described in detail in the statistics section. In short, it was a latent variable built up from over-time monocyte, neutrophil, and NK-cell counts in the blood of patients after HCT. For this, time-dependent dynamics of all subsets were assumed to be similar and directly related to the underlying latent variable. Low or high innate recovery was defined as having reached a maximum re-

covery below (<100%) or above (>100%) the standard innate cell reference value (100%), respectively. This reference value was defined as the median observed cell count (adjusted per cell type) after reaching full recovery (50-200 days after transplantation). IR was defined as having $> 50 \cdot 10^6$ CD4+ T-cell/L blood in two consecutive measurements within the first 100 days after transplantation, as this cut off was found to be the best predictor for outcomes[1]. In the second part, we aimed to study the differences in innate recovery over the first 200 days between CBT and BMT. Additionally, we evaluated whether the differences found could be explained biologically.

Immune monitoring

Absolute numbers of neutrophils and monocytes were measured in EDTA-treated whole blood at least 3 times per week during the admission period of 3-5 weeks after HCT, until a count of $0.4 \cdot 10^9$ cells/L was reached using a Sapphire Coulter Counter (Abbott). From then on cell levels were measured at least every other week up to twelve weeks post-HCT and monthly thereafter up to six months, followed by every 3 months until 1 year, and twice in the 2nd year after HCT with additional analyses of absolute NK-cells (CD3-CD16+CD56+) and CD4+ T-cells (CD3+CD4+CD8-) using TruCOUNT tubes (BD Biosciences, Erembodegem, Belgium) / Sapphire (Abbott). For more in-depth evaluation of monocyte and NK-cell subsets after transplantation, biobanked samples taken from 15 CBT and 19 BMT recipients, at week 2, 4, 6, 8, 16, and 24 weeks, were thawed and stained for flow cytometry. Monocyte subsets were gated as CD3- cells; CD14++CD16- (classical), CD14++CD16+ (intermediate), and CD14+CD16++ (non-classical). For NK-cell subset evaluation, CD3- cells were gated as CD56++CD16- (naïve), CD56+CD16+ (transitional), and CD56dimCD16+ (effector). Samples for neutrophil subset evaluation were not available. Antibodies used for in-depth immune monitoring were CD3-AF700 (clone UCHT1, Biolegend), CD56-PE-Cy7 (clone NCAM16.2, BD Bioscience), CD16-BV510 (clone 3G8, BD Horizon), and CD14-eFluor780 (clone 61D3, eBioscience).
For assessment of proliferation of cells within grafts, 6 CB-grafts and 6 BM-grafts were thawed and stained with extracellular CD3-AF700 (clone UCHT, Biolegend) and CD14-APC-eFluor780, and intracellular Ki-67-PerCP-Cy5.5 (clone B56, BD Bioscience) after permeabilization using the eBioscience kit (eBioscience, Thermo Fisher Scientific Inc). For assessment of proliferation of monocytes after transplantation, samples from 15 CBT and 19 BMT recipients were examined for intracellular Ki-67 expression in CD3-CD14+ cells, at week 2, 4, 6, 8, 16, and 24 after HCT. Samples were measured on

a BD LSR Fortessa (BD Bioscience). Results of innate cell proliferation in grafts were independent from innate cell proliferation after transplantation.

Statistics

Duration of the follow-up was defined as the time from HCT to the last assessment for surviving patients or death. Actual cell counts of monocytes, neutrophils, and NK-cells, in patients during the first 200 days after HCT were evaluated using non-linear mixed effects modeling (with NONMEM v.7.3)[14]. Models were parameterized in terms of fixed effects (population mean) and two levels of random effects; per subject (inter-individual variability [IIV]) and per measurement (residual variability). A logistic growth model was used as a structural model. Discrimination between nested models was based on the likelihood ratio test. The objective function value (OFV)(ΔOFV) between two hierarchical models follows a χ^2 distribution, with n degrees of freedom for n parameters added.
Covariates that were tested for inclusion in the models were: age at transplantation (continuous), diagnosis (malignant/benign), GvHD status (time-dependent: $<$ or $\geq$ grade 2), year of transplantation (continuous), and post-transplantation exposure of rATG. Because of the high correlation between the recovery of the different innate subsets (i.e. neutrophils, monocytes, and NK-cells), we developed a combined model for these three innate cell types describing the full dynamics of these cells after transplantation. For this, innate recovery was modelled as a latent variable, which was related to the different cell types using scaling to this latent variable, inclusion of relevant covariates, and inclusion of different residual errors for each subtype. A pre-defined reference value for innate recovery could not be derived from literature. Therefore, we assessed at which time cell counts stabilized for each individual patient, implying that the recovery is at its maximum. This time was set to where 95% of patients reached an observed (donor-derived) cell count above their model-estimated maximum. The variability of this maximum was the main independent determinant in the manuscript. We chose the median value for all patients as the threshold between good and bad maximal recovery. Subsequently, the difference in innate recovery for patients with an early or late IR was tested as a covariate in the innate recovery model. To test the predictive value of innate recovery on recovery of the adaptive system, individual Bayesian estimates for innate recovery were generated using the POSTHOC option of NONMEM. These parameters were tested as predictors in an IR parametric time-to-event (TTE) model. The TTE-model was constructed by 1) choosing the optimal base-

line hazard function and 2) finding predictors for the hazard of IR. The baseline hazard resulting in the lowest Akaike information criterion was regarded as optimal. Predictors, including the individual estimates from the innate recovery model, were chosen by stepwise forward addition and backward deletion, with a significance level of 0.05. Continuous covariates were evaluated with a linear relationship and as a reduced cubic spline with 3, 4 and 5 degrees of freedom. The optimal relationship in the univariate setting was kept throughout the TTE-model building process.
To evaluate the influence of cell source on innate recovery, cell counts were modelled per innate cell subset. Covariates, other than cell source, found to be significant in the innate recovery model were *a priori* included in these models for each subset. R[15] version 3.3 was used to make LOESS-regression curves for visualization of innate cell recovery in different patient groups and for the building of the TTE-models of IR, and to illustrate and analyse differences in Ki-67 expression in CD3-CD14+ cells within CB- and BM-grafts. We used Graphpad v.6 to perform two-tailed paired t-tests to evaluate subsets, and CD3-CD14+ Ki-67 expression differences after CBT or BMT.

3.4 Results

Patients

Between 2007 and 2016, data from 312 allogeneic transplants was available. To optimize the cohort for data availability and homology, 33 patients were excluded for unsuccessful engraftment of their first transplant, 7 were excluded for receiving serotherapy other than Thymoglobulin (i.e. Alemtuzumab or ATG Fresenius), and 66 were excluded for lacking immune recovery data for one of the analyzed subsets. Thus, 205 patients were included with a median age of 7.3 years (range 0.16-23) at time of HCT. In this cohort, 129 patients received CB and 76 received BM. Median time to follow-up was 1412 days. In total, 3785 neutrophil samples (median 17, range 3-58 samples per patient) were available. For monocytes and NK-cells these numbers were 5887 (median 25, range 3-84) and 1844 (median 9, range 0-22) respectively. Patient characteristics and differences between transplantation groups are summarized in table 1.

Table 1: Patient characteristics

Variable	Total	BM	CB
Age at HCT***	7.3 (0.16-23)	11 (0.45-20)	4.9 (0.16-23)
Sex			
F	80 (39%)	30 (39%)	50 (39%)
M	125 (61%)	46 (61%)	79 (61%)
Diagnosis***			
Autoimmune disease	2 (0.98%)	0 (0%)	2 (1.6%)
Bonemarrow failure	22 (11%)	14 (18%)	8 (6.2%)
Immunedeficiency	37 (18%)	7 (9.2%)	30 (23%)
Malignancy	105 (51%)	50 (66%)	55 (43%)
Metabolic/inborn	39 (19%)	5 (6.6%)	34 (26%)
Serotherapy***			
No serotherapy	65 (32%)	41 (54%)	24 (19%)
Thymoglobulin	140 (68%)	35 (46%)	105 (81%)
Conditioning**			
Busulfan based	168 (82%)	53 (70%)	115 (89%)
Fludarabine + cyclophosphamide	13 (6.3%)	9 (12%)	4 (3.1%)
Other	6 (2.9%)	5 (6.6%)	1 (0.78%)
TBI based	18 (8.8%)	9 (12%)	9 (7%)
Year of HCT			
2007-2011	64 (31%)	25 (33%)	39 (30%)
2011-2014	67 (33%)	26 (34%)	41 (32%)
2014-2016	74 (36%)	25 (33%)	49 (38%)
Median Follow-up	1838	1851.5	1795

Statistically significant variables between CB & BM at levels <0.05, <0.01 and <0.001 are marked by * , **, *** respectively
BM = Bonemarrow
CB = Cord blood

Innate immune cell recovery in relation to CD4+ T-cell reconstitution probability

We evaluated the association between recovery of innate immune cells and IR probability in both directions (figure 1). In figure 1A, a percentage of <100% indicates a decreased innate recovery, and >100% an increased innate recovery, compared to the predicted innate cell reference value (which is 100%). In our model, each innate immune cell subset contributed to the model, thereby retaining valuable information on innate immune recovery. The magnitude of contribution to the model per subset was assessed by the residual error per subset. Monocyte recovery associated with the lowest residual error (65%) compared to neutrophils (75%) and NK-cells (75%), indicating a strong contribution to the innate recovery model. In this model, we found that patients who had a successful CD4+ T-cell recovery also reached higher innate cell counts of $> 100\%$ of the predicted innate cell reference value, which was significantly different from patients who had delayed CD4+ T-cell recovery ($p < 0.001$, figure 1A). The IR probability over time was best characterized by a log-logistic distribution. During the forward inclusion and backwards deletion procedure it was found that age at transplantation, year of transplantation, and post-transplantation rATG exposure were independent predictors for the IR probability. In this multivariate setting (correcting for these covariates), we found that reaching a higher innate immune cell recovery strongly related to IR (continuous, $p < 0.001$; figure 1B). Stem cell source did not have a significant impact on IR probability in this cohort, after correcting for the above mentioned covariates. In addition, cumulative incidence of IR was increased in patients reaching innate cell counts above the previously defined standard reference value. Thus, our model indicates that patients with either monocytes $> 0.89 \cdot 10^9$, and/or neutrophils $> 4.2 \cdot 10^9$/L, and/or NK-cells $> 0.34 \cdot 10^9$/L within the first 50 days after HCT have higher chance of having successful IR.

Innate immune cell recovery according to cell source

Differences in innate immune cell recovery after BMT and CBT were observed. The recovery of monocytes (figure 2A), neutrophils (figure 2B), and NK-cells (figure 2C) is similar during the first ~2030 days after either BMT or CBT. After this first month, the numbers of monocytes ($p < 0.001$; 0.99 after CBT vs. $0.55 \cdot 10^9$/L after BMT), neutrophils (p<0.001; 4.05 after CBT vs. $3.05 \cdot 10^9$/L after BMT), and NK-cells ($p < 0.001$; 0.36 after CBT vs. $0.24 \cdot 10^9$/L after BMT) were significantly higher in CBT compared to

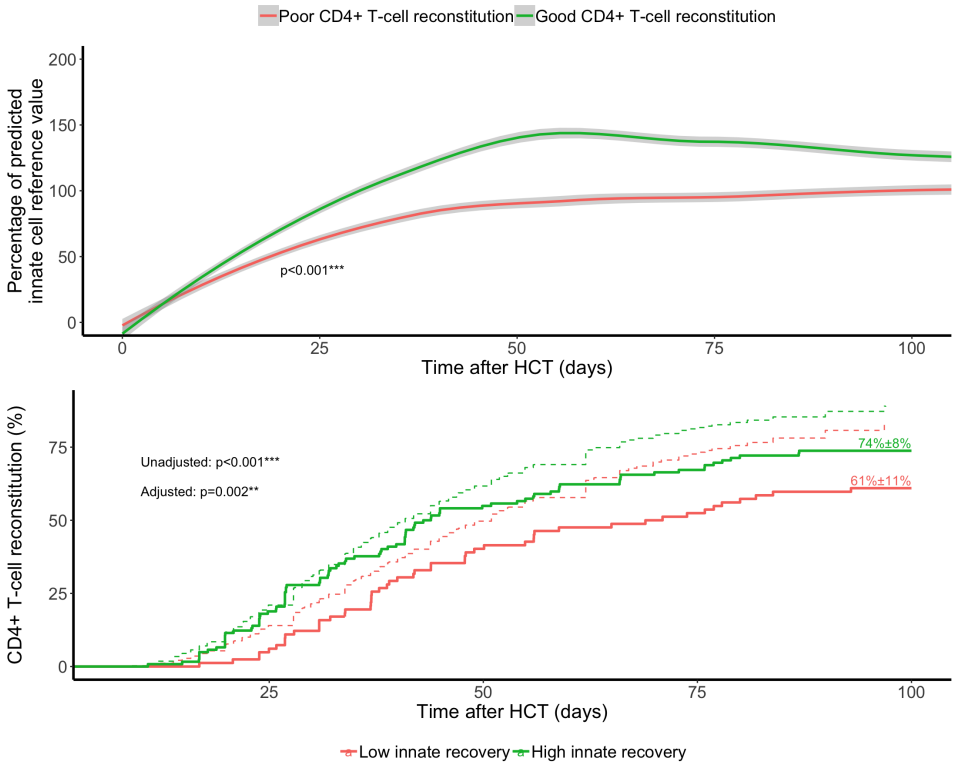

Figure 1: Innate recovery in relation to IR. (A) The percentage of observed innate cell counts relative to the population predicted reference value in patients who did not have successful IR (red lines) have different dynamics, compared to patients who did have successful IR (blue lines). A percentage of <100% indicates a decreased innate recovery, and >100% an increased innate recovery, compared to the predicted innate cell reference value (which is 100%). Grey areas represent 95% confidence intervals of LOESS-regression curves and p-values are derived from backwards deletion of adaptive immune recovery implemented in the innate cell model as a discrete variable (low/high). Successful CD4+ T-cell recovery was defined as having $>50*10^6$ CD4+ T-cells/L blood, in 2 consecutive measurements, within 100 days after HCT. Innate recovery was defined as neutrophil, monocyte, and NK-cell reconstitution. (B) Cumulative incidence curves depict higher probability of IR in patients with higher innate recovery ($p < 0.001$). Dashed lines reflect the model predictions.

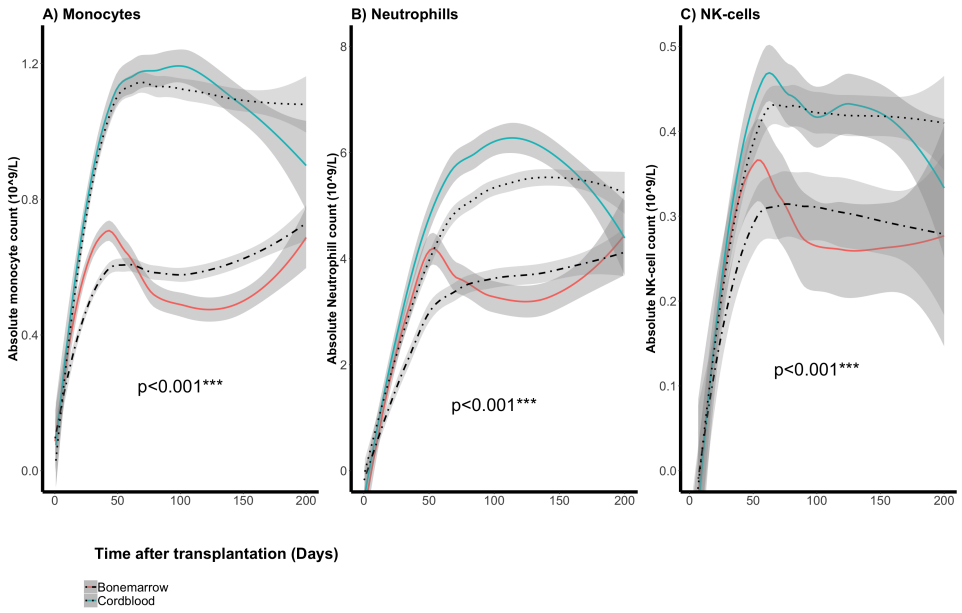

Figure 2: Recovery of innate immune cells after HCT with CB or BM. Innate subset recovery after transplantation with bone marrow (red lines, n=82) or CB (blue lines, n=131), grey areas represent 95% confidence intervals of LOESS-regression curves. Model predictions for recovery after BMT (intermitted line) and after CBT (dotted line) are included. (A) Absolute monocyte counts ($\cdot 10^9$ monocytes/L blood) after CBT and BMT. (B) Absolute neutrophil counts ($\cdot 10^9$ monocytes/L blood) in CB and bone marrow recipients. (C) Absolute NK-cell counts ($\cdot 10^9$ monocytes/L blood) after BMT and CBT.

BMT recipients. This difference in innate recovery between BMT and CBT diminished by day 200 after transplantation. Interestingly, in our pediatric cohort, CBT was also related to a higher chance of having successful IR, compared to BMT ($p = 0.018$), when residual ATG exposure after transplantation was low (< 10 active-ATG·day/mL) or absent[16]. Nevertheless, the higher innate recovery after CBT only increases the chance of having successful IR ($> 50 \cdot 10^6$ cells/L blood within 100 days after transplantation) by 3.0%, compared to BMT, when correcting for all relevant covariates (age at transplantation, year of transplantation, and post-transplantation ATG exposure).

Because of these differences, we performed additional in-depth immune monitoring in 15 CBT and 19 BMT recipients. We evaluated which subsets within monocytes (figure 3A) and NK-cells (figure 3B) were most prominent

during recovery. Data from neutrophil subsets were not available. Classical monocytes (CD14++CD16-) predominated (±70-80%) within 2 months and 4-6 months after CBT and BMT, over intermediate (CD14++CD16+) and non-classical monocytes (CD14+CD16++). Although absolute monocyte numbers differed after CBT and BMT, no differences were observed in the percentages of monocyte subsets. NK-cells within the first 2 months after both CBT and BMT mostly consisted of naïve (CD56++CD16-; ±30-40%) and effector NK-cells (CD56dimCD16+; ±50-60%), with only low percentages of transitional NK-cells (CD56+CD16+) present. 4-6 Months after both CBT and BMT, the majority of NK-cell recovery consisted of the effector subset (±70-80%).

Innate immune recovery relates to graft proliferation potency

We further investigated what biological explanation could be found for the higher monocyte level after CBT. Differences between proliferation potential of BM-monocytes and CB-monocytes would provide a possible explanation for the difference in recovery. These data were not available for neutrophils and NK-cells. Therefore, we investigated differences in Ki-67 expression between CD3-CD14+ cells ex vivo during the first 6 months in 15 CBT and 19 BMT recipients (figure 4A), and CD3-CD14+ cells in 6 CB- and 6 BM-grafts (figure 4B). Data from grafts and samples after transplantation were not related. We found a higher percentage of Ki-67 in CB-graft CD3-CD14+ cells compared to BM-grafts ($p = 0.041$). An enhanced expression of Ki-67 in CD3-CD14+ cells was also found in blood of patients after CBT compared to BMT ($p = 0.003$). During the first 6 months after transplantation, proliferation decreased over time, while higher percentages of Ki-67 remained in CB-derived CD3-CD14+ cells compared to BM. Interestingly, patients who received CB- (figure 5A; $p = 0.004$) or BM-grafts (figure 5B; $p = 0.01$) with higher innate Ki-67+ expression, had higher innate recovery (above median) during the first 200 days after CBT or BMT, and vice versa. These results indicate that enhanced innate recovery after CBT and BMT is most probably due to the enhanced proliferation potential of graft-innate cells.

3.5 Discussion

This study is the first to both analyze innate recovery in context of adaptive reconstitution and compare innate recovery after CBT and BMT. We showed that superior innate immune cell recovery predicts higher probabil-

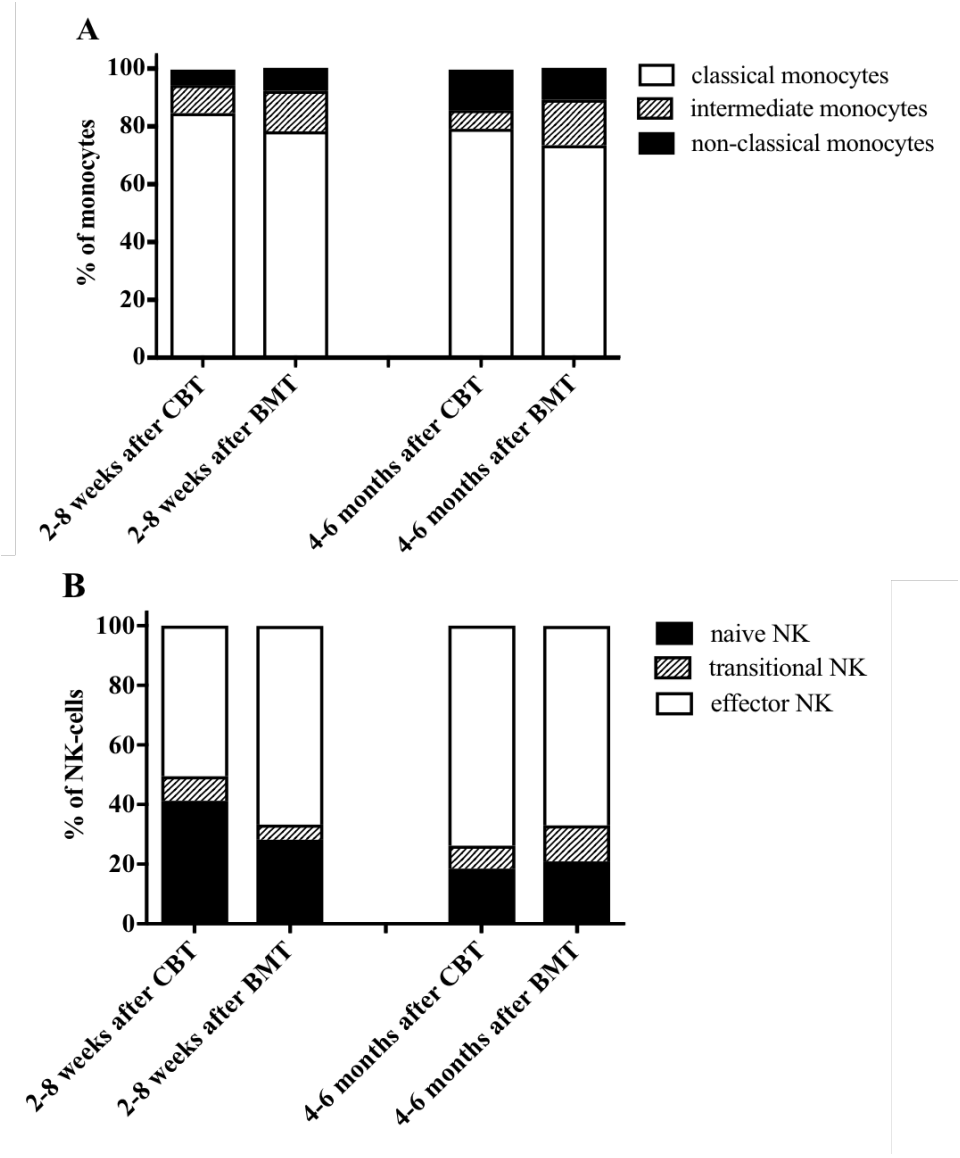

Figure 3: Recovery of innate cell subsets after HCT with CB or bone marrow. (A) Monocyte subsets after CBT (n=15) and BMT (n=19), as percentages of total monocytes. (B) NK-cell subsets after CBT (n=15) and BMT (n=19), as percentages of total NK-cells.

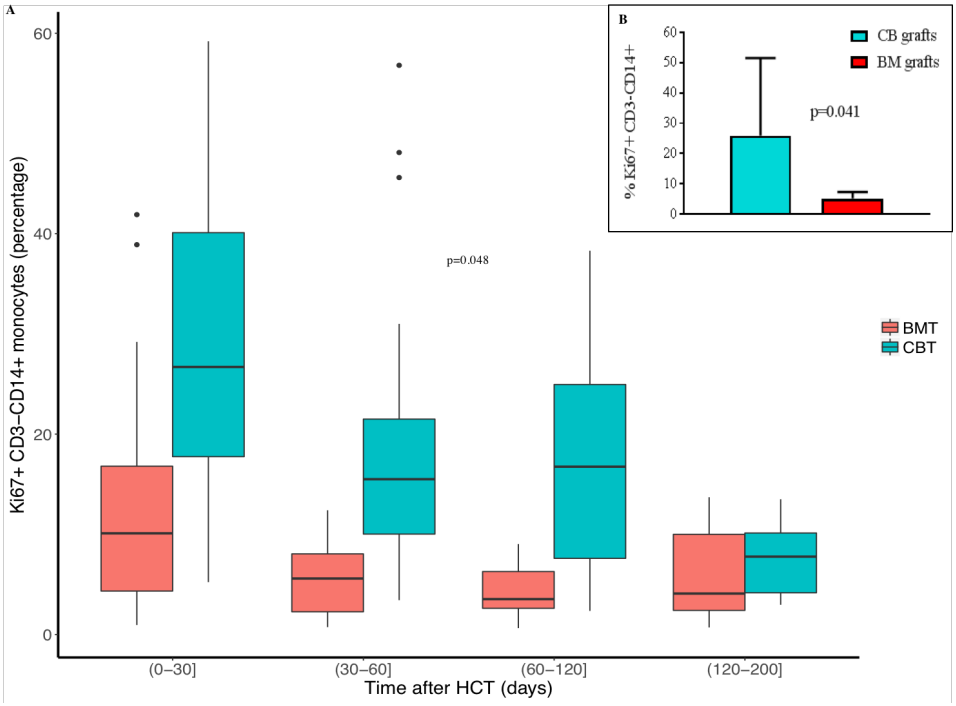

Figure 4: Proliferation of innate cells after transplantation with, and in grafts of, CB and BM. Proliferation of innate cells after transplantation with, and in grafts of, CB and BM.

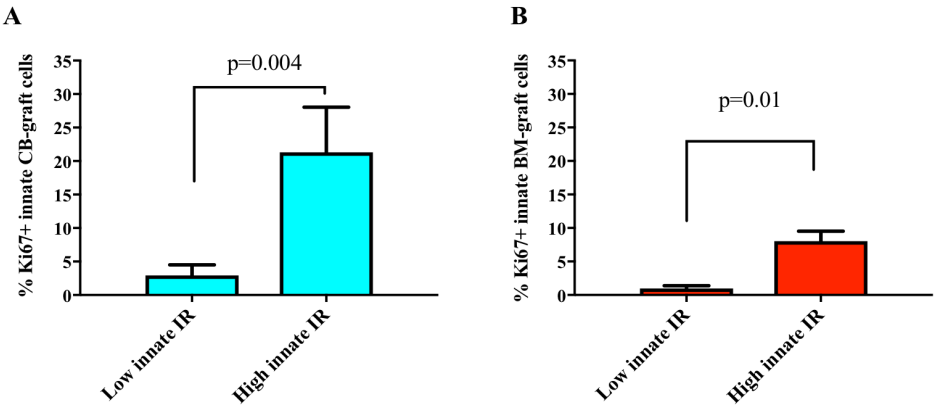

Figure 5: Innate immune recovery relates to proliferation rate in CB- and BM-graft innate cells. Percentages of Ki-67+ CD3-CD14+ cells within 6 CB- (A) and 6 BM-grafts (B). Groups are made based on low innate immune recovery (below median; left bars, n=3 CB and n=3 BM) and high innate immune recovery (above median; right bars, n=3 CB and n=3 BM), relative to graft source. The percentage of Ki-67+ innate cells was higher in CB- (p=0.004) and BM-grafts (p=0.01) in patients with high innate IR, compared to patients with low innate IR, as analyzed with two-tailed paired t-tests.

ity of IR after pediatric HCT. This appeared to be cell source dependent, since we also observed higher levels of monocytes, neutrophils, and NK-cells during the first 200 days after CBT compared to BMT. This emphasizes the potency of monitoring early innate immune cell recovery after HCT for early prediction of IR probability, which is highly related to better survival chances after HCT[1–9].

The association between innate immune cell levels and IR might be a result from enhanced graft activity. This hypothesis is supported by the relation we found between innate immune cell recovery and proliferation potential in graft cells. Grafts containing more proliferating cells resulted in more robust innate recovery, which might indicate a more optimal graft condition. The generally higher CB graft-cell proliferation capacity could also explain why IR is faster after CBT compared to BMT, in patients with no/low serotherapy exposure after transplantation[16]. Innate cell recovery might thus represent a healthy niche for homeostatic peripheral expansion (HPE) mediated T-cell reconstitution. Moreover, innate immune cells may also directly stimulate HPE of T-cells via TCR-stimulation and/or cytokine production[17–20]. Therefore, further research is needed to assess what mechanisms exactly explain the relation between the recovery of innate immune cells and early IR after HCT.

Our finding that monocytes recover to higher levels after CBT compared to BMT has been reported before[21], but no other studies so far have described this difference for neutrophil and NK-cell recovery. We found a biologically relevant explanation for the strikingly higher monocyte levels after CBT compared to BMT. The number of proliferating monocytes within CB-grafts was significantly higher than in BM-grafts. After transplantation, monocyte proliferation remained higher at least during the first 6 months after CBT compared to BMT. In addition, we found a relation between the proliferation potential in both CB- and BM-graft cells and innate recovery after transplantation in the respective individual patients. Of note, all CBT patients in our cohort were treated with G-CSF (from day +7 till neutrophils reach $> 2000/\mu$L) as standard-of-care to enhance neutrophil levels, while BMT recipients were not[16]. This might also affect the recovery of both neutrophils and monocytes, since monocytes have functional G-CSF receptors as well[22]. However, it is unclear if, and to what extent, G-CSF treatment after CBT would affect long-term monocyte recovery as it is given only for on average 2-2.5 weeks, and higher monocyte proliferation is already evident within CB-grafts. Altogether, the observations of higher proliferation in CB-graft innate cells and after CBT provide a possible explanation for the higher innate levels observed in CBT patients.

The higher innate immune cell recovery after CBT, compared to BMT, can be related to higher normal levels of innate cells in the blood of healthy newborns, compared to healthy adults[23,24]. For instance, monocyte levels in blood of healthy adults are between $0 - 0.8 \cdot 10^9$/L, while we measured around $0.55 \cdot 10^9$/L after BMT[24]. For healthy newborns the reference range is $0.4 - 3.1 \cdot 10^9$/L[24], and levels of monocytes after CBT reached around $0.99 \cdot 10^9$/L. In addition, normal neutrophil levels are between $1.8 - 7.7 \cdot 10^9$/L in adults; we observed $3.05 \cdot 10^9$/L after BMT, and between $6 - 26 \cdot 10^9$/L in newborns; after CBT we measured $4.05 \cdot 10^9$/L neutrophils[24]. Also for NK-cells, recovery after CBT ($0.36 \cdot 10^9$/L) fits best within the range of normal NK-cell counts in newborns of $0.3 - 1.7 \cdot 10^9$/L, while after BMT ($0.24 \cdot 10^9$/L) the levels meet the reference range in adults of $0.1 - 0.4 * 10 \cdot 10^9$/L[23]. CBT, therefore, seems to recapitulate fetal ontogeny in terms of immune reconstitution, as described by others[25].

Since immune reconstitution data consist of values over time, it demands over-time evaluation to be able to optimally relate immune reconstitution to outcomes after HCT. We here show that immune-modeling is feasible and potentially provides a more accurate and unbiased approach compared to dichotomous evaluations as presented by others[26,27]. Ultimately, more insights into the mechanisms and variables that influence immune cell re-

constitution after HCT could lead to a model that is able to predict immune reconstitution and clinical outcomes per patient, based on individual data. Our findings suggest importance of monitoring early innate immune cell recovery to predict IR later on, and show differences in innate recovery after CBT and BMT in pediatric patients. Finally, these findings are of use to achieve more predictable immune reconstitution to enhance outcome in future HCT recipients.

References

1. Admiraal, R., de Koning, C., Lindemans, C. A., Bierings, M. B., Wensing, A. M., Versluys, A. B., *et al.* Viral Reactivations and Associated Outcomes in Context of Immune Reconstitution after Pediatric Hematopoietic Cell Transplantation. *J Allergy Clin Immunol* (2017).
2. Bartelink, I. H., Belitser, S. V., Knibbe, C. A., Danhof, M., de Pagter, A. J., Egberts, T. C., *et al.* Immune reconstitution kinetics as an early predictor for mortality using various hematopoietic stem cell sources in children. *Biol Blood Marrow Transplant* **19,** 305–13 (2013).
3. Berger, M., Figari, O., Bruno, B., Raiola, A., Dominietto, A., Fiorone, M., *et al.* Lymphocyte subsets recovery following allogeneic bone marrow transplantation (BMT): CD4+ cell count and transplant-related mortality. *Bone Marrow Transplant* **41,** 55–62 (2008).
4. Buhlmann, L., Buser, A. S., Cantoni, N., Gerull, S., Tichelli, A., Gratwohl, A., *et al.* Lymphocyte subset recovery and outcome after T-cell replete allogeneic hematopoietic SCT. *Bone Marrow Transplant* **46,** 1357–62 (2011).
5. Fedele, R., Martino, M., Garreffa, C., Messina, G., Console, G., Princi, D., *et al.* The impact of early CD4+ lymphocyte recovery on the outcome of patients who undergo allogeneic bone marrow or peripheral blood stem cell transplantation. *Blood Transfus* **10,** 174–80 (2012).
6. Kim, D. H., Sohn, S. K., Won, D. I., Lee, N. Y., Suh, J. S. & Lee, K. B. Rapid helper T-cell recovery above 200 x 10 6/l at 3 months correlates to successful transplant outcomes after allogeneic stem cell transplantation. *Bone Marrow Transplant* **37,** 1119–28 (2006).
7. Pourgheysari, B., Piper, K. P., McLarnon, A., Arrazi, J., Bruton, R., Clark, F., *et al.* Early reconstitution of effector memory CD4+ CMV-specific T cells protects against CMV reactivation following allogeneic SCT. *Bone Marrow Transplant* **43,** 853–61 (2009).
8. Ritter, J., Seitz, V., Balzer, H., Gary, R., Lenze, D., Moi, S., *et al.* Donor CD4 T Cell Diversity Determines Virus Reactivation in Patients After HLA-Matched Allogeneic Stem Cell Transplantation. *Am J Transplant* **15,** 2170–9 (2015).

9. Admiraal, R., Chiesa, R., Lindemans, C. A., Nierkens, S., Bierings, M. B., Versluijs, A. B., *et al.* Leukemia-free survival in myeloid leukemia, but not in lymphoid leukemia, is predicted by early CD4+ reconstitution following unrelated cord blood transplantation in children: a multicenter retrospective cohort analysis. *Bone Marrow Transplant* **51,** 1376–1378 (2016).

10. Admiraal, R., Lindemans, C. A., van Kesteren, C., Bierings, M. B., Versluijs, A. B., Nierkens, S., *et al.* Excellent T-cell reconstitution and survival depend on low ATG exposure after pediatric cord blood transplantation. *Blood* **128,** 2734–2741 (2016).

11. De Koning, C., Plantinga, M., Besseling, P., Boelens, J. J. & Nierkens, S. Immune Reconstitution after Allogeneic Hematopoietic Cell Transplantation in Children. *Biol Blood Marrow Transplant* **22,** 195–206 (2016).

12. Lindemans, C. A., Chiesa, R., Amrolia, P. J., Rao, K., Nikolajeva, O., de Wildt, A., *et al.* Impact of thymoglobulin prior to pediatric unrelated umbilical cord blood transplantation on immune reconstitution and clinical outcome. *Blood* **123,** 126–32 (2014).

13. Oshrine, B. R., Li, Y., Teachey, D. T., Heimall, J., Barrett, D. M. & Bunin, N. Immunologic recovery in children after alternative donor allogeneic transplantation for hematologic malignancies: comparison of recipients of partially T cell-depleted peripheral blood stem cells and umbilical cord blood. *Biol Blood Marrow Transplant* **19,** 1581–9 (2013).

14. Sheiner, L. B. & Beal, S. L. Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. *J Pharmacokinet Biopharm* **8,** 553–71 (1980).

15. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

16. De Koning, C., Gabelich, J. A., Langenhorst, J., Admiraal, R., Kuball, J., Boelens, J. J., *et al.* Filgrastim enhances T-cell clearance by antithymocyte globulin exposure after unrelated cord blood transplantation. *Blood Adv* **2,** 565–574 (2018).

17. Gordon, S. & Taylor, P. R. Monocyte and macrophage heterogeneity. *Nat Rev Immunol* **5,** 953–64 (2005).

18. Kalyan, S. & Kabelitz, D. When neutrophils meet T cells: beginnings of a tumultuous relationship with underappreciated potential. *Eur J Immunol* **44,** 627–33 (2014).

19. Leliefeld, P. H., Koenderman, L. & Pillay, J. How Neutrophils Shape Adaptive Immune Responses. *Front Immunol* **6,** 471 (2015).

20. Leon, B., Lopez-Bravo, M. & Ardavin, C. Monocyte-derived dendritic cells. *Semin Immunol* **17,** 313–8 (2005).

21. Turcotte, L. M., Cao, Q., Krahn, E., Curtsinger, J., Yingst, A. M., Cooley, S., *et al.* Monocyte Subpopulation Recovery and Outcomes Following Hematopoietic Cell Transplantation. *Blood* **128(22):2232** (2016).

22. Boneberg, E. M., Hareng, L., Gantner, F., Wendel, A. & Hartung, T. Human monocytes express functional receptors for granulocyte colony-stimulating factor that mediate suppression of monokines and interferon-gamma. *Blood* **95,** 270–6 (2000).

23. De Vries, E., de Bruin-Versteeg, S., Comans-Bitter, W. M., de Groot, R., Hop, W. C., Boerma, G. J., *et al.* Longitudinal survey of lymphocyte subpopulations in the first year of life. *Pediatr Res* **47,** 528–37 (2000).

24. KD, M. *Clinical Laboratory Medicine* (2002).

25. Hiwarkar, P., Hubank, M., Qasim, W., Chiesa, R., Gilmour, K. C., Saudemont, A., *et al.* Cord blood transplantation recapitulates fetal ontogeny with a distinct molecular signature that supports CD4(+) T-cell reconstitution. *Blood Adv* **1,** 2206–2216 (2017).

26. Barker, C. I., Germovsek, E., Hoare, R. L., Lestner, J. M., Lewis, J. & Standing, J. F. Pharmacokinetic/pharmacodynamic modelling approaches in paediatric infectious diseases and immunology. *Adv Drug Deliv Rev* **73,** 127–39 (2014).

27. Hoare, R. L., Veys, P., Klein, N., Callard, R. & Standing, J. F. Predicting CD4 T-Cell Reconstitution Following Pediatric Hematopoietic Stem Cell Transplantation. *Clin Pharmacol Ther* **102,** 349–357 (2017).

Conditioning: Optimizing the busulfan-fludarabine regimen

4 **Busulfan and fludarabine in conditioning:** a potent duo adjustable to precision?

Jurgen B Langenhorst, Charlotte van Kesteren, Alwin DR Huitema, Jaap Jan Boelens

Submitted (March 2019)

4.1 Abstract

Busulfan (Bu) and fludarabine (Flu) are increasingly combined in conditioning regimens prior to allogeneic hematopoietic cell transplantation (HCT). Therefore, the current review aims to summarize the suitability, clinical pharmacology and the optimal use of BuFlu in HCT. BuFlu has been proven a less toxic alternative to the historically used regimens cyclophosphamide (Cy) + total body irradiation (TBI) and BuCy. Owing to the synergistic anti-leukemic effect of Bu and Flu, while preserving graft-versus-leukemia effect, the relapse rates remain similar compared to BuCy and CyTBI. Toxicity and consequently non-relapse mortality (NRM) are substantially lower. As is common for high-dose chemotherapeutic agents, the therapeutic windows for both drugs are narrow. Minimal exposure of both agents is essential to prevent graft failure (Bu and Flu) and relapse (Bu). Over-exposure to either agent leads to an unnecessary increase in toxicity and even NRM. Bu target attainment can be achieved in increasing order by intravenous dosing based on body weight, dosing based on population-pharmacokinetic models and therapeutic drug monitoring (TDM). Current dosing for Flu has been shown to yield unfavorable exposures in a large part of the population, thus, similarly to Bu, new dosing algorithms and TDM are needed.

4.2 Introduction

Busulfan (Bu) and fludarabine (Flu) are often combined in conditioning regimens prior to allogeneic hematopoietic cell transplantation (HCT). Bu is an alkylating agent, which forms the backbone of HCT conditioning regimens, because of its profound toxic effect on non-dividing marrow cells. Bu is currently used as an alternative to total body irradiation (TBI), to mediate both short- and especially long-term toxicity (such as cognitive impairment, growth hormone deficiencies, chronic pulmonary disease) associated with TBI.
Flu is a purine analogue, which is added to a variety of regimens, including Bu-based regimens. Flu is used in this setting for its toxic effect on mature lymphocytes causing strong immunosuppression. It was initially introduced in non-myeloablative regimens, but is mainly known as a less toxic alternative for cyclophosphamide (Cy) in Bu-based regimens.
Because both agents are currently used to reduce toxicity associated with their predecessor, the BuFlu combination is now often referred to as a reduced-toxicity regimen. As the BuFlu conditioning regimen is increasingly used in the HCT setting[1], the current review summarizes the suitability, clinical pharmacology and the optimal use of this combination in HCT.

4.3 Evolution of BuFlu regimen: why should we use it?

In the early HCT era (1970-1990), TBI combined with Cy (CyTBI) has been the conditioning regimen of choice. Disease relapse (~20%) and transplant related complications and mortality (~40%) occurred frequently, prompting researchers to explore alternative regimens. In 1983, Bu was introduced to replace TBI[2], showing promising results with survival rates reported of 40-60%[2,3]. However, the available oral formulation was inconvenient and hepatic toxicity in the form of sinusoidal obstruction syndrome (SOS) was notable (~40%).

Oral BuCy versus CyTBI

After introduction of Bu in HCT, studies started focusing on comparing oral Bu with TBI with regard to disease control and toxicity. These studies are summarized in table 1. In the early 90's four RCTs were published comparing oral Bu at 16 mg/kg to fractionated TBI at 12 Gy, acute lymphocytic leukemia (ALL) regimens also contained 120 mg/kg Cy. As can be seen in the table, BuCy compared favourably to TBI in chronic myeloid leukemia

(CML): similar 2-year overall survival (OS) (70-80%) with less toxicity in BuCy[4,5]. For acute myeloid leukemia (AML)[6], both relapse and non-relapse mortality (NRM) were found to be less in CyTBI, resulting in superior OS for this regimen. A randomized controlled trial (RCT) combining AML, ALL and CML[7] showed increased NRM after BuCy especially in patients with advanced disease, which was confirmed in a long-term analysis[8].
The data of aforementioned trials were combined in 2001 to analyze long-term outcomes stratified for AML and CML[9]. Herein, no statistical differences were found in either diagnosis, but a trend was found for increased survival in AML patients after CyTBI conditioning compared to BuCy. One smaller trial (n=43) in pediatric ALL found superior outcomes for TBI, but regimens were combined with an unusual regimen of Cy at 60 mg/kg and etoposide at 40 mg/kg. In a similar setting, a large registry-based study found inferior survival for BuCy, due to increased NRM and comparable disease control to CyTBI[10].
Overall, CyTBI tends to associate with superior survival for acute leukemia's (AML/ALL), while oral BuCy leads to reduced toxicity in CML and similar survival. Drawbacks of BuCy are likely to be partly attributable to unpredictable bio-availability of the oral formulation of Bu[11].

IV Bu versus TBI

Around 2000 an intravenous (IV) formulation of Bu was approved, which potentially circumvents limitations of oral Bu[11], such as high inconvenience for patients, added to the high variability in oral bio-availability. The latter might be responsible for unpredictable toxicity of oral Bu. Indeed, SOS and early mortality could be markedly reduced by replacing oral for IV Bu[12]; SOS probability was significantly reduced from 33% to 8% as well as 100-day NRM from 33% to 13%. This re-opened the discussion of CyTBI versus BuCy.
Although no new RCT's were initiated, several retrospective cohort studies were performed (table 2). Many of these studies only included a limited number of patients over long inclusion periods, which may potentially introduce period-bias in such studies[13–17]. To avoid such bias, fairly large registry-based studies including parallel transplantation programs using the different regimens, have been performed.
These comparisons showed, in contrast to oral Bu vs TBI, a trend to superior outcomes for IVBu-based regimens; especially in CML BuCy associated with increased survival[18]. For both ALL[19] and AML[20] survival was about similar for both regimens[20]. The advantage in CML was attributable to a

lower relapse incidence, where in the acute leukemia's a lower NRM in BuCy was compensated by increased relapse. In a combined malignancy setting, CyTBI was also compared to Bu combined with Flu (BuFlu). Here, the decrease in NRM remained, but relapse was comparable, thus resulting in increased OS[21]. In contrast, another study in several myeloid malignancies found reduced relapse in Bu-based regimens and comparable NRM[22]. The study started inclusion later and, as a possible consequence, over 50% of patients received Bu dosing based on individual Bu-pharmacokinetics, possibly explaining the difference in relapse. In pediatrics, the need for alternatives to TBI is even greater given its association with secondary malignancies[23] and cognitive impairment[24–26]. The two studies comparing IV BuCy to CyTBI showed similar survival[27,28], but increased long-term morbidity associated with TBI[28]. The more predictable IV BuCy regimen in general performed much better compared to its oral counterpart. This may indicate a narrow therapeutic window, with the possibility of using even more precise dosing to further increase Bu efficacy and reduce toxicity.

BuFlu versus BuCy

A more recent evolution in Bu-based regimens was replacing Cy for Flu. Major concerns for BuCy have always been raised regarding toxicity such as SOS because both Bu[29] and Cy[30] are well known hepatotoxic agents. Owing to the mild non-hematologic toxicity of Flu[31,32], the hypothesis was that NRM could be reduced.

Several studies compared BuCy to BuFlu (table 3). Initial retrospective studies again suffered from period bias, with IVBuFlu being compared to historical controls of (sometimes oral) BuCy[33–36], though BuFlu remained superior to BuCy after period-adjustment[37].

To provide more definitive answers, three RCT's were conducted[38–40]. The first study concluded superiority of BuCy in both relapse and NRM. However, the BuFlu arm contained more patients with ALL, associated with higher relapse probability. In addition, Flu and Bu were partly administered on different days, thereby potentially losing synergistic anti-leukemic effect of Bu and Flu[41]. The other RCT's found decreased NRM with the BuFlu regimen[39,40]. Only the final (largest) study included a power calculation, where a much larger sample size was found necessary than included in preceding studies[40]. As primary end-points were similar, the first two trials were probably under-powered.

Interestingly, neither of these trials used pharmacokinetic targeted Bu, which could aid in preventing SOS-related NRM especially in the BuCy group. In-

deed, survival was comparable in other retrospective studies[42–44] comparing Cy and Flu with a targeted Bu backbone, though (non-lethal) toxicity was still lower in the BuFlu arm.
Bu and Flu are at least comparable to their respective counterparts TBI and Cy regarding survival. Late toxicity associated with TBI, such as cataract, stunted growth and cognitive impairment[25,45,46] could be avoidable by using IV Bu. Also Flu appears to have a more favourable toxicity profile compared to Cy, with sufficient disease control, especially in AML.

4.4 **Clinical pharmacology of BuFlu:** how should we use it?

As BuFlu is now a widely used, relatively safe regimen, this section focuses on the clinical pharmacology and optimal use of these agents.

Busulfan pharmacology

Figure 1 displays a schematic overview of Bu metabolism[47,48]. Bu contains two labile methane sulfonate groups on opposite ends of a butane chain. When these groups are hydrolyzed, highly reactive, positively charged carbonium ions are formed. These ions possess the DNA-damaging and alkylating activity, resulting in cytotoxicity[41]. Of the administered and circulating moiety, approximately 2% is excreted unchanged in urine. The primary route of Bu clearance is through extensive metabolism in the liver. Initial inactivation occurs by conjugation to glutathione (GSH), both spontaneously and aided by the enzyme glutathione-S-transferase (GST). The inactive Bu-GSH-conjugate is further metabolized through the mercapturic acid pathway and beta-elimination, with the three main metabolites (sulfolane, 3-hydroxysulfolane and tetrahydrothiophene 1-oxide) recovered in urine[49].

Busulfan exposure response relationship

The relationship between Bu systemic exposure and treatment outcome has been of interest for almost 30 years. Target exposure has been quantified as area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($\mathrm{AUC}_{t_0-\infty}$) in $\mu Mol \cdot min$ per day of administration or $mg \cdot h/L$ for all (usually 4) administrations combined. First observations have shown increased risk of SOS on the high end of exposure, where the low end has been associated with high graft failure and relapse rates[50]. Given the detrimental effect of both under-exposure and over-exposure, an optimal exposure window can be imagined, where the combined probabilities

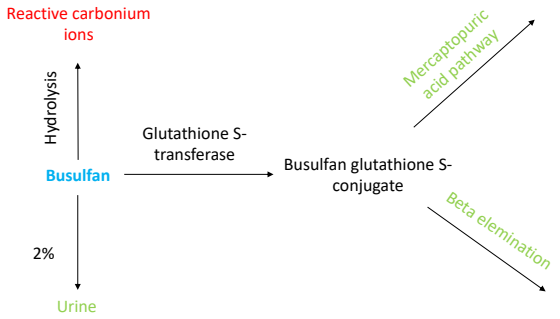

Figure 1: Pharmacology of busulfan. Blue depicts the parent drug, red active metabolites, and green definitive elimination routes.

of these events are minimized (figure 2). Studies, that sought to derive an optimal exposure are listed in table 4. Here it can be seen that the upper limit is mostly set around 97 $mg \cdot h/L \approx 6000$ $\mu Mol \cdot min$[51–53], though one study found 80 $mg \cdot h/L \approx 5000$ $\mu Mol \cdot min$[54]. In addition, most studies used non-compartmental pharmacokinetic analysis, which is known to potentially under-estimate the true exposure[55]. For the minimally necessary exposure more variable limits have been found ($60 - 90$ $mg \cdot h/L \approx 3800 - 5400$ $\mu Mol \cdot min$)[51,54–56].

Though the difference in methods for pharmacokinetic analyses (table 4) makes it difficult to compare absolute numbers between studies, possibly, the lower limit of exposure varies with different co-conditioning agents or diagnosis. This is however purely speculative as this effect appears to be absent in a large multi-center trial. Here, a similar target for all different indications included (malignant and benign) as well as conditioning regimens was found[55]. Interestingly, a recent RCT compared therapeutic drug monitoring (TDM) to standard dose in (mostly adult) AML/myelodysplastic syndrome (MDS) patients (median age: 50 years, range: 13-66 years) using a relatively high target exposure of 97 $mg \cdot h/L \approx 6000$ $\mu Mol \cdot min$[57]. Put together, it seams that Bu $AUC_{t_0 - \infty}$ target should be set in the range of $90 - 100$ $mg \cdot h/L \approx 5500 - 6000$ $\mu Mol \cdot min$ to achieve good engraftment, with minimal risk of relapse and NRM.

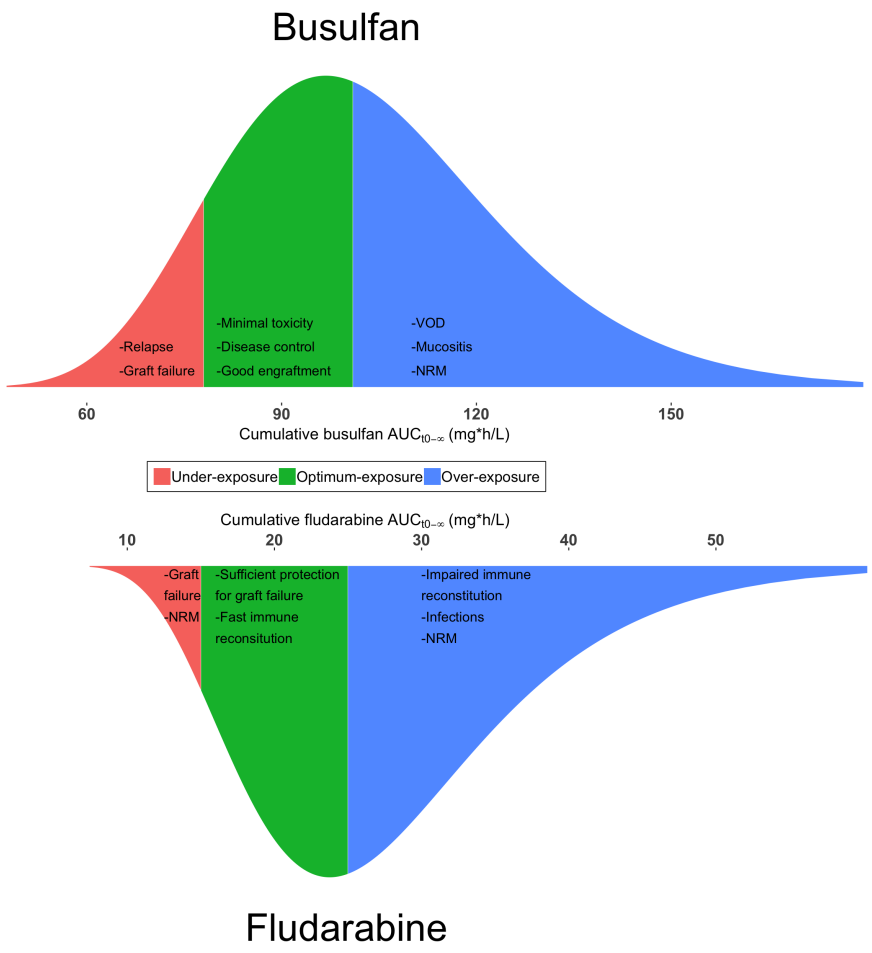

Figure 2: Exposure response relationships for buulfan and fludarabine. The x-axis represents the cumulative exposure ($\mu M*min/day$ for busulfan, $mg*h/L$ for fludarabine) and the y-axis the theoretical density of the corresponding exposure in an (on average) optimal body-size adjusted (body weight for busulfan; body-surface-area for fludarabine) dosing regimen, assuming a log-normal distribution of drug clearance.

Optimal busulfan exposure attainment

Three different dosing approaches to achieve the optimal Bu exposure can be imagined: dosing based on body-weight, an approach based on the predicted clearance calculated from patient characteristics and a population pharmacokinetic model (model-based), or dosing based on TDM. Currently,

both the Food and Drug Administration (FDA)[47] and European Medicine Agency (EMA)[58] recommend a cumulative dose of 3.2 mg/kg IV for adults and up to 4.4 mg/kg in children. Both labels suggest but do not oblige the use of TDM to adjust the dose according to individual pharmacokinetics.
McCune et al. showed that these recommended dosing algorithms lead to inadequate target attainment (57% and 70% for FDA and EMA, respectively), especially in children, where target attainment decreased to 61% and even 33% in FDA and EMA recommended dosing respectively[59].
An extensive population pharmacokinetic model-based algorithm using both age and weight increased target attainment to approximately 70% for all ages. However, it should be noted that the target area assessed was target ±25%, where deviations over 10% of the optimal exposure already show relevant differences in survival probability[55]. Including GST polymorphisms[60–65] did not result in substantially lower inter-individual variability (~20%). In figure 2, the expected exposures of a theoretically perfect dosing algorithm are depicted, showing that undesirable exposures will persist.
Though TDM coincides with higher financial costs and burden for staff and personnel, the unpredicted inter-individual variability can be largely overcome. Indeed, TDM was shown to increase target attainment to 82% (at the ±10% level)[66]. This was confirmed by a (previously mentioned) RCT (section 4.4). Superior OS (Δ20%) was found for the TDM guided dosing and, strikingly, this difference was nullified when an exposure matched comparison was performed. This confirms that Bu exposure was the main predictor for the survival difference.

Pharmacology of fludarabine

Flu (figure 3) is intravenously administered as a monophosphate prodrug (F-ara-AMP). It is very rapidly fully converted to the circulating metabolite F-ara-A, as which it is distributed intracellularly. Subsequently, intracellular phosphorylation takes place to the active metabolite fludarabine as the intracellular active moiety F-ara-ATP (FluTP) which is built into DNA and RNA, thereby inhibiting DNA/RNA-synthesis. This leads to apoptosis in both CML cells[67] and in cell types targeted in the HCT setting[68,69]. As FluTP is the active metabolite, exposure to it is expected to best predict efficacy. However, in some protocols anti-thymocyte globulin is also administered prior to Flu, causing a low number of cells to be available for quantification. Therefore, ex-vivo quantification of FluTP accumulation in pre-treatment samples has been proposed in HCT recipients[70], though a relationship between this accumulation and outcome has not been found[71].

Given these complexities and an apparent correlation between F-ara-A concentration and FluTP formation in cells[72], circulating F-ara-A has been used primarily to represent drug exposure.

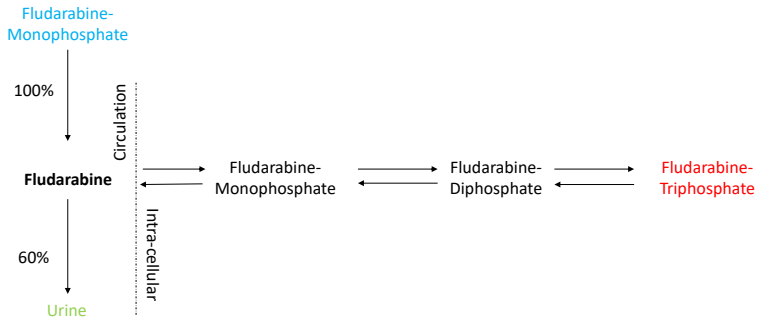

Figure 3: Pharmacology of fludarabine. Blue depicts the parent drug, red active metabolites, and green definitive elimination routes. Arrows reference to transformation directions and percentages depict the fraction of the parent that is transformed to the metabolite. The dashed line depicts the border between extra- and intra-cellular processes.

Fludarabine exposure response relationship

In contrast to Bu, less is known about the optimal dosing of Flu in HCT. Flu has a single circulating metabolite that is cleared mainly by the kidney[67,73] (figure 3) allowing for possible adjustability for further optimization. Flu is given for its toxicity to circulating lymphocytes[69] to prevent graft failure. Therefore, insufficient exposure is expected to increase the probability of graft failure, where high exposure is expected to associate with toxicity or NRM (figure 2). Indeed, the low end of F-ara-A exposure observed at 160 mg/m^2 in children has been associated with increased graft failure and decreased survival[72,74], concluding a minimal efficacious exposure of 15 mg*h/L. In a study with a higher dose of 200 mg/m^2, adverse effects of high Flu exposures could be observed and indeed an increase in NRM was observed with increasing Flu exposure[75]. Here, an upper limit of 26 mg*h/L

was advised for reduced risk of NRM. These limits were confirmed in a study combining both adults and children where the optimal exposure was set at 20 (15-25) mg*h/L[74].

Optimal Fludarabine exposure attainment

Neither the EMA nor the FDA label state specific dosing for Flu in HCT[76,77]. Generally, Flu is dosed 150-160 mg/m^2 on either the same days as Bu or before/after. Current body-surface-area based dosing however, leads to highly variable exposure[72,78], which is also shown in figure 2. This highly variable exposure seems also to influence the outcomes, including survival. This makes alternative dosing to better achieve optimal exposure, as done for Bu, the logical next step. Population pharmacokinetic studies have found actual body weight to better predict Flu clearance (thus AUC) than body-surface-area[72,78]. In addition, renal function seems to be an important predictor as can be expected for an almost fully renally cleared metabolite[72,78,79]. Dosing individualization can be achieved using body weight and renal function or TDM. The latter is expected to further decrease exposure variability, but comes with increased financial costs and logistics for medical staff and personnel. Possibly, part of this could be overcome by using the proposed method by Punt et al.[80], where quantification can occur in the same samples as Bu. This of course is only easily implemented, when Bu TDM is already routinely used.

4.5 Discussion

Bu Flu has been proven a less toxic alternative to the historically used regimens CyTBI and BuCy. Owing to the synergistic anti-leukemic effect and preserved graft-versus-leukemia effect, the relapse rates remain similar (to BuCy and CyTBI), while toxicity and subsequent NRM are substantially lower. As is common for high-dose chemotherapeutic agents, the therapeutic windows for both drugs are narrow. Minimal exposure of both agents is essential to prevent graft failure (Bu and Flu) and relapse (Bu).
Indeed, the outcome of an HCT seems highly influenced by Bu exposure, making optimal Bu dosing warranted. Owing to the very small therapeutic window (±10%) relative to unpredicted inter-individual variability in Bu clearance (20%), *a priori* dosing adjustments lead to high proportion of under- and over-exposed patients. TDM vastly increases target attainment and has been proven to lead to increased survival.
The evidence regarding Flu optimal exposure is not yet equally strong as

for Bu, though similarly steep exposure-response relationships have been observed, warranting similarly strict dosing. Implementation of Flu dosing individualization needs to be confirmed in a prospective trial.
With all development in transplantation medicine and cellular therapies, a predictable framework surrounding the key components is of utmost importance. With conditioning being the starting point of these therapies, this needs to be optimal for each individual, in order to better study the efficacy of cellular therapy itself. BuFlu has potent synergistic activity and both agents can be readily tuned to the individual. This makes BuFlu a potent duo adjustable to precision.

Table 1: Studies comparing oral busulfan-based regimens to total body irradiation-based regimens*

Author	Year	Design	N	Age group	Comparison	Indication	Survival	Author conclusions
Sakellari et al.	2018	Retrospective cohort 1990-2016	67 vs 84	All ages	POBuCy vs CyTBI	ALL	Similar	Similar outcomes in general for both regimens. TBI superior in patients younger than 40 years of age.
Lucchini et al.	2017	Registry based 2000-2010	389 vs 109 vs 133	Pediatric	PO/IVBuCy vs CyTBI vs POBuCyMel	AML	Similar	Similar outcomes for BuCy and CyTBI. BuCyMel had superior outcomes due to less relapses.
Beranger et al.	2014	Registry based: 1980-2004	142 vs 84	Pediatric	PO/IVBuCy vs CyTBI	AML	Similar	Superior outcomes BuCy, due to decreased NRM, cGvHD and aGvHD
Kalaycio et al.	2011	Retrospective cohort 1989-2008	51 vs 35	All ages	BuCy* vs CyEto	ALL	Similar	Superior outcomes CyTBI due to significantly higher DFS and non-significantly higher OS
Bunin et al	2003	RCT	21 vs 22	Pediatric	POBuCyEto vs EtoCyTBI	ALL	TBI-based	Superior survival for CyTBI (EFS: 58% vs 29%) due to decreased NRM and similar relapse
Litzow et al.	2002	Registry based 1988-1996	381 vs 200	Adults	POBuCy vs CyTBI	AML	Similar	Equal survival for both regimens, but increased relapse associated with BuCy.
Kim et al	2001	Retrospective cohort 1990-1997	27 vs 26	Adults	POBuCy vs CyTBI	CML	Similar	No statistical difference in long-term outcomes, though a (non-significant) trend for reduced relapse rate in BuCy group was found
Socié et al.	2001	Pooled RCT's 1987-1992	AML: 92 vs 80 CML: 168 vs 148	Adults	POBuCy vs CyTBI	Stratified: AML /CML	Similar	No statistical difference in long-term outcomes, though a (non-significant) trend for improved survival in AML patients after CyTBI was found
Davies et al.	2000	Registry based 1988-1995	176 vs 451	Pediatric	POBuCy vs CyTBI	ALL	TBI-based	Superior survival for CyTBI due to decreased NRM (relative risk 0.59) and similar relapse
Ringden et al	1999	RCT 1997-2000	88 vs 79	Adults	POBuCy vs CyTBI	ALL/AML/CML	Similar	Increased toxicity in Bu-based (VOD, cystitis) and increased cGvHD. Long-term NRM was higher as well: 21/64% for BuCy vs 12/22% for CyTBI in early/ advanced disease.
Devergie et al.	1995	RCT 1988-1991	55 vs 55	Adults	POBuCy vs CyTBI	CML	Similar	Similar survival and transplant-related complications for BuCy and CyTBI, but a trend towards more anti-leukemic activity in BuCy leading to less relapses.
Ringden et al	1994	RCT 1988-1992	88 vs 79	Adults	POBuCy vs CyTBI	ALL/AML/CMl	TBI-based	NRM was increased in BuCy group (23% vs 8%), especially for those with advanced disease. TBI is regimen of choice, but BuCy can be alternative for early disease.
Clift et al.	1994	RCT 1988-1992	73 vs 69	Adults	POBuCy vs CyTBI	CML	Similar	Equal survival for both regimens, but increased morbidity (i.e. Renal dysfunction, GvHD, infections) in TBI-group.
Blume et al.	1993	RCT 1987-1991	59 vs 55	All ages	POBuCy vs EtoTBI	ALL/AML/CML	Similar	No conclusions can be drawn, regarding superiority of any regimen.
Blaise et al.	1992	RCT 1987-1990	51 vs 50	Adults	POBuCy vs CyTBI	AML	TBI-based	TBI-based regimen gives superior survival (75% vs 51%), due to decreased NRM- (8% vs 23%) and relapse probability (8% vs 27%)

*RCT: randomized controlled trial; PO: oral; IV: intravenous; Bu: busulfan; Cy: cyclophosphamide; TBI: total body irradiation; Eto: etoposide; Mel: melphalan; ALL: acute lymphocytic leukemia; AML: acute myeloid leukemia; CML: chronic myeloid leukemia; NRM: non-relapse mortality; cGvHD: chronic graft-versus-host-disease; aGvHD: acute graft-versus-host-disease

Table 2: Studies comparing intravenous busulfan-based regimens to total body irradiation-based regimens*

Author	Year	Design	N	Age group	Comparison	Indication	Survival	Author conclusions
Gooptu et al.	2018	Retrospective cohort 2010-2015	158 vs 229	Adults	IVBuFlu vs CyTBI	Malignancies	Bu-based	Superior survival BuFlu (65 vs 55%) due to less NRM (HR 0.18) and comparable relapse
Kibriaei et al	2018	Registry based 2005-2014	299 vs 819	Adults	Bu-based (+Flu/ Cy/Clo/Mel) vs TBI-based (+Cy/Eto)	ALL	Similar	Similar survival for both regimens. A trend towards lower NRM in Bu-based and significantly lower relapse in TBI-based regimens.
Mitsuhashi et al.	2016	Registry based 2000-2012	42 vs 2028 vs 60	Adults	IVBuCy vs CyTBI vs POBuCy	ALL	Similar	Similar outcomes BuCy CyTBI, when IVBu is used
Copelan et al.	2015	Registry based 2000-2006	97 vs 222 vs 354	Adults	IVBuCy vs CyTBI vs POBuCy	CML	Similar	Less relapses in IVBuCY (HR 0.38) and superior LFS (HR 0.55) among MSD recipients compared to CyTBI
Ishida	2015	Registry based 2006-2011	69 vs 151	Pediatric	IVBuCy vs CyTBI	AML	Similar	Outcomes for BuCy and CyTBI are similar.
Nagler et al.	2013	Registry based 2004-2010	795 vs 864	Adults	IVBuCy vs CyTBI	AML	Similar	Comparable OS through balance in decreased NRM (15 to 12%) and increased Relapse (21 to 26%) in IVBuCy. aGvHD and cGvHD were both decreased in IVBuCy
Bredeson et al.	2013	Prospective cohort 2009-2011	1025 vs 458	All ages	Bu-based (+Flu/ Cy) vs TBI-based (+Cy/ Eto)	AML/CML/ MDS	Bu-based	Superior survival for Bu-based regimens (relative risk 0.82) due to decreased relapse rate (39 to 34%), especially in AML and CML
Eroglu et al.	2013	Retrospective 2000-2011	50 vs 45	All ages	IVBuCy vs CyTBI	ALL	TBI-based	Superior outcomes CyTBI due to significantly longer median DFS (13 vs 4 months) and OS (37 vs 12 months), resulting from a decreased relapse rate (51 vs 75%).
Sisler et al.	2009	Registry based 2000-2008	61 vs 90	Pediatric	IV/PO BuCy vs CyTBI	AML (beyond CR1)	Similar	Similar survival for both regimens, but increased morbidity (i.e. Renal dysfunction, GvHD, infections) in TBI-group.

*PO: oral; IV: intravenous; Bu: busulfan; Flu: fludarabine; Cy: cyclophosphamide; TBI: total body irradiation; Clo: clofarabine; Eto: etoposide; Mel: melphalan; ALL: acute lymphocytic leukemia; AML: acute myeloid leukemia; CML: chronic myeloid leukemia; MDS: myelodysplastic syndrome; CR1: first complete remission; NRM: non-relapse mortality; LFS: leukemia-free survival; MSD: matched sibling donor; cGvHD: chronic graft-versus-host-disease; aGvHD: acute graft-versus-host-disease; HR: hazard ratio

Table 3: Studies comparing busulfan + fludarabine to busulfan + cyclophosphamide*

Author	Year	Design	N	Age group	Comparison***	Indication	Survival	Author conclusions
Harris et al.	2018	Registry-based 2008-2014	381 vs 1400	Pediatrics	BuFlu vs BuCy +stratified for malignant/ benign	Malignant and benign	Similar	Similar survival for both regimens with reduced toxicity in BuFlu (especially in Benign disorders). TRM similar in Maligant & benign. Equal relapse in maligancies.
Pasquini et al.	2016	Registry-based 2009-20	91 vs 331 vs 495 vs 96	All ages	BuFluQ6 vs BuFluQ24 vs BuCy Q6 vs BuCy Q24	AML/MDS/ CML	Similar	There is no difference between dosing Bu each 6 or 24 hours, neither between adding Cy or Flu.
Rambaldi et al.	2015	RCT 2008-2012	127 vs 125	40-65 years	BuFlu160** vs BuCy	AML	Similar	Improved OS (HR 0.69 , n.s.) due to decreased NRM (HR 0.42, p=0.02) in the BuFlu group. NRM difference mainly due to organ toxicity in BuCy.
Fedele et al.	2014	Retrospective cohort: 1993-2009	17 vs 48	Adult	BuFlu120** vs BuCy	AML	Similar	Similar outcomes with a trend to improved outcome in high-risk patients when given BuCy.
Bartelink et al.	2014	Retrospective cohort: 2009-2012	64 vs 50	Pediatrics	T90BuFlu160** vs T80BuCy	Malignant and benign	Similar	Similar survival outcomes. Increased toxicity (GvHD, IPS, VOD)
Lie et al.	2013	RCT 2009-2012	52 vs 53	14-55 years	mBuFlu150** vs mBuCy	AML/MDS/ CML/ALL	Similar	Similar survival for both regimens. More pneunomia observed in BuFlu arm.
Liu et al.	2013	RCT 2007-2012	54 vs 54	12-60 years	BuFlu150** vs BuCy	AML	Similar	Similar survival for both regimens. Trend towards decreased NRM in BuFlu (19% to 10%).
Lee et al.	2013	RCT 2005-2009	62 vs 64	15-70 years	BuFlu150** vs BuCy	AML/MDS/ CML/ALL	BuCy	Replacing Cy by Flu causes increased graft failure with subsequent NRM (14 to 24%) and relapse rate (19 to 34%)
Lee et al.	2010	Retrospectivve cohort: 1991-2009	17 vs 25	Adults	BuFlu150** vs BuCy	AML/ALL/MDS	Similar	Similar outcomes for both regimens with less mucositis in BuFlu.
Andersson et al.	2008	Retrospective cohort 1997-2005	148 vs 67	Adults	BuFlu160** vs BuCy	AML/MDS	BuFlu	Improved outcome in the BuFlu group, with higher OS (51% vs 42% in BuCy) due to decreased NRM probability (12% vs 27% in BuCy). Especially young patients and those transplanted in CR1 benefitted from BuFlu
Bredeson et al.	2008	Registry-based matched case-control study 1999-2003	120 vs 215	Adults	IVBuFlu250** vs POBuCy	Malignancies	BuFlu	Slightly superior OS in BuFlu (58% vs 51% in BuCy) through reduced NRM (34 to 12%), but increased disease relapse (20 vs 42%).
Chae et al.	2007	Retrospective cohort 1998-2004	40 vs 55	Adults	IVBuFlu180** vs POBuCy	Malignancies	BuFlu	BuFlu is the superior regimen with increased OS (HR 0.168) due to lower NRM (reduced from 30 to 10%)

*RCT: randomized controlled trial; PO: oral; IV: intravenous; Bu: busulfan; Flu: fludarabine; Cy: cyclophosphamide; Q6: administered every 6 hours; Q24: administered every 24 hours; T80: targeted to 4-day busulfan exposure of 80 mg*h/L; T90: targeted to 4-day busulfan exposure of 90 mg*h/L; m: modified; ALL: acute lymphocytic leukemia; AML: acute myeloid leukemia; CML: chronic myeloid leukemia; MDS: myelodysplastic syndrome; CR1: first complete remission; OS: overall survival; NRM: non-relapse mortality; GvHD: graft-versus-host-disease; IPS: idiopathic pulmonary syndrome; VOD: veno-occlusive disease
**Number indicates the fludarabine dose per m2 of body-surface-area
***Busulfan is used in IV formulation, unless otherwise specified

Table 4: Studies comparing busulfan + fludarabine to busulfan + cyclophosphamide*

Author	Year	N	Age group	Regimen	Diagnosis	Design & objective	Pharmacokinetic quantification	Target
Andersson et al.	2017	107 vs 111	13-66 years	BuFlu	AML & MDS	Randomized controlled trial to compare fixed dosing to targeted dosing	Parametric non population-based.	Pre-defined target exposure: 97 $mg \cdot h/L \equiv$ 6000 $\mu Mol \cdot min$
Bartelink et al	2016	674	0-22 years	BuFlu /BuCy /BuCyMel	Maligant & Benign	Retrospective multicentre to establish an optimal exposure	Parametric population-based	Exposure associated with minimal event probablity: 90 $(78-101)$ $mg \cdot h/L \equiv$ 5488 $(4756-6158)$ $\mu Mol \cdot min$
Russel et al	2013	158	Adults	BuFlu	Malignancies	Retrospective study of ¡6000 uM*min targeted population to find an optimal exposure	Non-parametric: trapezoidal	3rd quartile: 71 $(61-81)$ $mg \cdot h/L \equiv$ 4404 $(3814-4994)$ $\mu Mol \cdot min$
Geddes et al	2007	130	Adults	BuFlu	Malignancies	Post-hoc analysis of a prospective study to determine the need for TDM	Non-parametric: trapezoidal	$<$ 97 $mg \cdot h/L \equiv$ 6000 $\mu Mol \cdot min$
Andersson et al.	2002	36	Adults	BuCy	CML	Post-hoc analysis of a prospective study to find an optimal exposure	Parametric non population-based.	80 $(62-99)$ $mg \cdot h/L \equiv$ 4940 $(3800-6080)$ $\mu Mol \cdot min$
Slattery et al	1997	45	Adults	POBuCy	CML	Post-hoc analysis of a prospective study to find an exposure-response relationship	Non-parametric: mean of observed concentrations	$>$ 87 $mg \cdot h/L \equiv$ 5382 $\mu Mol \cdot min$

* PO: oral; Bu: busulfan; Flu: fludarabine; Cy: cyclophosphamide; Mel: melphalan; AML: acute myeloid leukemia; CML: chronic myeloid leukemia; MDS: myelodysplastic syndrome

References

1. Eapen, M., Brazauskas, R., Hemmer, M., Perez, W. S., Steinert, P., Horowitz, M. M., *et al.* Hematopoietic cell transplant for acute myeloid leukemia and myelodysplastic syndrome: conditioning regimen intensity. *Blood Adv* **2,** 2095–2103 (2018).
2. Santos, G. W., Tutschka, P. J., Brookmeyer, R., Saral, R., Beschorner, W. E., Bias, W. B., *et al.* Marrow transplantation for acute nonlymphocytic leukemia after treatment with busulfan and cyclophosphamide. *N Engl J Med* **309,** 1347–53 (1983).
3. Nevill, T. J., Barnett, M. J., Klingemann, H. G., Reece, D. E., Shepherd, J. D. & Phillips, G. L. Regimen-related toxicity of a busulfan-cyclophosphamide conditioning regimen in 70 patients undergoing allogeneic bone marrow transplantation. *J Clin Oncol* **9,** 1224–32 (1991).
4. Clift, R. A., Buckner, C. D., Thomas, E. D., Bensinger, W. I., Bowden, R., Bryant, E., *et al.* Marrow transplantation for chronic myeloid leukemia: a randomized study comparing cyclophosphamide and total body irradiation with busulfan and cyclophosphamide. *Blood* **84,** 2036–43 (1994).
5. Devergie, A., Blaise, D., Attal, M., Tigaud, J. D., Jouet, J. P., Vernant, J. P., *et al.* Allogeneic bone marrow transplantation for chronic myeloid leukemia in first chronic phase: a randomized trial of busulfan-cytoxan versus cytoxan-total body irradiation as preparative regimen: a report from the French Society of Bone Marrow Graft (SFGM). *Blood* **85,** 2263–8 (1995).
6. Blaise, D., Maraninchi, D., Archimbaud, E., Reiffers, J., Devergie, A., Jouet, J. P., *et al.* Allogeneic bone marrow transplantation for acute myeloid leukemia in first remission: a randomized trial of a busulfan-Cytoxan versus Cytoxan-total body irradiation as preparative regimen: a report from the Group d'Etudes de la Greffe de Moelle Osseuse. *Blood* **79,** 2578–82 (1992).
7. Ringden, O., Ruutu, T., Remberger, M., Nikoskelainen, J., Volin, L., Vindelov, L., *et al.* A randomized trial comparing busulfan with total body irradiation as conditioning in allogeneic marrow transplant recipients with leukemia: a report from the Nordic Bone Marrow Transplantation Group. *Blood* **83,** 2723–30 (1994).

8. Ringden, O., Remberger, M., Ruutu, T., Nikoskelainen, J., Volin, L., Vindelov, L., *et al.* Increased risk of chronic graft-versus-host disease, obstructive bronchiolitis, and alopecia with busulfan versus total body irradiation: long-term results of a randomized trial in allogeneic marrow recipients with leukemia. Nordic Bone Marrow Transplantation Group. *Blood* **93,** 2196–201 (1999).

9. Socie, G., Clift, R. A., Blaise, D., Devergie, A., Ringden, O., Martin, P. J., *et al.* Busulfan plus cyclophosphamide compared with total-body irradiation plus cyclophosphamide before marrow transplantation for myeloid leukemia: long-term follow-up of 4 randomized studies. *Blood* **98,** 3569–74 (2001).

10. Davies, S. M., Ramsay, N. K., Klein, J. P., Weisdorf, D. J., Bolwell, B., Cahn, J. Y., *et al.* Comparison of preparative regimens in transplants for children with acute lymphoblastic leukemia. *J Clin Oncol* **18,** 340–7 (2000).

11. Andersson, B. S., Madden, T., Tran, H. T., Hu, W. W., Blume, K. G., Chow, D. S., *et al.* Acute safety and pharmacokinetics of intravenous busulfan when used with oral busulfan and cyclophosphamide as pre-transplantation conditioning therapy: a phase I study. *Biol Blood Marrow Transplant* **6,** 548–54 (2000).

12. Kashyap, A., Wingard, J., Cagnoni, P., Roy, J., Tarantolo, S., Hu, W., *et al.* Intravenous versus oral busulfan as part of a busulfan/cyclophosphamide preparative regimen for allogeneic hematopoietic stem cell transplantation: decreased incidence of hepatic venoocclusive disease (HVOD), HVOD-related mortality, and overall 100-day mortality. *Biol Blood Marrow Transplant* **8,** 493–500 (2002).

13. De Berranger, E., Cousien, A., Petit, A., Peffault de Latour, R., Galambrun, C., Bertrand, Y., *et al.* Impact on long-term OS of conditioning regimen in allogeneic BMT for children with AML in first CR: TBI+CY versus BU+CY: a report from the Societe Francaise de Greffe de Moelle et de Therapie Cellulaire. *Bone Marrow Transplant* **49,** 382–8 (2014).

14. Eroglu, C., Pala, C., Kaynar, L., Yaray, K., Aksozen, M. T., Bankir, M., *et al.* Comparison of total body irradiation plus cyclophosphamide with busulfan plus cyclophosphamide as conditioning regimens in patients with acute lymphoblastic leukemia undergoing allogeneic hematopoietic stem cell transplant. *Leuk Lymphoma* **54,** 2474–9 (2013).

15. Kalaycio, M., Bolwell, B., Rybicki, L., Absi, A., Andresen, S., Pohlman, B., *et al.* BU- vs TBI-based conditioning for adult patients with ALL. *Bone Marrow Transplant* **46,** 1413–7 (2011).

16. Mitsuhashi, K., Kako, S., Shigematsu, A., Atsuta, Y., Doki, N., Fukuda, T., *et al.* Comparison of Cyclophosphamide Combined with Total Body Irradiation, Oral Busulfan, or Intravenous Busulfan for Allogeneic Hematopoietic Cell Transplantation in Adults with Acute Lymphoblastic Leukemia. *Biol Blood Marrow Transplant* **22,** 2194–2200 (2016).

17. Sakellari, I., Gavriilaki, E., Chatziioannou, K., Papathanasiou, M., Mallouri, D., Batsis, I., *et al.* Long-term outcomes of total body irradiation plus cyclophosphamide versus busulfan plus cyclophosphamide as conditioning regimen for acute lymphoblastic leukemia: a comparative study. *Ann Hematol* **97,** 1987–1994 (2018).

18. Copelan, E. A., Avalos, B. R., Ahn, K. W., Zhu, X., Gale, R. P., Grunwald, M. R., *et al.* Comparison of outcomes of allogeneic transplantation for chronic myeloid leukemia with cyclophosphamide in combination with intravenous busulfan, oral busulfan, or total body irradiation. *Biol Blood Marrow Transplant* **21,** 552–8 (2015).

19. Kebriaei, P., Anasetti, C., Zhang, M. J., Wang, H. L., Aldoss, I., de Lima, M., *et al.* Intravenous Busulfan Compared with Total Body Irradiation Pretransplant Conditioning for Adults with Acute Lymphoblastic Leukemia. *Biol Blood Marrow Transplant* **24,** 726–733 (2018).

20. Nagler, A., Rocha, V., Labopin, M., Unal, A., Ben Othman, T., Campos, A., *et al.* Allogeneic hematopoietic stem-cell transplantation for acute myeloid leukemia in remission: comparison of intravenous busulfan plus cyclophosphamide (Cy) versus total-body irradiation plus Cy as conditioning regimen–a report from the acute leukemia working party of the European group for blood and marrow transplantation. *J Clin Oncol* **31,** 3549–56 (2013).

21. Gooptu, M., Kim, H. T., Ho, V. T., Alyea, E. P., Koreth, J., Armand, P., *et al.* A Comparison of the Myeloablative Conditioning Regimen Fludarabine/Busulfan with Cyclophosphamide/Total Body Irradiation, for Allogeneic Stem Cell Transplantation in the Modern Era: A Cohort Analysis. *Biol Blood Marrow Transplant* **24,** 1733–1740 (2018).

22. Bredeson, C., LeRademacher, J., Kato, K., Dipersio, J. F., Agura, E., Devine, S. M., *et al.* Prospective cohort study comparing intravenous busulfan to total body irradiation in hematopoietic cell transplantation. *Blood* **122,** 3871–8 (2013).

23. Vaxman, I., Ram, R., Gafter-Gvili, A., Vidal, L., Yeshurun, M., Lahav, M., *et al.* Secondary malignancies following high dose therapy and autologous hematopoietic cell transplantation-systematic review and meta-analysis. *Bone Marrow Transplant* **50,** 706–14 (2015).

24. Jankovic, M., Brouwers, P., Valsecchi, M. G., Van Veldhuizen, A., Huisman, J., Kamphuis, R., *et al.* Association of 1800 cGy cranial irradiation with intellectual function in children with acute lymphoblastic leukaemia. ISPACC. International Study Group on Psychosocial Aspects of Childhood Cancer. *Lancet* **344,** 224–7 (1994).

25. Krull, K. R., Zhang, N., Santucci, A., Srivastava, D. K., Krasin, M. J., Kun, L. E., *et al.* Long-term decline in intelligence among adult survivors of childhood acute lymphoblastic leukemia treated with cranial radiation. *Blood* **122,** 550–3 (2013).

26. Phipps, S., Dunavant, M., Srivastava, D. K., Bowman, L. & Mulhern, R. K. Cognitive and academic functioning in survivors of pediatric bone marrow transplantation. *J Clin Oncol* **18,** 1004–11 (2000).

27. Ishida, H., Kato, M., Kudo, K., Taga, T., Tomizawa, D., Miyamura, T., *et al.* Comparison of Outcomes for Pediatric Patients With Acute Myeloid Leukemia in Remission and Undergoing Allogeneic Hematopoietic Cell Transplantation With Myeloablative Conditioning Regimens Based on Either Intravenous Busulfan or Total Body Irradiation: A Report From the Japanese Society for Hematopoietic Cell Transplantation. *Biol Blood Marrow Transplant* **21,** 2141–2147 (2015).

28. Sisler, I. Y., Koehler, E., Koyama, T., Domm, J. A., Ryan, R., Levine, J. E., *et al.* Impact of conditioning regimen in allogeneic hematopoetic stem cell transplantation for children with acute myelogenous leukemia beyond first complete remission: a pediatric blood and marrow transplant consortium (PBMTC) study. *Biol Blood Marrow Transplant* **15,** 1620–7 (2009).

29. Hassan, Z., Hellstrom-Lindberg, E., Alsadi, S., Edgren, M., Hagglund, H. & Hassan, M. The effect of modulation of glutathione cellular content on busulphan-induced cytotoxicity on hematopoietic cells in vitro and in vivo. *Bone Marrow Transplant* **30,** 141–7 (2002).

30. Milford, D. V., Butler, N., Clarke, R. & Vaid, J. Reversible hepatic dysfunction in association with cyclophosphamide therapy. *Eur J Pediatr* **154,** 411–2 (1995).

31. Khouri, I. F., Keating, M., Korbling, M., Przepiorka, D., Anderlini, P., O'Brien, S., *et al.* Transplant-lite: induction of graft-versus-malignancy using fludarabine-based nonablative chemotherapy and allogeneic blood progenitor-cell transplantation as treatment for lymphoid malignancies. *J Clin Oncol* **16,** 2817–24 (1998).

32. O'Brien, S., Kantarjian, H., Beran, M., Smith, T., Koller, C., Estey, E., *et al.* Results of fludarabine and prednisone therapy in 264 patients with chronic lymphocytic leukemia with multivariate analysis-derived prognostic model for response to treatment. *Blood* **82,** 1695–700 (1993).

33. Bredeson, C. N., Zhang, M. J., Agovi, M. A., Bacigalupo, A., Bahlis, N. J., Ballen, K., *et al.* Outcomes following HSCT using fludarabine, busulfan, and thymoglobulin: a matched comparison to allogeneic transplants conditioned with busulfan and cyclophosphamide. *Biol Blood Marrow Transplant* **14,** 993–1003 (2008).

34. Chae, Y. S., Sohn, S. K., Kim, J. G., Cho, Y. Y., Moon, J. H., Shin, H. J., *et al.* New myeloablative conditioning regimen with fludarabine and busulfan for allogeneic stem cell transplantation: comparison with BuCy2. *Bone Marrow Transplant* **40,** 541–7 (2007).

35. Fedele, R., Messina, G., Martinello, T., Gallo, G. A., Pontari, A., Moscato, T., *et al.* Tolerability and efficacy of busulfan and fludarabine as allogeneic pretransplant conditioning therapy in acute myeloid leukemia: comparison with busulfan and cyclophosphamide regimen. *Clin Lymphoma Myeloma Leuk* **14,** 493–500 (2014).

36. Lee, J. H., Choi, J., Kwon, K. A., Lee, S., Oh, S. Y., Kwon, H. C., *et al.* Fludarabine based myeloablative regimen as pretransplant conditioning therapy in adult acute leukemia/myelodysplastic syndrome: comparison with oral or intravenous busulfan with cyclophosphamide. *Korean J Hematol* **45,** 102–8 (2010).

37. Andersson, B. S., de Lima, M., Thall, P. F., Wang, X., Couriel, D., Korbling, M., *et al.* Once daily i.v. busulfan and fludarabine (i.v. BuFlu) compares favorably with i.v. busulfan and cyclophosphamide (i.v. BuCy2) as pretransplant conditioning therapy in AML/MDS. *Biol Blood Marrow Transplant* **14,** 672–84 (2008).

38. Lee, J. H., Joo, Y. D., Kim, H., Ryoo, H. M., Kim, M. K., Lee, G. W., *et al.* Randomized trial of myeloablative conditioning regimens: busulfan plus cyclophosphamide versus busulfan plus fludarabine. *J Clin Oncol* **31,** 701–9 (2013).

39. Liu, H., Zhai, X., Song, Z., Sun, J., Xiao, Y., Nie, D., *et al.* Busulfan plus fludarabine as a myeloablative conditioning regimen compared with busulfan plus cyclophosphamide for acute myeloid leukemia in first complete remission undergoing allogeneic hematopoietic stem cell transplantation: a prospective and multicenter study. *J Hematol Oncol* **6,** 15 (2013).

40. Rambaldi, A., Grassi, A., Masciulli, A., Boschini, C., Mico, M. C., Busca, A., *et al.* Busulfan plus cyclophosphamide versus busulfan plus fludarabine as a preparative regimen for allogeneic haemopoietic stem-cell transplantation in patients with acute myeloid leukaemia: an open-label, multicentre, randomised, phase 3 trial. *Lancet Oncol* **16,** 1525–1536 (2015).

41. Valdez, B. C., Li, Y., Murray, D., Champlin, R. E. & Andersson, B. S. The synergistic cytotoxicity of clofarabine, fludarabine and busulfan in AML cells involves ATM pathway activation and chromatin remodeling. *Biochem Pharmacol* **81,** 222–32 (2011).

42. Bartelink, I. H., van Reij, E. M., Gerhardt, C. E., van Maarseveen, E. M., de Wildt, A., Versluys, B., *et al.* Fludarabine and exposure targeted busulfan compares favorably with busulfan/cyclophosphamide based regimens in pediatric hematopoietic cell transplantation: maintaining efficacy with less toxicity. *Biol Blood Marrow Transplant* **20,** 345–53 (2014).

43. Harris, A. C., Boelens, J. J., Ahn, K. W., Fei, M., Abraham, A., Artz, A., *et al.* Comparison of pediatric allogeneic transplant outcomes using myeloablative busulfan with cyclophosphamide or fludarabine. *Blood Adv* **2,** 1198–1206 (2018).

44. Pasquini, M. C., Le-Rademacher, J., Zhu, X., Artz, A., DiPersio, J., Fernandez, H. F., *et al.* Intravenous Busulfan-Based Myeloablative Conditioning Regimens Prior to Hematopoietic Cell Transplantation for Hematologic Malignancies. *Biol Blood Marrow Transplant* **22,** 1424–30 (2016).

45. Kunkele, A., Engelhard, M., Hauffa, B. P., Mellies, U., Muntjes, C., Huer, C., *et al.* Long-term follow-up of pediatric patients receiving total body irradiation before hematopoietic stem cell transplantation and post-transplant survival of ¿2 years. *Pediatr Blood Cancer* **60,** 1792–7 (2013).

46. Thomas, O., Mahe, M., Campion, L., Bourdin, S., Milpied, N., Brunet, G., *et al.* Long-term complications of total body irradiation in adults. *Int J Radiat Oncol Biol Phys* **49,** 125–31 (2001).

47. *BUSULFEX (busulfan) for injection FDA Label* `https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/020954s014lbl.pdf`. Accessed: 2018-09-21.

48. Myers, A. L., Kawedia, J. D., Champlin, R. E., Kramer, M. A., Nieto, Y., Ghose, R., *et al.* Clarifying busulfan metabolism and drug interactions to support new therapeutic drug monitoring strategies: a comprehensive review. *Expert Opin Drug Metab Toxicol* **13,** 901–923 (2017).

49. Hassan, M., Oberg, G., Ehrsson, H., Ehrnebo, M., Wallin, I., Smedmyr, B., *et al.* Pharmacokinetic and metabolic studies of high-dose busulphan in adults. *Eur J Clin Pharmacol* **36,** 525–30 (1989).

50. Hassan, M., Ehrsson, H. & Ljungman, P. Aspects concerning busulfan pharmacokinetics and bioavailability. *Leuk Lymphoma* **22,** 395–407 (1996).

51. Andersson, B. S., Thall, P. F., Madden, T., Couriel, D., Wang, X., Tran, H. T., *et al.* Busulfan systemic exposure relative to regimen-related toxicity and acute graft-versus-host disease: defining a therapeutic window for i.v. BuCy2 in chronic myelogenous leukemia. *Biol Blood Marrow Transplant* **8,** 477–85 (2002).

52. Bartelink, I. H., Lalmohamed, A., Reij, E. M. L. v., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *The Lancet Haematology* **3,** e526–e536 (2016).

53. Geddes, M., Kangarloo, S. B., Naveed, F., Quinlan, D., Chaudhry, M. A., Stewart, D., *et al.* High busulfan exposure is associated with worse outcomes in a daily i.v. busulfan and fludarabine allogeneic transplant regimen. *Biol Blood Marrow Transplant* **14,** 220–8 (2008).

54. Russell, J. A., Kangarloo, S. B., Williamson, T., Chaudhry, M. A., Savoie, M. L., Turner, A. R., *et al.* Establishing a target exposure for once-daily intravenous busulfan given with fludarabine and thymoglobulin before allogeneic transplantation. *Biol Blood Marrow Transplant* **19,** 1381–6 (2013).

55. Bartelink, I. H., Lalmohamed, A., van Reij, E. M., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *Lancet Haematol* **3,** e526–e536 (2016).

56. Slattery, J. T., Clift, R. A., Buckner, C. D., Radich, J., Storer, B., Bensinger, W. I., *et al.* Marrow transplantation for chronic myeloid leukemia: the influence of plasma busulfan levels on the outcome of transplantation. *Blood* **89,** 3055–60 (1997).

57. Andersson, B. S., Thall, P. F., Valdez, B. C., Milton, D. R., Al-Atrash, G., Chen, J., *et al.* Fludarabine with pharmacokinetically guided IV busulfan is superior to fixed-dose delivery in pretransplant conditioning of AML/MDS patients. *Bone Marrow Transplant* **52,** 580–587 (2017).

58. *Busilvex 6 mg/ml concentrate for solution for infusion Summary of Product Characteristics (SmPC)* `http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000472/WC500052066.pdf`. Accessed: 2018-09-21.

59. McCune, J. S., Bemer, M. J., Barrett, J. S., Scott Baker, K., Gamis, A. S. & Holford, N. H. Busulfan in infant to adult hematopoietic cell transplant recipients: a population pharmacokinetic model for initial and Bayesian dose personalization. *Clin Cancer Res* **20,** 754–63 (2014).

60. Ansari, M. & Krajinovic, M. Can the pharmacogenetics of GST gene polymorphisms predict the dose of busulfan in pediatric hematopoietic stem cell transplantation? *Pharmacogenomics* **10,** 1729–32 (2009).

61. Ansari, M., Lauzon-Joset, J. F., Vachon, M. F., Duval, M., Theoret, Y., Champagne, M. A., *et al.* Influence of GST gene polymorphisms on busulfan pharmacokinetics in children. *Bone Marrow Transplant* **45,** 261–7 (2010).

62. Choi, B., Kim, M. G., Han, N., Kim, T., Ji, E., Park, S., *et al.* Population pharmacokinetics and pharmacodynamics of busulfan with GSTA1 polymorphisms in patients undergoing allogeneic hematopoietic stem cell transplantation. *Pharmacogenomics* **16,** 1585–94 (2015).

63. Kim, S. D., Lee, J. H., Hur, E. H., Lee, J. H., Kim, D. Y., Lim, S. N., *et al.* Influence of GST gene polymorphisms on the clearance of intravenous busulfan in adult patients undergoing hematopoietic cell transplantation. *Biol Blood Marrow Transplant* **17,** 1222–30 (2011).

64. Yin, J., Xiao, Y., Zheng, H. & Zhang, Y. C. Once-daily i.v. BU-based conditioning regimen before allogeneic hematopoietic SCT: a study of influence of GST gene polymorphisms on BU pharmacokinetics and clinical outcomes in Chinese patients. *Bone Marrow Transplant* **50,** 696–705 (2015).

65. Zwaveling, J., Press, R. R., Bredius, R. G., van Derstraaten, T. R., den Hartigh, J., Bartelink, I. H., *et al.* Glutathione S-transferase polymorphisms are not associated with population pharmacokinetic parameters of busulfan in pediatric patients. *Ther Drug Monit* **30,** 504–10 (2008).

66. Philippe, M., Neely, M., Bertrand, Y., Bleyzac, N. & Goutelle, S. A Nonparametric Method to Optimize Initial Drug Dosing and Attainment of a Target Exposure Interval: Concepts and Application to Busulfan in Pediatrics. *Clin Pharmacokinet* **56,** 435–447 (2017).

67. Plunkett, W., Huang, P. & Gandhi, V. Metabolism and action of fludarabine phosphate. *Semin Oncol* **17,** 3–17 (1990).

68. Consoli, U., El-Tounsi, I., Sandoval, A., Snell, V., Kleine, H. D., Brown, W., *et al.* Differential induction of apoptosis by fludarabine monophosphate in leukemic B and normal T cells in chronic lymphocytic leukemia. *Blood* **91,** 1742–8 (1998).

69. McCune, J. S., Vicini, P., Salinger, D. H., O'Donnell, P. V., Sandmaier, B. M., Anasetti, C., *et al.* Population pharmacokinetic/dynamic model of lymphosuppression after fludarabine administration. *Cancer Chemother Pharmacol* **75,** 67–75 (2015).

70. Woodahl, E. L., Wang, J., Heimfeld, S., Sandmaier, B. M., O'Donnell, P. V., Phillips, B., *et al.* A novel phenotypic method to determine fludarabine triphosphate accumulation in T-lymphocytes from hematopoietic cell transplantation patients. *Cancer Chemother Pharmacol* **63,** 391–401 (2009).

71. McCune, J. S., Mager, D. E., Bemer, M. J., Sandmaier, B. M., Storer, B. E. & Heimfeld, S. Association of fludarabine pharmacokinetic/dynamic biomarkers with donor chimerism in nonmyeloablative HCT recipients. *Cancer Chemother Pharmacol* **76,** 85–96 (2015).

72. Ivaturi, V., Dvorak, C. C., Chan, D., Liu, T., Cowan, M. J., Wahlstrom, J., *et al.* Pharmacokinetics and Model-Based Dosing to Optimize Fludarabine Therapy in Pediatric Hematopoietic Cell Transplant Recipients. *Biol Blood Marrow Transplant* **23,** 1701–1713 (2017).

73. Gandhi, V. & Plunkett, W. Cellular and clinical pharmacology of fludarabine. *Clin Pharmacokinet* **41,** 93–103 (2002).

74. Langenhorst, J. B., van Kesteren, C., van Maarseveen, E. M., Dorlo, T. P. C., Nierkens, S., Lindemans, C. A., *et al. Fludarabine exposure in the conditioning prior to allogeneic hematopoietic cell transplantation predicts outcomes* Submitted. Dec. 2018.

75. Long-Boyle, J. R., Green, K. G., Brunstein, C. G., Cao, Q., Rogosheske, J., Weisdorf, D. J., *et al.* High fludarabine exposure and relationship with treatment-related mortality after nonmyeloablative hematopoietic cell transplantation. *Bone Marrow Transplant* **46,** 20–6 (2011).

76. *Fludara 2 (fludarabine phosphate) FOR INJECTION FDA Label* `https://www.accessdata.fda.gov/drugsatfda_docs/label/2009/020038s032lbl.pdf`. Accessed: 2018-09-26.

77. *Fludara 50mg powder for solution for injection or infusion Summary of Product Characteristics (SmPC)* `https://www.medicines.org.uk/emc/product/1125/smpc`. Accessed: 2018-05-14.

78. Langenhorst, J. B., Dorlo, T. P. C., van Maarseveen, E. M., Nierkens, S., Kuball, J., Boelens, J. J., *et al.* Population Pharmacokinetics of Fludarabine in Children and Adults during Conditioning Prior to Allogeneic Hematopoietic Cell Transplantation. *Clinical pharmacokinetics.* Epub ahead of print (Oct. 2018).

79. Malspeis, L., Grever, M. R., Staubus, A. E. & Young, D. Pharmacokinetics of 2-F-ara-A (9-beta-D-arabinofuranosyl-2-fluoroadenine) in cancer patients during the phase I clinical investigation of fludarabine phosphate. *Semin Oncol* **17,** 18–32 (1990).

80. Punt, A. M., Langenhorst, J. B., Egas, A. C., Boelens, J. J., van Kesteren, C. & van Maarseveen, E. M. Simultaneous quantification of busulfan, clofarabine and F-ara-A using isotope labelled standards and standard addition in plasma by LC-MS/MS for exposure monitoring in hematopoietic cell transplantation conditioning. *J Chromatogr B Analyt Technol Biomed Life Sci* **1055-1056,** 81–85 (2017).

5 Simultaneous quantification of busulfan, clofarabine and F-ara-A using isotope labelled standards and standard addition in plasma by LC-MS/MS for exposure monitoring in hematopoietic cell transplantation conditioning

Arjen M Punt, Jurgen B Langenhorst, Annelies C Egas, Jaap Jan Boelens, Charlotte van Kesteren, Erik M van Maarseveen

Journal of Chromatography B. 1055-1056, 81–85 (2017).

5.1 Abstract

In allogeneic hematopoietic cell transplantation (HCT) it has been shown that over- or underexposure to conditioning agents have an impact on patient outcomes. Conditioning regimens combining busulfan (Bu) and fludarabine as the prodrug F-ara-AMP (FluP) with or without clofarabine (Clo) are gaining interest worldwide in HCT. To evaluate and possibly adjust full conditioning exposure a simultaneous analysis of Bu, fludarabine (Flu) (F-ara-A, circulating metabolite of FluP) and Clo in one analytical run would be of great interest. However, this is a chromatographical challenge due to the large structural differences of Bu compared to Flu and Clo. Furthermore, for the bioanalysis of drugs it is common to use stable isotope labelled standards (SILS). However, when SILS are unavailable (in case of Clo and Flu) or very expensive, standard addition may serve as an alternative to correct for recovery and matrix effects. This study describes a fast analytical method for the simultaneous analysing of Bu, Clo and Flu with liquid chromatography-tandem mass spectrometry (LCMS) including standard addition methodology using 604 spiked samples. First, the analytical method was validated in accordance with European Medicines Agency guidelines. The lower limit of quantification (LLOQ) were for Bu 10 μg/L and for Clo and Flu 1 μg/L, respectively. Variation coefficients of LLOQ were within 20% and for low medium and high controls were all within 15%. Comparison of Bu, Clo and Flu standard addition results correspond with those obtained with calibration standards in calf serum. In addition for Bu, results obtained by this study were compared with historical data analyzed within therapeutic drug monitoring. In conclusion, an efficient method for the simultaneous quantification of Bu, Clo and Flu in plasma was developed. In addition, a robust and cost-effective method to correct for matrix interference by standard addition was established.

5.2 Introduction

Allogeneic hematopoietic cell transplantation (HCT) is a last resort for a variety of malignant and nonmalignant disorders (e.g. immuno-deficiencies and inherited metabolic diseases). Prior to HCT, preparative conditioning is incorporated to enable engraftment of donor cells. Although no consensus on optimal conditioning exists, the combination of busulfan (Bu), fludarabine as the prodrug F-ara-AMP (FluP) with or without clofarabine (Clo) is used more and more in current HCT conditioning[1]. For Bu a target exposure has been defined[2,3]. However, for Clo and FluP (fludarabine (Flu)) limited pharmacokinetic (PK) and pharmacodynamic (PD) (PKPD) data in the HCT setting are available. Recent reports do suggest that overexposure to Flu or Clo may be a predictor of poor transplant outcomes. Yet, optimal exposure to Flu and Clo is dependent on underlying disease and the selected conditioning regimen among others making it difficult to define a therapeutic window[4–6].

As therapeutic drug monitoring (TDM) of Bu is rapidly gaining interest worldwide, numerous bioanalytical methods predominantly using a stable isotope labelled standards (SILS) are available in literature[7–10]. For quantification of Clo combined with Flu in plasma an analytical method was published recently[11]. For the latter method 2-chloroadenosine was used as an IS, since SIL variants of Clo and SIL-Flu are not commercial available. To correct for matrix interference for Clo and Flu, the use of standard addition commonly used in pesticide analysis may serve as an alternative[12]. Standard addition methodology can be employed to correct for the whole process efficacy including matrix effects and recovery losses, without the need for expensive labeled standards. In contract to pesticide analysis, the expected concentration range of the analytes is known and relatively small facilitating addition of optimal selected spike concentrations. Following their combined use and concomitant administration in the clinic, a simultaneous analysis of Bu, Flu and Clo in one analytical run would be of great interest to determine the optimal exposure for HCT. The latter can be considered a chromatographical challenge due to the large structural differences of Flu and Clo compared to Bu. The lipophilicity of Bu (0.5 logP) is more suitable for reverse phase chromatography whereas the 1.5 logP of Flu fits a normal phase chromatography (HILIC). Therefore, we aimed to develop a combined and fast analytical method for simultaneous quantification of Bu, Clo, Flu in plasma with liquid chromatography-tandem mass spectrometry (LCMS).

5.3 Materials and Methods

Chemicals and reagents

All chemicals and compounds were purchased from Sigma-Aldrich (St. Louis, MO, USA), unless stated differently. Gibco Newborn Calfs Serum, heat inactivated, was obtained from Thermo Fisher Scientific (Waltman, MA, USA). Drug Free Serum was obtained from Bio-Rad (Hercules, CA, USA). Blank bovine serum was obtained from Drug analysis and toxicology (KKGT) studies (SKML, Nijmegen, Netherlands). Random blank patient plasma samples were randomly selected from residual material obtained in routine TDM specimens with patients consent.

Patients samples

After collection of blood, samples were stored in the refrigerator, plasma samples were obtained after centrifugation at the same day or the next day if samples were taken after working hours. After centrifugation samples were directly analyzed and stored in the -80 °C freezer. Samples stored in the freezer (-80 °C) from pharmacokinetic studies in patients undergoing HCT conditioning with Bu, FluP combined with or without Clo and/or rabbit anti-thymocyte globulin (rATG) were used for this study[13,14]. From January 2012 till December 2015 four time points were routinely drawn for the purpose of Bu TDM (t = 5 min, t = 1 h, t = 2 h and t = 3 h) following infusion of Bu. Between December 2015 and September 2016 Clo and Flu were included in the pharmacokinetic study and therefore an additional sample was taken between the end of FluP infusion and the start of Bu infusion (t = -3 h) in each patient. All included patients received Bu and FluP, and in 20% of the patients Clo was administered prior to infusion of both FluP and Bu. The study was approved by the local medical ethical committee. Broad informed consent was obtained from all patients.

Sample preparation

Plasma samples (50 μL) were pipetted into 1.5 mL eppendorf tubes and in order of succession the following solvents were added; 12.5 μL IS dissolved in acetonitrile (ACN) : H2O (1:1), 12.5 μl solvent ACN: H2O (1:1) and 25 μl TCA (20%) and vortexed for 60 seconds. Thereafter, specimens were centrifuged for 5 min at 10.000 g. Finally, 60 μl of the supernatant was transferred into glass vials with 540 μl 5% ACN.

Validation

Validation was performed in accordance with European Medicine Agency (EMA) guidelines[15]. For Bu, Clo and Flu separate stock solutions were made for calibration curves (CC) and quality controls (QC). CC and QC's were freshly prepared on each instance. For the CC eight levels of Bu were made by dilution in N,N,-Dimethylacetamide and for Clo and Flu in ACN/H2O (1:1). For Bu, Clo and Flu 10 μl of each level of the standard solution was added to 970 μl blank calf serum. Standard calibration concentrations for Bu were 10, 50, 250, 1000, 5000, 7500 and 10000 μg/L and for Clo and Flu 1, 5, 25, 100, 500, 1000, 2500 and 5000 μg/L. Bu-D8 (10000 μg/L) and 2-chloroadenosine (2000 μg/L) dissolved in ACN/H2O (1:1) were used as IS for respectively Bu, Flu and Clo. For accuracy and precision testing QC samples for Bu were made at lower limit of quantification (LLOQ) (10 μg/L), low (25 μg/L), medium (4000 μg/L) and high (8000 μg/L) in calf serum and for Clo and Flu respectively 1 μg/l, 2.5 μg/l, 2000 μg/L and 4000 μg/L were used as QC's. For selectivity testing, three samples obtained from an interproficiency testing program and 4 samples obtained from other pharmacokinetic studies were analysed (section 5.3).
For stability testing at room temperature, 18 patient samples were selected, 12 samples contained Bu and Flu and 6 samples contained all three compounds. These samples were stored in a fume hood and four time points were taken at 0 h, 24 h, 48 h and 72 h. Autosampler stability was tested by analysing standard level 3 in calf serum every hour for a total 68 h. Long-term stability and stability after an extra freeze thaw cycle was tested when samples were stored at -80°C. For Flu long-term stability was tested for 6 months and for Clo for 4 months. In addition, Bu concentrations were compared with historical routine monitoring results analysed with an in house prospectively validated LCMS method.
Statistical analyses for within run coefficient of variation (CV), between run CV and overall CV were performed using one-way analysis of variance (ANOVA). Linear regression analysis was used for autosampler stability. For comparing results obtained by standard addition with calibration curve, Deming regression was used in EP Evaluator (Build 10.3.0.556).

Standard addition

Patient samples were spiked to study recovery, ion suppression and matrix effects, Bu (5000 μg/L), Clo (200 μg/L) and Flu (1000 μg/L) were added to 604 samples. The primary goal of standard addition was optimal correction

for potential matrix effects of Clo and Flu. For Bu, a stable labelled D8-Bu was used as an IS. Therefore, T = -3 samples (taken between December 2015 and September 2016) or T = 5 samples (taken from January 2012 till December 2015) were spiked. These samples were prepared in duplicate, by spiking one sample by replacing 12,5 μl ACN:H2O (1:1) (section 2.2.2) with 12,5 μl spike solution made in ACN:H2O (1:1).

Instrumentation

All samples were analysed with an Ultimate 3000 UHPLC Dionex (Sunnyvale, CA) coupled to a triple quadrupole TSQ Quantiva, Thermo Fisher Scientific (Waltman, MA, USA). The method was validated with the following settings: 3 μl sample was injected onto an UPLC Acquity (BEH 2.1 * 50 mm, 1.7 μm particle size) analytical column (column temperature 40 °C), Waters (Milford, MA, USA). Eluents were 0.1% formic acid in water (eluent A) and ACN with 0.1% ammonium acetate (eluent B). The eluent profile consisted of 0-2.4 min isocratic 5% B, 2.4 - 2.75 min linear gradient from 5-95% B, 2.75 – 3.5 min isocratic gradient 95% B, 3.5 - 4.0 min linear gradient from 95-5%, and 4.0 – 4.5 min isocratic gradient 5% B, used flow rate was 0.7ml/min. During method development different columns were tested (Atlantis Hilic 3μm 2.1 x 100 mm (Waters), Hypersil GOLD PFP 5μm 2.1 x 50 mm (Thermo Scientific), Kinetex Biphenyl 2.6 μm 2.1 x 50 mm (Phenomenex)). Based on retention times and peak shapes the above settings were used. Compounds were analysed by selected reaction monitoring (SRM) with the follow transitions: Bu 264>151.2 m/z (12 CE, 60 RF), Bu-D8 272>159 m/z (12 CE, 60 RF) Flu 286.1>154.1 m/z (17 CE, 49 RF) and 286.1>134.1 m/z (38 CE, 49 RF), Clo 304.1>107.1 m/z (23 CE, 77 RF) and 304.1>134.1 m/z (39 CE, 77 RF) and 2-cloroadenosine 302.1>170 m/z (21 CE, 62 RF) and 302.1>134.1 m/z (30 CE, 62 RF). Analytes were quantified with the following MS conditions: in positive mode (3700V), ion transfer tube temperature of 250°C and vaporizer temperature of 375°C.

5.4 Results

LC-MS/MS optimization

Initial focus was on achieving low LLOQ for Flu and Clo, with a fast and easy analytical method. Therefore, protein denaturation was performed by

using TCA (see section 2.2.1) and the LLOQ for Clo and Flu were respectively 0.5 μg/L and 2 μg/L. After the described sample preparation an extra dilution step was needed. A threefold dilution resulted in an LLOQ of <1 μg/L, but the linear range was not fit for the concentrations of Bu, Clo and Flu in human plasma. Moreover, a high linear range was needed due to implementation of standard addition. An optimum was reached by a tenfold dilution after sample preparation. By dilution, additional benefits like reducing matrix effects were obtained[16].

LC-MS/MS optimization

Clo and Flu have similar molecular structures and are more polar compared to Bu. Hence, Bu showed long retention times on a C18 column compared to Clo and Flu. Flu had the highest retention using a high percentage of eluent A. Due to limitations of the C18 column a maximum of 95% could be used for eluent A. An optimum was reached for Flu when using 95% water with 0.1% formic acid and 5% ACN (eluent C) after the elution of Flu using a linear gradient up to 90% eluent B 0.1% ammonium acetate in ACN (eluent B) for the ionisation of Bu. However, under these conditions the IS Bu-D8 showed significant variation between the CC in calf serum, blank human plasma and positive samples for Bu in human plasma. This was resolved by adding 0.1% ammonium acetate to ACN at t= 0 and using an isocratic eluent until all analytes were eluted followed by a wash program. Moreover, addition of 0.1% ammonium acetate to ACN at the start of the analytical run showed an undesired shift in retention time of Flu from 0.5 min to 0.33 min. Then, Flu still eluted after the dead volume of the column. Furthermore, Bu was more retained due to the low percentage of organic solvent with good peak shape (figure 1) and elution profile.

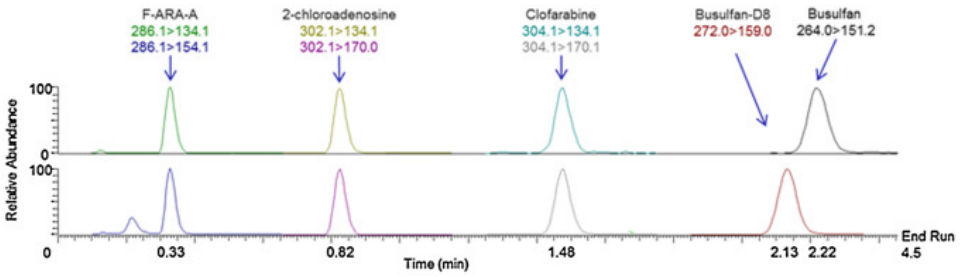

Figure 1: Elution profile, Chromatogram of Bu, Clo and Flu in calf serum.

Validation

Validation of the analytical method was performed according to the EMA guidelines, with incorporation of standard addition. 604 different patients were spiked to monitor matrix effects and recovery losses in plasma. Accuracy, precision and linearity results are presented in table 1. All results for LLOQ (<20%), LOW (<15%), MED (<15%) and HIGH (<15%) were within predefined levels of acceptation. Also, the coefficients of determination (R^2) over three days for each analyte were within acceptable range according the EMA guidelines[15].

Table 1: Validation results for accuracy, precision and linearity.

Bu	**μg/L**	**Within-Run CV (%)**	**Between-Run CV (%)**	**Overall CV (%)**	**Overall Bias (%)**	**R^2**
LLOQ	10	10.5	1.2	10.5	10.3	0.995
LOW	25	5.9	3.5	6.8	-3	
Medium	4000	1	2.2	2.4	-4	
High	8000	1.3	5.5	5.7	-4.3	
Clo	**μg/L**	**Within-Run CV**	**Between-Run CV**	**Overall CV**	**Overall Bias**	**R^2**
LLOQ	1	13	1	13	6.1	0.998
LOW	2.5	8.2	0	8.2	-0.4	
Medium	2000	1.6	9.7	9.9	4	
High	4000	1.3	7.8	7.9	7.3	
F-ARA-A	**μg/L**	**Within-Run CV**	**Between-Run CV**	**Overall CV**	**Overall Bias**	**R^2**
LLOQ	1	17.2	0	17.2	9.9	0.995
LOW	2.5	8.7	8.7	12.4	-1.9	
Medium	2000	1.8	6.9	7.2	-7	
High	4000	2.2	3.5	4.1	-3.4	

Auto sampler, room temperature and long term stability

Autosampler stability of Bu showed a high difference between absolute area and IS ratio. The absolute area of Bu decreased with 20.4% after 68 h based regression analysis. Yet, with the correction using IS it decreased 4.3% after 68 h. For Clo and Flu no differences between compound area and IS ratio were noted. The autosampler stability for Clo and Flu , calculated by regression analysis were respectively -4.2% and 9.2% after 68 h. Results obtained by autosampler stability for Bu, Clo and Flu, were compared with

stability at room temperature as shown in figure 2a-b. As expected for Bu (figure 2a) samples prepared in ACN:H2O were more stable compared to patient plasma samples. For Clo and Flu no differences were found between autosampler and room temperature stability (figure 2b). Stability testing did not show a significant decrease in recovery (less than 15%) after 72 h. For long-term stability (stored at -80°C), samples were reanalysed after 6 months for Bu and Flu with a recovery of 98%, CV 4%, bias -2% and a recovery 93%, CV 8%, bias -7%, respectively. The recovery for long-term stability of Clo after 4 months was 106%, with a CV of 2% and a bias 6%.

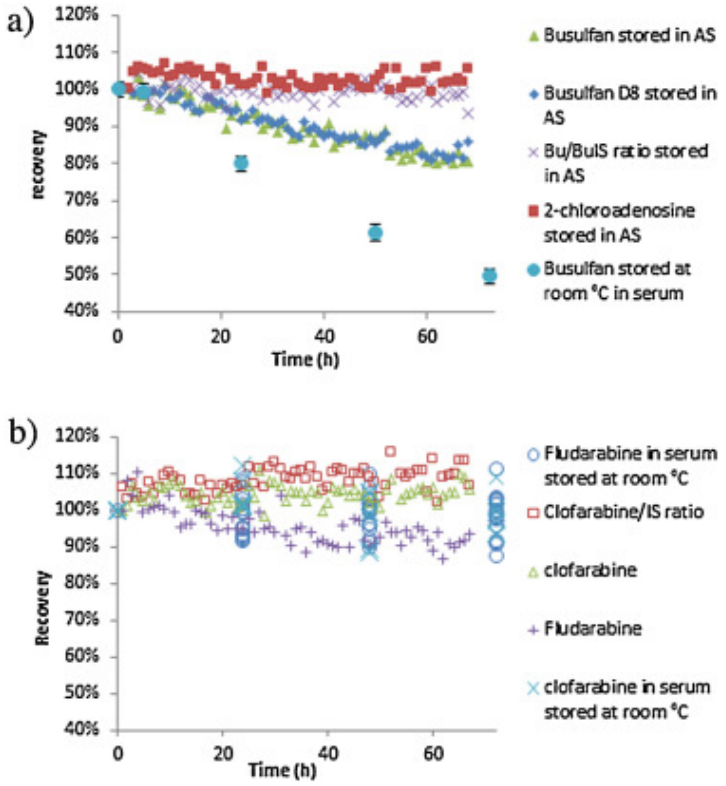

Figure 2: a) Bu stability, Autosampler stability for Bu compared to Bu in plasma stored at room temperature. Samples for autosampler stability were taken every hour; plasma samples stored at room temperature were taken every 24 h. As a reference 2-chloroadenosine (IS for Clo and Flu) was included. Abbreviations: AS = autosampler, IS = internal standard. b) Clo and Flu stability, Autosampler stability for Clo and Flu compared to Clo and Flu in plasma stored at room temperature. Samples for autosampler stability were taken every hour; plasma samples stored at room temperature were taken every 24 h. Abbreviations: AS = autosamler, IS = internal standard.

Also, a within-sample comparison of Bu results using historical routine monitoring data as a reference was performed in 2392 samples (figure 3). The later method was composed of ACN for denaturation of the proteins and also LCMS with an isocratic eluent profile for Bu. For most samples the recovery was between 80% and 100% compared to the historical data. In total 0.2% of the sample had a recovery higher than 115% and 23% of the samples had a recovery below 85%. Assigning the low recovery of these samples to matrix effects, due to a slightly difference in retention time between Bu-D8 and Bu was unlikely. Small differences in retention time between SILS and analyte can result in different analyte response, as described by Wang et al.[17]. However, the reanalysed Bu results were confirmed by standard addition as shown in figure 3. It is most likely that the high percentage of samples with a lower recovery than 85%, can again be explained by the instability of Bu at room temperature. The instability of Bu was previously reported by other authors in literature[18,19] and was confirmed in our study (section 5.4).

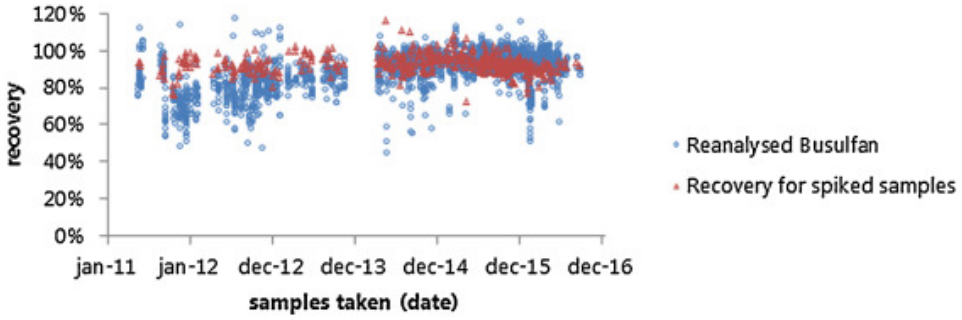

Figure 3: Reanalyzed Bu, In the last 5 years samples were collected and analysed for Bu. All samples were reanalysed with the new method and compared, results expressed by recovery. Results obtained by standard addition were included and used as confirmation of the current method.

Standard addition

In contrast to TDM analysis, for pesticide analyses it is common to use standard addition and guidelines for using standard addition are regulated by EU[12]. As standard addition necessitates an efficient sample preparation, not every sample could be spiked at the optimal spike-to-analyte ratio. As shown in Boer et al.[17] adding a lower spike amount than the CC concentration can result in a high deviation in outcome. The latter was also confirmed for pesticide analysis and the EU guidelines dictates that the ratio spike con-

centration:estimated amount should be at least 1:1 to 5:1.
As Bu concentrations t = 5 min specimens were between 3500 μg/L and 5000 μg/L, a spike concentration of 5000 μg/L was used for Bu. The concentrations for Clo were much lower compared to Bu. For Clo detected concentration range calculated by CC were between 30 μg/L and 120 μg/L (time point dependent). The Flu concentration range was between 50 μg/L and 1200 μg/L, this large range was time point dependant and/or if Clo was included. Like Bu, also for Clo and Flu one level was used for spiking and was directed at T= -3 h (spike concentrations: 200 μg/L for Clo and 1000 μg/L for Flu).
For the majority of the 503 samples spiked for Bu, the recovery calculated by standard addition was between 85% (6% was <85%) and 115% (1% was >115%). In total 2% of the 604 spiked samples for Flu had a recovery lower than 85% and 4% had a higher recovery then 115%. For the 121 spiked samples for Clo, 5% of the samples had a recovery lower than 85% and none of the samples had a recovery higher than 115%. As shown in figure 4, application of a relatively high spike concentration compared to the CC concentration did not show more deviation compared to the optimal spike range. However, adding a low spike amount to high concentrations of Flu resulted a high deviation as shown in figure 4. The later was not noted for Bu or Clo. Even though not all sample was spiked in the optimal range, it can be concluded that results obtained by standard addition compared to CC were in the same range. This was also confirmed when using Deming regression which showed for Bu a correlation R 0.9877, slope 0.966 (0.953 to 0.979) and a intercept of -115,485 (-162,993 to -67,977), for Clo a correlation R 0.9976, slope 0.998 (0.986- to 1.010) and a intercept -0.1104 (-0.157 to -0.063) and for Flu a correlation of R 0.9917, slope 0.982 (0.972- to 0.993) and a intercept of -7.157 (-11.704 to -2.609). Consequently, matrix effects were limited for Bu, Clo and Flu, which was also shown by Huang at al.[11]. However, the advantage of the standard addition method is that matrix effect can be monitored in each individual patient sample. Standard addition methodology is of interest due to high cost efficiency and robust process efficiency, In particular, for multi-component analysis which includes unavailable or expensive SILSs, standard addition serves as a valid alternative. Therefore, development of international guidelines of standard addition for quantification of drugs in biomatrices is warranted.

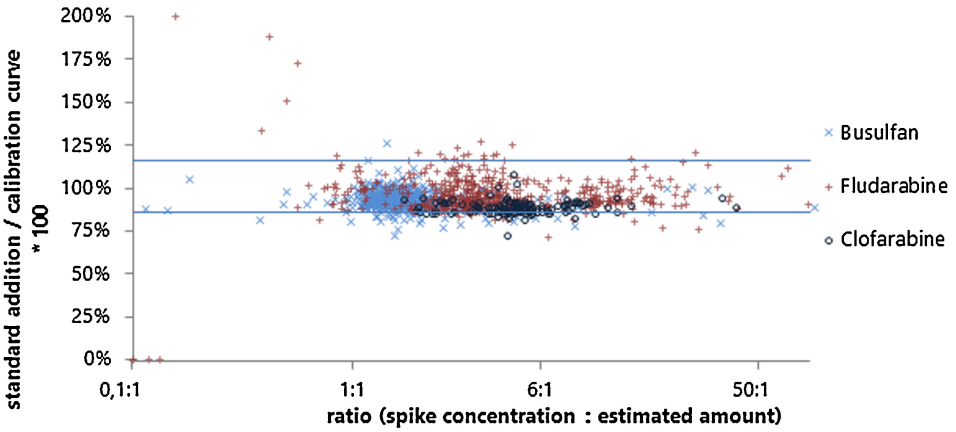

Figure 4: Spike-to-analyte concentration ratio, x-axis is the ratio spike concentration added to sample by concentration calculated by calibration curve. Y-axis calculated standard addition values divided by calibration curve values. Values with a ratio below 0,5:1 are negative or above 200%.

5.5 Conclusion

In conclusion, a fast, easy and robust method for the simultaneous quantification of Bu, Clo and Flu was developed. Next, a robust method to correct for matrix interference by standard addition was established. Also, the analytical method can facilitate routine drug monitoring for patient care and pharmacokinetic studies.

References

1. Admiraal, R. & Boelens, J. J. Individualized conditioning regimes in cord blood transplantation: Towards improved and predictable safety and efficacy. *Expert Opinion on Biological Therapy* **16,** 801–813 (2016).
2. Bartelink, I. H., Lalmohamed, A., Reij, E. M. L. v., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *The Lancet Haematology* **3,** e526–e536 (2016).
3. Palmer, J., McCune, J. S., Perales, M.-A., Marks, D., Bubalo, J., Mohty, M., *et al.* Personalizing Busulfan-Based Conditioning: Considerations from the American Society for Blood and Marrow Transplantation Practice Guidelines Committee. *Biology of Blood and Marrow Transplantation* **22,** 1915–1925 (2016).
4. Lindemans, C. A., Te Boome, L. C., Admiraal, R., Jol-van der Zijde, E. C., Wensing, A. M., Versluijs, A. B., *et al.* Sufficient Immunosuppression with Thymoglobulin Is Essential for a Successful Haplo-Myeloid Bridge in Haploidentical-Cord Blood Transplantation. *Biol Blood Marrow Transplant* **21,** 1839–45 (2015).
5. McCune, J. S., Vicini, P., Salinger, D. H., O'Donnell, P. V., Sandmaier, B. M., Anasetti, C., *et al.* Population pharmacokinetic/dynamic model of lymphosuppression after fludarabine administration. *Cancer Chemother Pharmacol* **75,** 67–75 (2015).
6. Salinger, D. H., Blough, D. K., Vicini, P., Anasetti, C., O'Donnell, P. V., Sandmaier, B. M., *et al.* A limited sampling schedule to estimate individual pharmacokinetic parameters of fludarabine in hematopoietic cell transplant patients. *Clin Cancer Res* **15,** 5280–7 (2009).
7. Bunch, D. R., Heideloff, C., Ritchie, J. C. & Wang, S. A fast and simple assay for busulfan in serum or plasma by liquid chromatography-tandem mass spectrometry using turbulent flow online extraction technology. *J Chromatogr B Analyt Technol Biomed Life Sci* **878,** 3255–8 (2010).
8. Chen, L., Zhou, Z., Shen, M. & Ma, A. Quantitative Analysis of Busulfan in Human Plasma by LC-MS–MS. *Chromatographia* **70,** 1727–1732 (2009).

9. Nadella, T. R., Suryadevara, V., Lankapalli, S. R., Mandava, V. B. & Bandarupalli, D. LC-MS/MS method development for quantification of busulfan in human plasma and its application in pharmacokinetic study. *J Pharm Biomed Anal* **120,** 168–74 (2016).

10. Quernin, M. H., Duval, M., Litalien, C., Vilmer, E. & Aigrain, E. Quantification of busulfan in plasma by liquid chromatography–ion spray mass spectrometry Application to pharmacokinetic studies in children. *Journal of Chromatography B* **763,** 61–69 (2001).

11. Huang, L., Lizak, P., Dvorak, C. C., Aweeka, F. & Long-Boyle, J. Simultaneous determination of fludarabine and clofarabine in human plasma by LC-MS/MS. *J Chromatogr B Analyt Technol Biomed Life Sci* **960,** 194–9 (2014).

12. Guidance document on analytical quality control and method validation procedures for pesticides residues analysis in food and feed. *SANCO/12571/2013* (2016).

13. Admiraal, R., Kesteren, C. v., Zijde, C. M. J.-v. d., Tol, M. J. D. v., Bartelink, I. H., Bredius, R. G. M., *et al.* Population Pharmacokinetic Modeling of Thymoglobulin in Children Receiving Allogeneic-Hematopoietic Cell Transplantation: Towards Improved Survival Through Individualized Dosing. *Clin Pharmacokinet* **2015,** 435–446 (2014).

14. Bartelink, I. H., Belitser, S. V., Knibbe, C. A., Danhof, M., de Pagter, A. J., Egberts, T. C., *et al.* Immune reconstitution kinetics as an early predictor for mortality using various hematopoietic stem cell sources in children. *Biol Blood Marrow Transplant* **19,** 305–13 (2013).

15. Guideline on bioanalytical method validation. **EMEA/CHMP/EWP/192217/2009 Rev. 1 Corr. 2** (2011).

16. Stahnke, H., Kittlaus, S., Kempe, G. & Alder, L. Reduction of Matrix Effects in Liquid Chromatography-Electrospray ionization-Mass Spectrometry by Dilution of the Sample Extracts: How Much Dilution is Needed? *Analytical Chemistry* **84,** 1474-1482 (2011).

17. Wang, S., Cyronak, M. & Yang, E. Does a stable isotopically labeled internal standard always correct analyte response? A matrix effect study on a LC/MS/MS method for the determination of carvedilol enantiomers in human plasma. *Journal of Pharmaceutical and Biomedical Analysis* **43,** 701–707 (2007).

18. Hassan, M. & Ehrsson, H. Degradation of busulfan in aqueous solution. *Journal of Pharmaceutical and Biomedical Analysis* **4,** 95–101 (1986).

19. Houot, M., Poinsignon, V., Mercier, L., Valade, C., Desmaris, R., Lemare, F., *et al.* Physico-chemical stability of busulfan in injectable solutions in various administration packages. *Drugs R D* **13,** 87–94 (2013).

6 A semi-mechanistic model based on glutathione depletion to predict intra-individual reduction in busulfan clearance

Jurgen B Langenhorst, Jill Boss, Charlotte van Kesteren, Arief Lalmohammed, Jürgen Kuball, Jaap Jan Boelens, Alwin DR Huitema, Erik M van Maarseveen

Manuscript in preparation

6.1 Abstract

Aim

To develop a semi-mechanistic model, based on glutathione depletion, to predict a previously identified intra-individual reduction in busulfan clearance to aid in more dosing.

Methods

Busulfan concentration data, measured as part of regular care for allogeneic hematopoietic cell transplantation (HCT) patients, were used to develop a semi-mechanistic model and compare it to a previously developed empirical model. The latter included an empirically estimated time-effect, where the semi-mechanistic model included theoretical glutathione depletion. Here glutathione was assumed to decrease proportional to busulfan metabolism and busulfan clearance was assumed proportional to the amount of glutathione. As older age has been related to lower glutathione levels, this was tested as a covariate in the semi-mechanistic model. Lastly, a therapeutic drug monitoring (TDM) simulation was performed comparing the two models in target attainment.

Results

In both models, a similar clearance decrease of 7% (range: -82% to 44%) with a proportionality to busulfan metabolism was found. The latter was inherently implemented in the semi-mechanistic model. After 40 years of age, the time-effect increased proportionally with 4% per year (0.6-8%, $p = 0.009$), causing the effect to increase more than a 2-fold over the observed age-range (0-73 years). Compared to the empirical model, the final semi-mechanistic model increased target attainment from 74% to 76%, mainly through better predictions for adult patients.

Conclusion

These data suggest that the time-dependent decrease in busulfan clearance may be related to gluthathione depletion. This effect is more clear above the age of 40 years and is proportional to busulfan metabolism. The newly constructed semi-mechanistic model could be used to further improve TDM-guided exposure target attainment of busulfan in patients undergoing HCT.

6.2 Introduction

Allogeneic hematopoietic cell transplantation (HCT) is a high-risk, but potentially curative treatment for a variety of malignant and nonmalignant hematological disorders. Unfortunately, treatment related mortality is substantial (10-40%), implying an urgent need of further optimization of this procedure[1].

Prior to HCT, the bone marrow and immune system of the host are ablated by means of a preparative conditioning regimen. In these conditioning regimens, busulfan is the most frequently used drug[2]. Busulfan is usually administered over a 4-day period and has a narrow therapeutic window, where an exposure corresponding with an area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($\mathrm{AUC}_{t_0-\infty}$) of $80-100 mg \cdot h/L$ ($\approx 20000 - 25000 \mu Mol \cdot min$) has been associated with optimal treatment outcomes in a myeloablative setting[3–5]. Lower exposures have been associated with more frequent relapse or graft failure, and higher exposures with more often severe toxicity and treatment related mortality.

Therapeutic drug monitoring (TDM)-guided dosing is recommended to better attain this narrow target exposure[6]. Indeed, the use of TDM has been proven to increase overall survival by 20% compared to fixed dosing in a randomized controlled trial setting[7]. However, the attainment of the desired busulfan target exposure is still challenging due to an intra-individually variable clearance reduction from day 1 to day 4 with associated variability of 11-15% as has previously been shown[8–10]. These effects respectively limit the accuracy and precision of TDM based on samples measured at the first day of conditioning. As the mechanism behind this time-dependent decrease is unknown, the effect has been implemented empirically, where a more mechanistic approach may better predict inter-individual differences in clearance-reduction.

The primary route of busulfan clearance is through extensive metabolism in the liver: only 2% of busulfan is excreted unchanged in urine. Initial inactivation occurs by conjugation to glutathione (GSH), both spontaneously and aided by the enzyme[11,12]. The busulfan-GSH conjugate is further metabolized via two parallel routes: β-elimination, catalyzed by cystathionine γ-lyase, forming tetrahydrothiophene, pyruvate and ammonium; or through conversion to an N-acetylated cysteine conjugate by N-acetyltransferase. Polymorphisms of the enzyme glutathione-S-transferase (GST) have been used to better predict busulfan clearance *a priori*[13–18], but such predictions are obviated by the use of TDM. Interestingly, it has also been shown that, in patients treated with high dose busulfan, levels of the substrate GSH

decrease by approximately 75%[19]. In addition, higher baseline GSH concentration were correlated with an up to 2-fold increased busulfan-clearance[19]. Therefore, we hypothesize that busulfan-mediated GSH-depletion causes the observed reduction in clearance. In lack of GSH concentration-time-data, a fully mechanistic approach to test this hypothesis was not possible. Nevertheless, patients with a high initial busulfan clearance may exhibit a higher decrease in clearance, following from more pronounced GSH-depletion. This hypothesis was explored in a semi-mechanistic population pharmacokinetic model, ultimately aiming to achieve more predictable busulfan exposure, and thus more predictable outcomes.

6.3 Methods

Patients

Included patients were those who received (non)myeloablative conditioning before HCT, between September 2005 and January 2017 at the UMC Utrecht and of whom plasma concentration data were available. Data consisted of plasma concentrations, measured as part of regular care for HCT patients. The dataset then contained all UMC patients included for the previously developed empirical model[8], which contained data up to 2008, plus all adult patients and children transplanted after September 2009. Patients were included after written informed consent was acquired. Ethical approval by the institutional medical ethics committee of the UMC Utrecht was obtained under protocol number 11/063.

Procedures

The conditioning regimen consisted of intravenous busulfan combined with either fludarabine (+/- clofarabine) or cyclophosphamide. In selected patients transplanted before 2011, targeted busulfan was combined with cyclophosphamide at a cumulative dose of 120 or 200 mg/kg. Four days of busulfan were followed by 2 days (120 mg/m^2) or 4 days (200 mg/m^2) of cyclophosphamide, starting on day -7 and -9 respectively. Busulfan and fludarabine conditioning was administered on day -5 to -2 relative to HCT and consisted of a 1-hour-infusion of fludarabine-phosphate (40 mg/m^2) directly followed by a 3-hour-infusion of busulfan (Busilvex, Pierre Fabre). A 1-hour infusion of clofarabine (30 mg/m^2) preceded a reduced dose of fludarabine (10 mg/m^2) in children with hematological malignancies. Rabbit anti-thymocyte globulin rabbit anti-thymocyte globulin (rATG) was added in the unrelated donor HCT setting: 4-hour infusions at 4 consecutive days

from day -9 (10 mg/kg $< 30kg$, 7.5 mg/kg $> 30kg$) for children and 4 times a 12-hour infusion from day -12 (6 mg/kg) for adults. To patients receiving rATG, clemastine (0.03 mg/kg up to 2 mg), paracetamol (60 mg/kg up to 4g) and 2 mg/kg prednisolone with a maximum of 100 mg were given intravenously prior to rATG infusion.
Busulfan, was targeted using a dosing algorithm[8] and TDM to a myeloablative cumulative 4-day exposure of $80mg \cdot h/L \approx 20000\mu Mol \cdot min$ (target till 2011), $90mg \cdot h/L \approx 22000\mu Mol \cdot min$ (current target), $60mg \cdot h/L \approx 15000\mu Mol \cdot min$ (reduced intensity) or $30mg \cdot h/L \approx 7300\mu Mol \cdot min$ for Fanconi anemia patients (expressed as area under the curve for all doses $[AUC_{T_0-\infty}]$). According to the busulfan TDM protocol, plasma samples were drawn on the first and/or second day of conditioning. In case of large dose adjustments ($> 50\%$), samples were also drawn on subsequent days for confirmatory reasons. Additional samples were taken for all patients at day 4 of conditioning to evaluate target exposure attainment. In general, plasma samples were taken at 5 minutes, 1, 2, and 3 hours, after the end of busulfan infusion. For a subset of patients, additional samples were collected from 4 to 20 hours post-infusion. Samples were analyzed with a validated LC-MS method according to Langman et al.[20]. In children treated before September 2008 a previously published HPLC-UV method was used[21,22]. For dose adjustment, the individual clearance was estimated with a Bayesian approach using the individual samples and a 1-compartmental pharmacokinetic analysis in the software package of MwPharm[23]. The individual clearance was estimated and the doses for the subsequent days were calculated using equation 1.

$$Dose_{adjusted} = (Target(AUC_{t_0-\infty}) - \frac{Dose_{administered}}{CL_{estimated}}) \cdot CL_{estimated} \quad (1)$$

Pharmacokinetic model design and evaluation

The previously published model by Bartelink et al. was used as the basis for a structural model[8]. In that model, weight was included as a covariate on clearance and volume of distribution of the central compartment (V_1). Clearance was included using an empirical weight-changing allometric exponent (equation 2) and V_1 was described by a constant empirically estimated exponent. A stepwise effect of time on clearance was estimated using sepa-

rate values for $day = 1$ and $day > 1$. No further covariates were included.

$$CL_i = CL_{43kg} \cdot \frac{BW_i}{43kg}^{l \cdot BW_i^p} \tag{2}$$

In addition to this model, henceforth referred to as the empirical model, allometric scaling to weight was added for volume of distribution of the peripheral compartment (V_2) and inter-compartmental clearance between V_1 and V_2 (inter-compartmental clearance between V_1 and V_2), using fixed exponents of 1 and 0.75 respectively. Next, the stochastic model was optimized with a higher number of subjects. Inter-individual variability (IIV) and correlations between IIV were tested for clearance, V_1, V_2, and Q. Inter-occasion variability (IOV) was tested only on V_1 and clearance to preserve parsimony. Both IIV and IOV were assumed to be log-normally distributed. The proportional residual variability from the original model was retained, but variances estimated for the different methods of quantification (HPLC-UV and LC-MS) differed.
For the semi-mechanistic model, the above mentioned expanded empirical model was used as a basis, but without the empirical time-effect. A compartment was then added representing the relative amount of GSH available at any time, where the initial amount was assumed to be 1. The model assumed a zero-order synthesis and first-order elimination of GSH. As the relative amount of GSH at baseline is set at 1, the zero-order synthesis rate constant equals the first-order elimination rate constant at equilibrium. Busulfan in the central compartment was assumed to be metabolized in a GSH-dependent way. The full model is depicted in figure 1 and described in equations 3, 4, and 5.

$$\frac{dA_{bu_1}}{dt} = -A_{bu_1} \cdot (k_{bu_{10}} \cdot A_{GSH} + k_{12}) + A_{bu_2} \cdot k_{bu_{21}} \tag{3}$$

$$\frac{dA_{bu_2}}{dt} = A_{bu_1} \cdot k_{bu_{12}} - A_{bu_2} \cdot k_{bu_{21}} \tag{4}$$

$$\frac{dA_{GSH}}{dt} = -\frac{S_{GSH}}{V_1} \cdot A_{GSH} \cdot k_{10} \cdot A_{bu_1} - k_{GSH_{baseline}} \cdot A_{GSH} + k_{GSH_{synthesis}} \tag{5}$$

Compartments bu_1 and bu_2 represent the central and peripheral compartments of busulfan respectively, and GSH represents the theoretical GSH

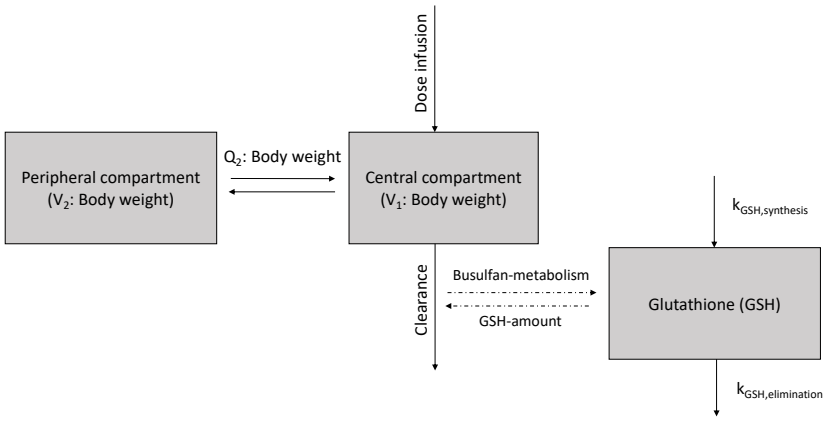

Figure 1: Semi-mechanistic busulfan model structure: busulfan is infused to and eliminated from (Clearance) the central compartment (V1); busulfan distributes reversibly to the peripheral compartment (V2) with a rate determined by the inter compartmental clearance: Q2. Glutathione is synthesized and eliminated according to $k_{GSH_{elimination}}$ and $A_{GSH} \cdot k_{GSH_{elimination}}$ respectively. These terms are assumed to be the same at steady state. The dashed lines indicate the influence busulfan-metabolism and glutathione amount have on each other, where the solid lines depict transport.

compartment. The elimination and distribution constants for busulfan are depicted by k_{bu}. The first-order elimination constant for GSH is depicted by $k_{GSH_{baseline}}$ and the zero-order synthesis constant by $k_{GSH_{synthesis}}$. S_{GSH} is a scaling factor between busulfan metabolism and relevant GSH depletion.

As no GSH concentrations were available in the current analysis, busulfan metabolism was used as a surrogate marker and full GSH dynamics could not be re-constructed. Therefore the following assumptions were made: (1) the last two terms of equation 5 represent the endogenous GSH turnover and were assumed sum up to 0 at baseline; (2) when A_{GSH} decreases as a result of busulfan metabolism, the sum of endogenous turnover terms exceeds zero, resulting in net GSH production; it was assumed that this net synthesis was negligible compared to busulfan dependent depletion, thus equation 5 was

simplified to equation 6.

$$\frac{dA_{GSH}}{dt} = -\frac{S_{GSH}}{V_1^*} \cdot A_{GSH} \cdot k_{10} \cdot A_{bu_1} \tag{6}$$

The factor S_{GSH} as such, scales busulfan metabolism to relevant depletion of GSH. Relevance is defined as depletion to an extent that it becomes a limiting factor in busulfan clearance (as this drives estimation of S_{GSH}). As GSH was set to an absolute amount equal for all individuals (1 at baseline), a scaling factor was necessary for consistent GSH amounts relative to busulfan amounts, to account for the highly variable body size and concurrent dosing in the current dataset. Therefore, S_{GSH} was scaled to individual values for V_1, thus assuming the volume of distribution for GSH to be proportional to V_1 of busulfan.
Age[24,25] has a reported relation to human GSH abundance and turn-over and was tested as a continuous covariate on S_{GSH}.

A population approach based on non-linear mixed-effects modeling was applied[26], using the software package NONMEM (version 7.3.0, Icon, Hanover, MD, USA). Pirana (version 2.9.5) and R (version 3.3.3) were used for workflow management and data handling and visualization, respectively.[27,28] The stochastic approximation and estimation maximization (SAEM) and Monte Carlo importance sampling estimation maximization assisted by mode a posteriori estimation (IMPMAP) as implemented in NONMEM, were respectively used for estimation and objective function calculation.
The structural and covariate model with corresponding estimates had to be scientifically and biologically plausible. A visual inspection of model performance was done through standard goodness-of-fit plots. Examples of these plots are observed concentrations plotted versus individual and population predicted concentrations, and conditional weighted residuals (CWRES) versus time and observed concentrations[29]. Particular emphasis was given to goodness-of-fit plots stratified for different days (occasions) to assess the time-dependent performance. Hierarchical models were statistically compared after backward deletion of the term or covariate of interest. This comparison was done by the objective function value (OFV)(ΔOFV), which follows a chi-square distribution. A ΔOFV of -3.84 then corresponds to a p-value of 0.05 for addition of one parameter (i.e. 1 degree of freedom).
Several other evaluation techniques were performed, all in accordance with European Medicine Agency (EMA) and Food and Drug Administration (FDA) guidelines for population pharmacokinetic analyses[30,31]. A sampling

importance resampling (SIR) evaluation (final step: 2000 samples, 1000 resamples) was performed to estimate parameter precision. To assess the simulation properties, prediction-corrected visual predictive checks (VPCs) were created to judge predictive performance of the final model as compared to the observed concentrations. The prediction-corrected VPC allows for variability in dosing[32]. In this analysis, the observed concentration data and its median and 95% confidence interval (CI) were compared to 95% CI of the predicted mean, 2.5^{th} and 95^{th} percentile, derived from 1000 model simulations.

TDM-guided target attainment evaluation

This analysis aimed to compare the Bayesian forecasting properties of both models in a TDM setting. For this, all patients targeted to an $AUC_{t_0-\infty}$ of 90 mg*h/L with samples available on at least day 1 and 4 were included. TDM was simulated by using both models to predict busulfan clearance throughout conditioning, using only the samples available on day 1. Subsequently, equation 1 was used to calculate the required dose for days 2, 3 and 4. The predicted $AUC_{t_0-\infty}$ was calculated by using the post-hoc estimates estimated using all available pharmacokinetic data and dosing as calculated from the day 1-TDM simulations. The target $AUC_{t_0-\infty}$ attainment rates were assessed for both models.

6.4 Results

Patient characteristics

A total of 386 patients were included with a median age of 14 years (range 0.16-73), from whom 3994 samples were collected. Of these patients, 292 received busulfan targeted to 90 mg*h/L and 94 received doses targeted to a lower exposure (as described in methods section 6.3). Most patients (n=259) received 160 mg/m^2 fludarabine next to busulfan. Alternative conditioning consisted mostly of either 120 mg/m^2 clofarabine with 40 mg/m^2 fludarabine (n=68) or 120-200 mg/kg cyclosphosphamide (n=54) in addition to the targeted busulfan. Serotherapy (rATG) was given to 78% (n=303) of patients. Detailed patient characteristics are shown in table 1.

Pharmacokinetic models

Empirical model Parameter estimates and 95% CI of the adjusted empirical model (as described in section 6.3) can be found in table 2. IIV on

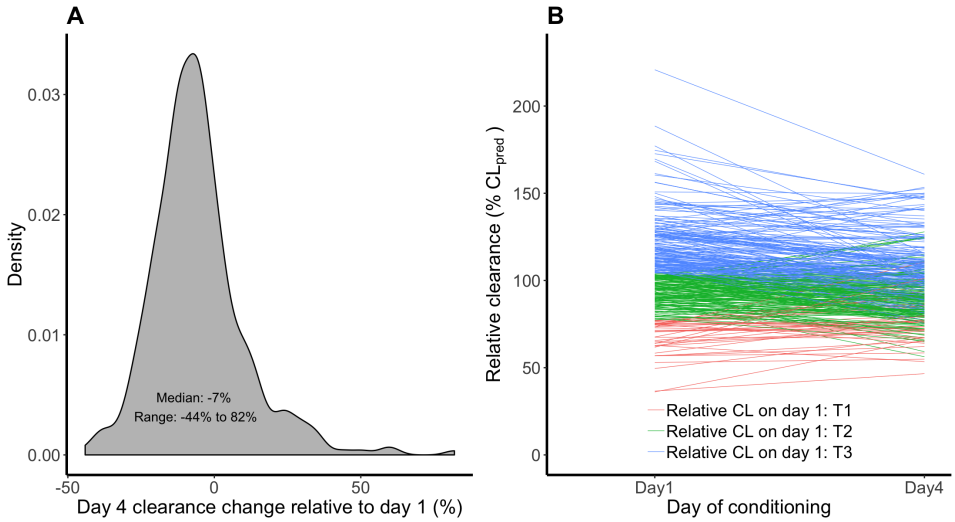

Figure 2: Observed variability in busulfan clearance change. Panel A depicts a density plot of the relative change of clearance from day 1 to day 4 (%). In panel B, this change is displayed per individual and stratified in tertiles for relative clearance day 1: defined as the individually estimated clearance divided by the weight-predicted clearance for that individual (%).

clearance, V_1 and V_2 were estimated at 14%, 19%, and 28% respectively. A correlation of 71% between IIV of V_1 and clearance was found. IOV was implemented on V_1 (11%) and clearance (11%). A proportional residual error was separately estimated for samples measured with UV (9.0%) and MS (6.6%). The population mean clearance for day>1 was estimated to be 7% (95% CI: 5-8%) lower than day 1. Compared to random IOV, the predicted decrease was limited illustrated by an estimated difference in clearance between day 1 and 4 ranging from 87% (increase) down to -44% (decrease) as is depicted in figure 2A. In figure 2B the intra-individual change in clearance from day 1 to 4 is depicted, stratified for tertiles of individual clearance relative to the population predicted clearance (equation 2). A general trend was observed that patients with a higher clearance have a relatively stronger reduction in clearance compared to patient with a low clearance on day 1, who regularly had an increased busulfan clearance over time.

Semi-mechanistic model The semi-mechanistic model was developed and final estimates with corresponding SIR-derived 95% CI are shown in table 2. A similar overall reduction of clearance was estimated compared to

Table 1: Patient characteristics

Weight at HCT (kg)	50 (3.7-130)
Age at HCT (years)	14 (0.16-73)
Age categorie at HCT	
Children: 0-12 years	159 (41%)
Adolescents: 12-20 years	91 (24%)
Adults: 20-40 years	35 (9%)
Adults: 40-60 years	59 (15%)
Adults: 60+ years	43 (11%)
Samples (No. per patient)	11 (2-20)
Sex	
Male	233 (60%)
Female	154 (40%)
Cell source	
Cord blood	179 (46%)
Peripheral blood stem cells	120 (31%)
Bone marrow	76 (20%)
Autologous	7 (1.8%)
Haplo-Cord	5 (1.3%)
Conditioning regimen	
Bu90/Flu	215 (56%)
Bu90/Clo/Flu	68 (18%)
Bu<90/Cy	46 (12%)
Bu<90/Flu	44 (11%)
Other	14 (3.6%)
Diagnosis	
Leukemia	179 (46%)
Benign	132 (34%)
MDS	33 (8.5%)
Plasma cell disorder	24 (6.2%)
Lymphoma	19 (4.9%)
Serotherapy	
Serotherapy	303 (78%)
No Serotherapy	84 (22%)

Table 2: Final model parameter estimates

Fixed effects: Empirical model			
Parameter	Estimate		95% CI
V_1 (L/43kg)	23.4		22 - 24
Exponent V_1	0.869		0.84 - 0.88
Clearance at Day 1 (L/h/43kg)	7.58		7.5 - 7.9
Exponent1 Cl: l	1.03		0.95 - 1.1
Exponent1 Cl: p	-0.138		-0.17 - -0.11
V_2 (L/43kg)	4.83		4 - 5.8
Q (L/h/43kg	5.6		4 - 8.1
$CL_{decrement}$ after Day1 (%)	0.0676		0.05 - 0.082
Random effects: Empirical model			
Parameter	Estimate	Correlation	95% CI
IIV V_1	13.8%		12 - 15
IIV CL	19.4%	74.9% (V1)	18 - 21
IIV V_2	27.4%		20 - 36
IOV CL	12.4%		11 - 13
IOV V_1	12%		10 - 13
Proportional error (HPLC)	9%		7.8 - 11
Proportional error (LCMS)	6.6%		6.4 - 6.8

Fixed effects: Semi-Mechanistic model			
Parameter	Estimate		95% CI
V_1 (L/43kg)	23.3		22 - 24
Exponent V_1	0.863		0.84 - 0.88
CL at T=0 (L/h/43kg)	7.61		7.4 - 7.8
Exponent1 CL: l	1.04		0.95 - 1.1
Exponent1 CL: p	-0.14		-0.17 - -0.1
V_2 (L/43kg)	4.73		3.9 - 5.7
Q (L/h/43kg	5.9		4.1 - 8.6
S_{GSH} (h/mg, age$\leq$40 years)	0.00259		0.0017 - 0.0032
Age effect on S_{GSH} (proportional/year, age>40 years)	0.0419		0.0062 - 0.081
Random effects: Semi-Mechanistic model			
Parameter	Estimate	Correlation	95% CI
IIV V_1	13.8%		12 - 15
IIV CL	19.5%	68.4% (V1)	18 - 21
IIV V_2	28.6%		20 - 35
IOV CL	12.1%		11 - 13
IOV V_1	11.5%		10 - 13
Proportional error (HPLC)	9.1%		7.7 - 10
Proportional error (LCMS)	6.6%		6.4 - 6.8

the empirical model (figure 3). However, with the semi-mechanistic model a gradual reduction in clearance was assumed as GSH was presumed to decrease proportional to busulfan metabolism. The S_{GSH} was estimated at 0.0026 h/mg, implying a net relevant GSH reduction of 0.26% per hour for each mg of busulfan metabolism scaled to 1 liter V_1.
Age appeared to have an effect on the time-dependent decline of clearance (figure 4A), but the effect was only relevant after the age of 40 years ($p = 0.009$). The effect was modelled as a proportional increase of the S_{GSH} of 4% for each year of age (equation 7). This resulted in a more than 2-fold increase of the effect from 40 to 73 years of age. (figure 4B).

$$S_{GSH,i} = S_{GSH,pop} \cdot (1 + (Age_i - 40) * Slope_{Age}) \tag{7}$$

Herein, $Slope_{Age}$ was assumed to be 0 below 40 years of age and was estimated for patients older than 40.
rATG was tested on the time-dependent decline of clearance as a surrogate covariate for paracetamol usage, but no improvement in the model was found.

Model evaluation Figure 5 depicts the goodness-of-fit plots for both models. In both models, no time-dependent trends could be observed. In the VPC stratified for days of conditioning no other misspecifications were seen for either model (data not shown).

TDM-guided target attainment evaluation

258 patients were available for TDM simulation, of which the results are reported in figure 6. Overall, target attainment was slightly better when the semi-mechanistic model was used (top panel, 75%), compared to the empirical model (74%). The final model with age on the S_{GSH} further increased target attainment to 76%. Severe over-exposure (>25% above target) was similar for all models, while $AUC_{t_0-\infty}$ of > 25% below target was not simulated for any of the scenario's.
Because of the apparent effect of age, a subset analysis was conducted in children and adults separately. Here it was found that target attainment in adults increased from 76% with the emperical model to 80% by using the final semi-mechanistic model (figure 6: bottom panel).

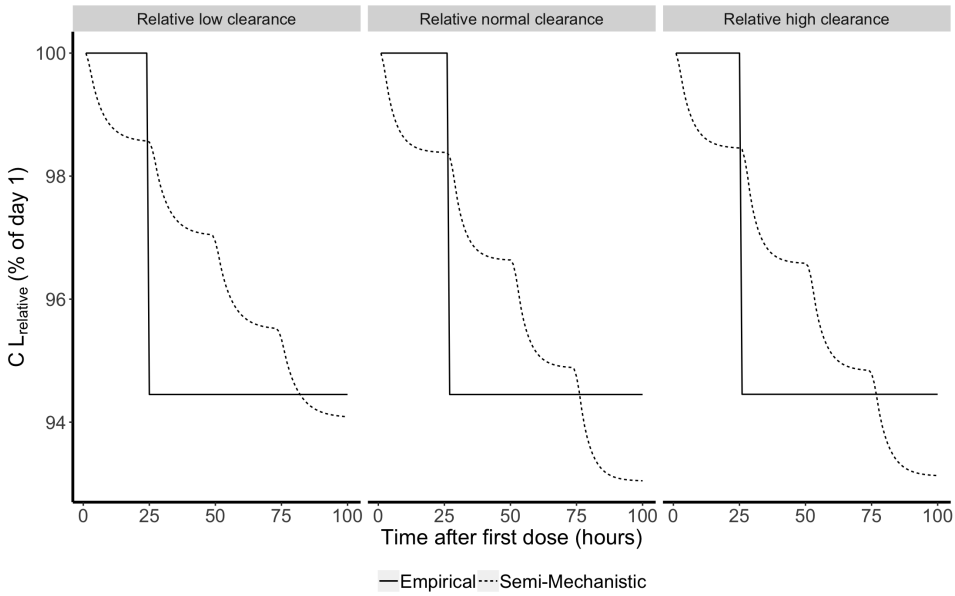

Figure 3: Time-effect. A display of the time-effect for both models. Three individuals were randomly drawn from the each tertile of relative clearance (as defined in figure 2). The clearance over time is depicted, as predicted per model, based on the day 1 clearance. Values are relative to the day 1 clearance (%).

6.5 Discussion

To our knowledge, this is the first pharmacokinetic model describing the decrease in busulfan clearance in a large cohort of both children and adult HCT recipients. We demonstrated an overall 7% decline in busulfan clearance over time, which was more pronounced in older adult patients (>40 years of age). The observed clearance reduction of busulfan is of clinical relevance due to a combination of the narrow therapeutic window and a large between-patient variation in clearance over time during the pre-HCT conditioning phase. Furthermore, patients with a high initial busulfan clearance showed a more pronounced decrease compared to patients with a lower initial clearance. These observations should be taken into account for precise and accurate targeting of busulfan using TDM. We hypothesized that this reduction in clearance is due to GSH depletion and constructed a semi-mechanistic model. Next to biology of GSH homeostasis and GST conjugation, the main arguments in support of the hypothesis are proportionality of the time-effect to busulfan metabolism and the increased effect in older

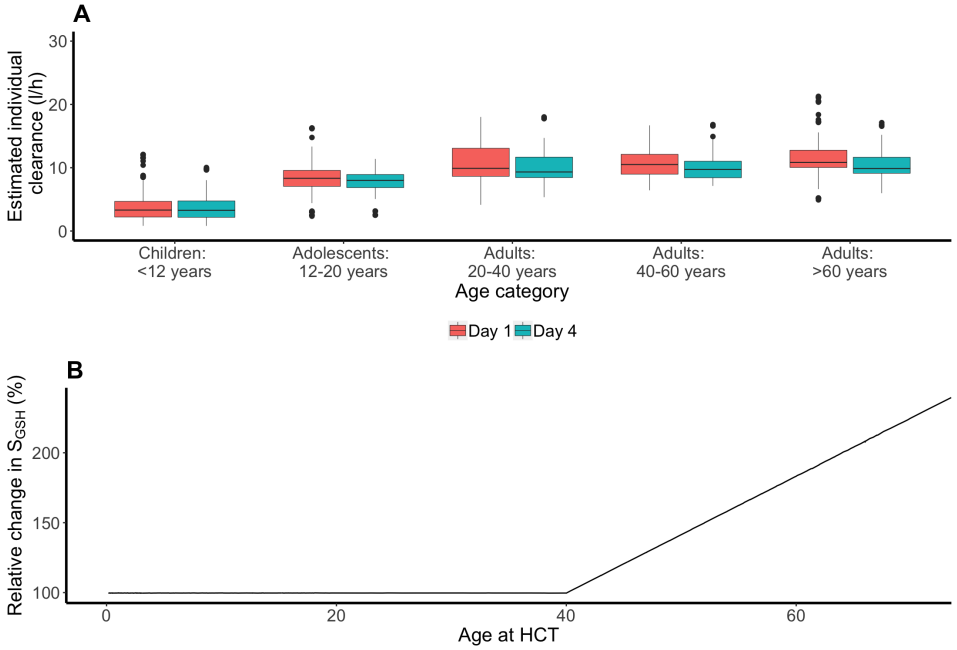

Figure 4: Covariate effects. Panel A depicts the observed clearance decrements from day 1 to 4 stratified for age at transplantation. Panel B depicts the model predicted decrease of S_{GSH} as implemented in the semi-mechanistic model.

age. In TDM simulations, adult patients showed the most improvement in target attainment using the semi-mechanistic model. Under-exposure was less in these patients, reducing the risk of relapse and graft failure.
Besides the direct association between GSH levels and busulfan clearance[19], there is also indirect evidence from metabolomics. Glycine levels, an important substrate in GSH sythesis, were positively associated with busulfan clearance[33]. Also the age effect is supported by literature and can biologically be explained with older age (60-80 years) being associated with decreased GSH-synthesis and thereby absolute levels[24,25]. Here, we found a linear increase in theoretical GSH-depletion from the age of 40. Perhaps, the latter effect is caused by a relatively low initial GSH reservoir, which results in the same absolute busulfan dependent depletion of GSH to have a more relevant effect on clearance in patients aged 40 years and older. In addition, GST polymorphisms have been linked to busulfan clearance with variable results[13–18], which can be explained using the presented hypothesis. Though patients with increased GST activity would initially have a higher clearance, they would also have faster GSH depletion. Thus the average

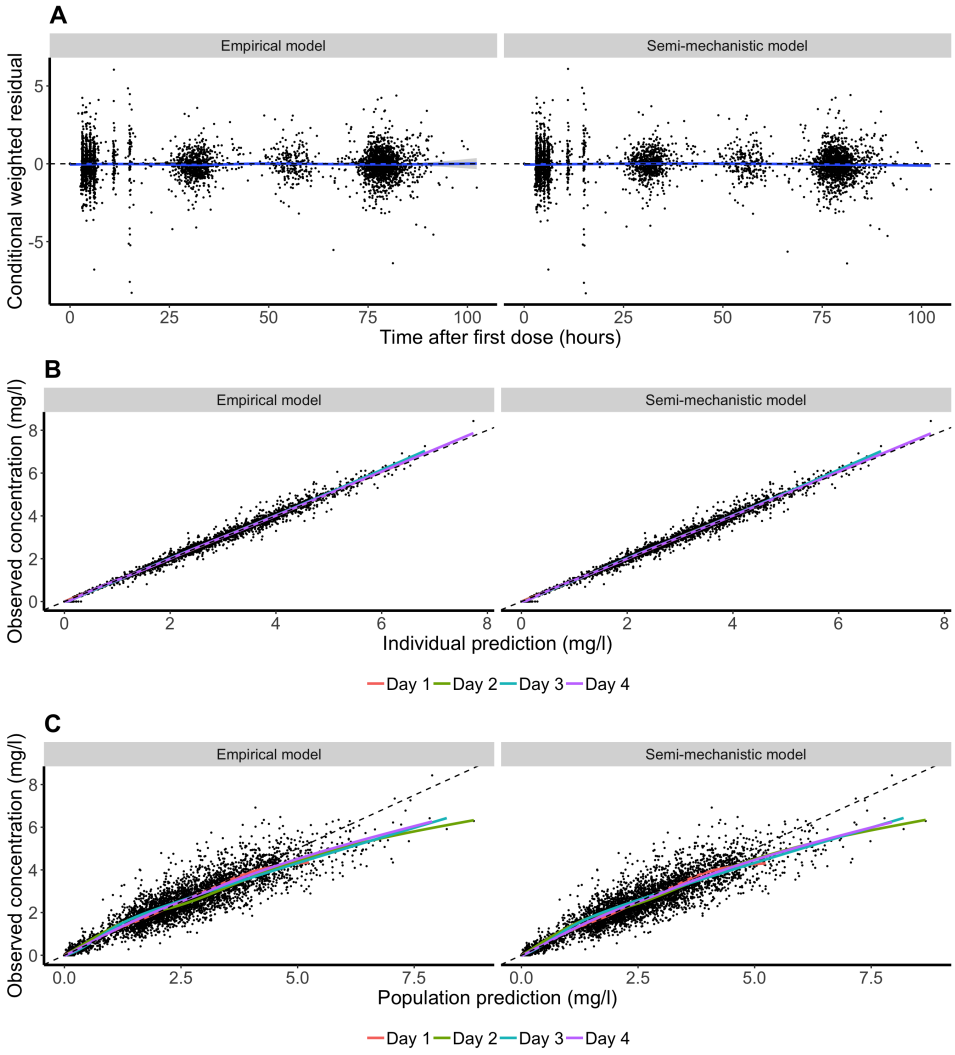

Figure 5: Goodness of fit plots for the empirical and semi-mechanistic model. Panel A depicts the CWRES versus time after first dose. Panel B and C show the observed plasma concentration versus the the individual and population predictions respectively. For all panels LOESS smooths are drawn, where in Panel B and C these are stratified by day of conditioning.

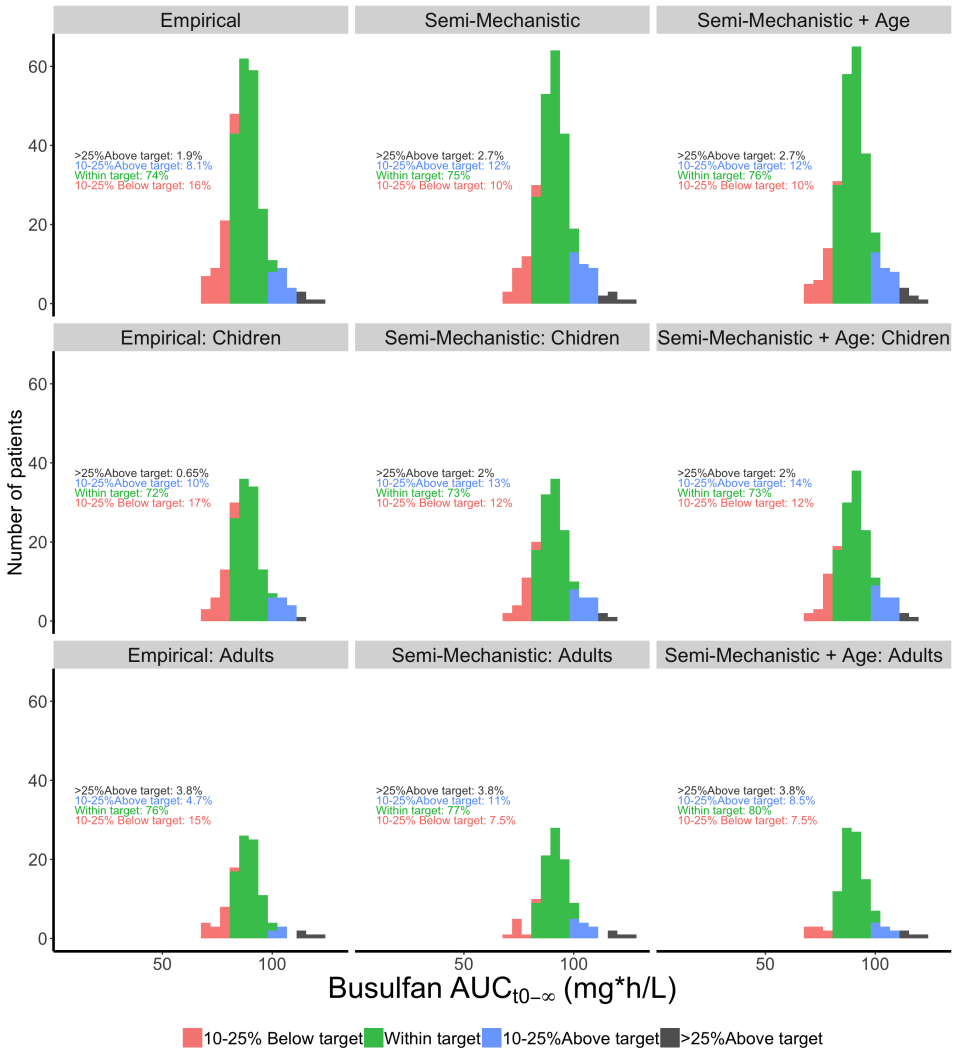

Figure 6: Target attainment for different models. Histograms of the simulated busulfan exposure using the empirical model or the semi-mechanistic model with or without age. Results are shown for the full population as well as for children (< 20 years) and adults (≥ 20 years) The target range is defined as 90 mg*h/L $\pm$ 10% (81-99 mg*h/L)

clearance over multiple doses may be similar to patients with less active GST subtypes. In the proposed setting, TDM accounts for the difference in initial clearance and the semi-mechanistic model predicts the concurrent extent of GSH depletion.
Furthermore, the GSH-dependent time-effect might have other implications, that were not quantified in this study. For example, other drugs that affect busulfan pharmacokinetics or GSH stores such as antifungal agents or paracetamol[34] could interact with the busulfan time-effect. In addition, treatment with N-acetylcysteine might be helpful in preventing severe side effects during treatment with busulfan. In previous research N-acetylcysteine was found to potentially serve as prophylactic agent against sinusoidal obstructive syndrome induced by busulfan[35,36]. Also, evidence was provided that N-acetylcysteine does not interfere with the myeloablative effect of busulfan. Thus N-acetylcysteine might be suitable to reduce the risk of hepatotoxic side effects of busulfan during conditioning regimens for hematopoietic stem cell transplantation.
A major strength of this study is the large sample size with a good distribution of patient over different age groups and limited missing data. As these time-dependent effects are subtle and variable, information including sufficient data over a wide age range was essential for proper quantification of effects. Nevertheless, some weaknesses remain. As direct measurement of GSH was unavailable, busulfan clearance was used as a surrogate. Future studies should focus on measuring active GSH levels before and during the conditioning and implement these in the constructed semi-mechanistic model. The developed mechanistic model can then be expanded with re-synthesis of GSH. GSH levels can be measured also in the time-course after a busulfan dose has been cleared and before administration of the subsequent dose, where most GSH re-synthesis is expected to take place. Preferably, this should be preceded by in vivo (animal) data to support a relationship between plasma and liver levels, as it is known that most GSH is stored in red blood and hepatic cells[37].
In summary, these data suggest the intra-individual decrease in busulfan clearance may be related to GSH depletion. This effect increases after an age of 40 years and is proportional to busulfan metabolism. Therefore, busulfan dosing guided by TDM, taking into account the decrease in clearance using the newly constructed semi-mechanistic model, can increase target attainment in patients undergoing conditioning prior to HCT.

References

1. Hahn, T., McCarthy P. L., J., Hassebroek, A., Bredeson, C., Gajewski, J. L., Hale, G. A., *et al.* Significant improvement in survival after allogeneic hematopoietic cell transplantation during a period of significantly increased use, older recipient age, and use of unrelated donors. *J Clin Oncol* **31,** 2437–49 (2013).
2. Ciurea, S. O. & Andersson, B. S. Busulfan in hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant* **15,** 523–36 (2009).
3. Andersson, B. S., Thall, P. F., Madden, T., Couriel, D., Wang, X., Tran, H. T., *et al.* Busulfan systemic exposure relative to regimen-related toxicity and acute graft-versus-host disease: defining a therapeutic window for i.v. BuCy2 in chronic myelogenous leukemia. *Biol Blood Marrow Transplant* **8,** 477–85 (2002).
4. Bartelink, I. H., Lalmohamed, A., van Reij, E. M., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *Lancet Haematol* **3,** e526–e536 (2016).
5. Russell, J. A., Kangarloo, S. B., Williamson, T., Chaudhry, M. A., Savoie, M. L., Turner, A. R., *et al.* Establishing a target exposure for once-daily intravenous busulfan given with fludarabine and thymoglobulin before allogeneic transplantation. *Biol Blood Marrow Transplant* **19,** 1381–6 (2013).
6. Palmer, J., McCune, J. S., Perales, M.-A., Marks, D., Bubalo, J., Mohty, M., *et al.* Personalizing Busulfan-Based Conditioning: Considerations from the American Society for Blood and Marrow Transplantation Practice Guidelines Committee. *Biology of Blood and Marrow Transplantation* **22,** 1915–1925 (2016).
7. Andersson, B. S., Thall, P. F., Valdez, B. C., Milton, D. R., Al-Atrash, G., Chen, J., *et al.* Fludarabine with pharmacokinetically guided IV busulfan is superior to fixed-dose delivery in pretransplant conditioning of AML/MDS patients. *Bone Marrow Transplant* **52,** 580–587 (2017).
8. Bartelink, I. H., Boelens, J. J., Bredius, R. G., Egberts, A. C., Wang, C., Bierings, M. B., *et al.* Body weight-dependent pharmacokinetics of busulfan in paediatric haematopoietic stem cell transplantation patients: towards individualized dosing. *Clin Pharmacokinet* **51,** 331–45 (2012).

9. Long-Boyle, J. R., Savic, R., Yan, S., Bartelink, I., Musick, L., French, D., *et al.* Population pharmacokinetics of busulfan in pediatric and young adult patients undergoing hematopoietic cell transplant: a model-based dosing algorithm for personalized therapy and implementation into routine clinical use. *Ther Drug Monit* **37,** 236–45 (2015).

10. McCune, J. S., Bemer, M. J., Barrett, J. S., Scott Baker, K., Gamis, A. S. & Holford, N. H. Busulfan in infant to adult hematopoietic cell transplant recipients: a population pharmacokinetic model for initial and Bayesian dose personalization. *Clin Cancer Res* **20,** 754–63 (2014).

11. *BUSULFEX (busulfan) for injection FDA Label* `https://www.accessdata.fda.gov/drugsatfda_docs/label/2015/020954s014lbl.pdf`. Accessed: 2018-09-21.

12. Myers, A. L., Kawedia, J. D., Champlin, R. E., Kramer, M. A., Nieto, Y., Ghose, R., *et al.* Clarifying busulfan metabolism and drug interactions to support new therapeutic drug monitoring strategies: a comprehensive review. *Expert Opin Drug Metab Toxicol* **13,** 901–923 (2017).

13. Ansari, M. & Krajinovic, M. Can the pharmacogenetics of GST gene polymorphisms predict the dose of busulfan in pediatric hematopoietic stem cell transplantation? *Pharmacogenomics* **10,** 1729–32 (2009).

14. Ansari, M., Lauzon-Joset, J. F., Vachon, M. F., Duval, M., Theoret, Y., Champagne, M. A., *et al.* Influence of GST gene polymorphisms on busulfan pharmacokinetics in children. *Bone Marrow Transplant* **45,** 261–7 (2010).

15. Choi, B., Kim, M. G., Han, N., Kim, T., Ji, E., Park, S., *et al.* Population pharmacokinetics and pharmacodynamics of busulfan with GSTA1 polymorphisms in patients undergoing allogeneic hematopoietic stem cell transplantation. *Pharmacogenomics* **16,** 1585–94 (2015).

16. Kim, S. D., Lee, J. H., Hur, E. H., Lee, J. H., Kim, D. Y., Lim, S. N., *et al.* Influence of GST gene polymorphisms on the clearance of intravenous busulfan in adult patients undergoing hematopoietic cell transplantation. *Biol Blood Marrow Transplant* **17,** 1222–30 (2011).

17. Yin, J., Xiao, Y., Zheng, H. & Zhang, Y. C. Once-daily i.v. BU-based conditioning regimen before allogeneic hematopoietic SCT: a study of influence of GST gene polymorphisms on BU pharmacokinetics and clinical outcomes in Chinese patients. *Bone Marrow Transplant* **50,** 696–705 (2015).

18. Zwaveling, J., Press, R. R., Bredius, R. G., van Derstraaten, T. R., den Hartigh, J., Bartelink, I. H., *et al.* Glutathione S-transferase polymorphisms are not associated with population pharmacokinetic parameters of busulfan in pediatric patients. *Ther Drug Monit* **30,** 504–10 (2008).

19. Almog, S., Kurnik, D., Shimoni, A., Loebstein, R., Hassoun, E., Gopher, A., *et al.* Linearity and stability of intravenous busulfan pharmacokinetics and the role of glutathione in busulfan elimination. *Biol Blood Marrow Transplant* **17,** 117–23 (2011).

20. Langman, L. J., Danso, D., Robert, E. & Jannetto, P. J. High-Throughput Quantitation of Busulfan in Plasma Using Ultrafast Solid-Phase Extraction Tandem Mass Spectrometry (SPE-MS/MS). *Methods Mol Biol* **1383,** 89–95 (2016).

21. Zwaveling, J., Bredius, R. G., Cremers, S. C., Ball, L. M., Lankester, A. C., Teepe-Twiss, I. M., *et al.* Intravenous busulfan in children prior to stem cell transplantation: study of pharmacokinetics in association with early clinical outcome and toxicity. *Bone Marrow Transplant* **35,** 17–23 (2005).

22. Cremers, S., Schoemaker, R., Bredius, R., den Hartigh, J., Ball, L., Twiss, I., *et al.* Pharmacokinetics of intravenous busulfan in children prior to stem cell transplantation. *Br J Clin Pharmacol* **53,** 386–9 (2002).

23. Proost, J. H. & Meijer, D. K. MW/Pharm, an integrated software package for drug dosage regimen calculation and therapeutic drug monitoring. *Comput Biol Med* **22,** 155–63 (1992).

24. Rathbun, W. B. & Murray, D. L. Age-related cysteine uptake as rate-limiting in glutathione synthesis and glutathione half-life in the cultured human lens. *Exp Eye Res* **53,** 205–12 (1991).

25. Van Lieshout, E. M. & Peters, W. H. Age and gender dependent levels of glutathione and glutathione S-transferases in human lymphocytes. *Carcinogenesis* **19,** 1873–5 (1998).

26. Sheiner, L. B., Rosenberg, B. & Marathe, V. V. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *J Pharmacokinet Biopharm* **5,** 445–79 (1977).

27. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

28. Keizer, R. J., van Benten, M., Beijnen, J. H., Schellens, J. H. & Huitema, A. D. Pirana and PCluster: a modeling environment and cluster infrastructure for NONMEM. *Comput Methods Programs Biomed* **101,** 72–9 (2011).

29. Hooker, A. C., Staatz, C. E. & Karlsson, M. O. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. *Pharm Res* **24,** 2187–97 (2007).

30. EMEA. Guideline on reporting the results of population pharmacokinetic analyses (2007).

31. FDA. Guidance for Industry, Population Pharmacokinetics. (1999).

32. Bergstrand, M., Hooker, A. C., Wallin, J. E. & Karlsson, M. O. Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. *AAPS J* **13,** 143–51 (2011).

33. Navarro, S. L., Randolph, T. W., Shireman, L. M., Raftery, D. & McCune, J. S. Pharmacometabonomic Prediction of Busulfan Clearance in Hematopoetic Cell Transplant Recipients. *Journal of proteome research* **15,** 2802–2811 (8 Aug. 2016).

34. Slattery, J. T., Wilson, J. M., Kalhorn, T. F. & Nelson, S. D. Dose-dependent pharmacokinetics of acetaminophen: evidence of glutathione depletion in humans. *Clin Pharmacol Ther* **41,** 413–8 (1987).

35. Sjöö, F., Aschan, J., Barkholt, L., Hassan, Z., Ringdén, O. & Hassan, M. N-acetyl-L-cysteine does not affect the pharmacokinetics or myelosuppressive effect of busulfan during conditioning prior to allogeneic stem cell transplantation. *Bone marrow transplantation* **32,** 349–354 (4 Aug. 2003).

36. Hassan, Z., Hellstrom-Lindberg, E., Alsadi, S., Edgren, M., Hagglund, H. & Hassan, M. The effect of modulation of glutathione cellular content on busulphan-induced cytotoxicity on hematopoietic cells in vitro and in vivo. *Bone Marrow Transplant* **30,** 141–7 (2002).

37. B, H. & JMC, G. *Free radicals in biology & medicine 5th ed. Oxford: .* 5th edition (Oxford university press, 2015).

7 Population pharmacokinetics of fludarabine in children and adults during conditioning prior to allogeneic hematopoietic cell transplantation

Jurgen B Langenhorst, Thomas PC Dorlo, Erik M van Maarseveen, Stefan Nierkens, Jürgen Kuball, Jaap Jan Boelens, Charlotte van Kesteren, Alwin DR Huitema

Clinical pharmacokinetics. Epub ahead of print (October 2018).

7.1 Abstract

Background

Fludarabine is often used as an important drug in reduced-toxicity conditioning regimens prior to allogeneic hematopoietic cell transplantation (HCT). As no definitive pharmacokinetic (PK) basis for HCT dosing for the wide age- and weight range in HCT is available, linear body-surface-area (BSA) dosing is still used.

Aim

We sought to describe the population PK of fludarabine in HCT recipients of all ages.

Methods

From 258 HCT recipients aged 0.3 to 74 years, 2605 samples were acquired on day 1 (42%), day 2 (17%), day 3 (4%) and day 4 (37%) of conditioning. Herein, the circulating metabolite of fludarabine was quantified and derived concentration-time data were used to build a population PK model using non-linear mixed effects modelling.

Results

Variability was extensive where area-under-the-curve ranged from 10-66 mg*h/L. A 3-compartment model with first-order kinetics best described the data. Actual body weight (BW) with standard allometric scaling was found to be the best body-size descriptor for all PK parameters. Estimated glomerular filtration rate (eGFR) was included as a descriptor of renal function. Thus, clearance was differentiated into a non-renal (3.24±20% L/h/70kg) and renal (eGFR*0.782±11% L/h/70kg) component. The typical volumes of distribution of the central (V_1), peripheral (V_2) and second peripheral (V_3) compartment were 39±8%, 20±11%, and 50±9% L/70kg respectively. Inter-compartmental clearances between V_1-V_2 and V_1-V_3 were 8.6±8% and 3.8±13% L/h/70 kg respectively.

Conclusion

BW and eGFR are important predictors of fludarabine PK. Therefore, current linear BSA-based dosing leads to highly variable exposure, which may lead to variable treatment outcome.

7.2 Introduction

Allogeneic hematopoietic cell transplantation (HCT) is a potentially curative treatment for a variety of malignant and benign hematological disorders. The preparative or conditioning regimen prior to HCT consists of a combination of cytotoxic agents, chemo- and serotherapy (antibodies against the host immune system), administered to ablate the bone marrow and the immune system[1]. Fludarabine has been evaluated in various studies as a replacement for cyclophosphamide, in combination with busulfan, aiming to decrease non-relapse mortality while maintaining the immunosuppressive and anti-cancer efficacy of cyclophosphamide[2]. Following confirmation of this hypothesis[3], fludarabine is currently used in various conditioning regimens, ranging from nonmyeloablative[4], to reduced intensity and myeloablative regimens.

Fludarabine is dosed based on body-surface-area (BSA) and intravenously administered as a monophosphate prodrug (F-ara-AMP). It is very rapidly fully converted to the circulating metabolite F-ara-A, as which it is distributed intracellularly. Subsequently, intracellular phosphorylation takes place to the active metabolite fludarabine as the intracellular active moiety F-ara-ATP (FluTP) which is built into the DNA and RNA, thereby inhibiting DNA/RNA-synthesis. This leads to apoptosis in both chronic lymphocytic leukemia cells[5] and (with different susceptibility) in cell types targeted in the HCT setting[6,7].

As FluTP is the active metabolite, intuitively it is the form of interest for pharmacokinetic (PK) studies. However, during conditioning only a limited number of target cells can be acquired, especially when anti-thymocyte globulin (rabbit anti-thymocyte globulin (rATG)) is administered prior to fludarabine in the conditioning, complicating accurate quantification of the intracellular FluTP. Therefore, ex-vivo quantification of FluTP accumulation in pre-treatment samples has been proposed in HCT recipients[8], though a relationship between this accumulation and outcome was not found[4]. Given these complexities and an apparent correlation between F-ara-A concentration and FluTP formation in cells[9], the freely circulating F-ara-A has been used primarily for PK analyses. During the phase I trial for fludarabine, triphasic first-order kinetics were found for F-ara-A. The main known route of elimination is through the kidney and indeed a correlation between clearance from the central compartment and estimated glomerular filtration rate (eGFR) has been found[10].

As this study was performed in adults with chronic lymphocytic leukemia, several HCT specific PK-studies were performed to further explore fludara-

bine PK in this setting[9,11–14]. However, so far none has led to a harmonized dosing regimen for both children and adults, which takes both renal function and body size into account. Previous study results could not be extrapolated to the general population, due to the use of non-population based methods[11,12], limited sample size[13,14], or containing only pediatric data[9]. This causes most centers to still use BSA-based dosing, although a PK rationale for this is lacking.
Therefore, this study aims to build a population PK model, using a very large heterogeneous dataset of both pediatric and adult HCT recipients. As such, this study can provide a rational base for optimal and harmonized dosing regimens for patients of all ages in this setting, whilst taking renal function into account.

7.3 Methods

Patients

A retrospective PK analysis was performed with data from patients who received myeloablative conditioning before HCT, between May 2010 and January 2017 at the UMC Utrecht and of whom PK samples were available. No restrictions were applied for co-morbidities, age and indication for HCT. Patients were included after written informed consent was acquired. Ethical approval by the institutional medical ethics committee of the UMC Utrecht was obtained under protocol number 11/063.

Procedures

The conditioning regimen consisted of 4 days of chemotherapy (given from day -5 to day -2 relative to HCT) consisting of a 1-hour-infusion of fludarabine-phosphate directly followed by a 3-hour-infusion of busulfan (Busilvex, Pierre Fabre). A 1-hour infusion of clofarabine preceded fludarabine infusion in children with malignancies. rATG was added in the unrelated donor HCT setting: 4-hour infusions at 4 consecutive days from day -9 (10 mg/kg <30 kg, 7,5 mg/kg >30 kg) for children and 4 times a 12-hour infusion from day -12 (6 mg/kg) for adults.
Patients received either a cumulative dose of 160 mg/m^2 of fludarabine-phosphate or 40 mg/m^2 fludarabine-phosphate combined with 120 mg/m^2 clofarabine. Intravenous busulfan, was targeted to a myeloablative cumulative 4-day exposure of 90 mg*h/L or 30 mg*h/L for Fanconi anemia patients (expressed as area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($AUC_{t_0-\infty}$)). For patients receiving rATG,

clemastine, paracetamol, and 2 mg/kg prednisolone (with a maximum of 100 mg) were given intravenously prior to rATG infusion.

Pharmacokinetic samples and analyses

Concentrations of the circulating metabolite of fludarabine (F-ara-A, hereafter referred to as fludarabine) were analyzed in PK samples taken for routine busulfan therapeutic drug monitoring (TDM) according to local protocol. Quantification of fludarabine concentrations was performed using a liquid chromatography mass spectrometry method validated according to FDA and EMA guidelines as described previously, with a lower limit of quantification of 0.001 mg/L[15]. In the TDM protocol, plasma samples were drawn on the first or second day of conditioning. If considered necessary for TDM purposes, samples were drawn on following days as well. Additional samples were taken at the final day of conditioning. In general, plasma samples were taken at 4, 5, 6, and 7 hours, after the end of fludarabine infusion. For a subset of patients, additional samples were collected from 7 to 24 hours post-infusion. From January 2016 onwards, additional samples were collected between the end of fludarabine infusion and the start of busulfan infusion: 15-45 minutes after the end of fludarabine infusion.
A population approach based on non-linear mixed-effects modeling was applied[16], using the software package NONMEM (version 7.3.0, Icon, Hanover, MD, USA). Pirana (version 2.9.5) and R (version 3.3.3) were used for workflow management and data handling and visualization, respectively.[17,18] The first order conditional estimation option with interaction between random and residual error components (FOCE-I) as implemented in NONMEM, was used as estimation method.

Pharmacokinetic model building procedure

For the structural model 1-, 2- and 3-compartment models with first-order kinetics were tested. inter-individual variability (IIV) was assumed to follow a log-normal distribution and was therefore implemented into the model as (equation 1):

$$P_i = P_{pop} * e^{\eta_i} \tag{1}$$

where P_i depicts the individual or post-hoc value of the parameter for the i^{th} individual, P_{pop} the population mean for the parameter, and η_i the empirical Bayes estimate of IIV for the i^{th} individual, sampled from a normal

distribution with a mean of zero and a variance of ω^2. Inter-occasion variability (IOV) was implemented similarly, with each dose and subsequent sampling defined as a separate occasion. This variability was evaluated for all parameters to diagnose potential time-dependent trends and to allow for random unaccounted variability between dosing moments.
Residual error was evaluated as a proportional or additive error, or as a combination of both (equation 2).

$$P_{obs} = P_i * (1 + \epsilon_{proportional}) + \epsilon_{additive} \tag{2}$$

where P_{obs} is the observed value, $\epsilon_{proportional}$ the proportional-, and $\epsilon_{additive}$ the additive-error component. Residual error components are sampled from a normal distribution with mean of zero and variance σ.

Covariate model

Following development of the structural and stochastic PK model, potential predictors (covariates) for variability in PK parameters were evaluated. Assessed covariates included patient-related (body size: i.e. actual body weight (BW)), fat-free mass (FFM), BSA; other: age, renal function) and treatment-related (serotherapy, additional co-conditioning agents) factors. FFM was calculated using the equation developed by McCune et al (equation 3).[19] and BSA according to the method developed by Dubois et al.[20]

$$FFM = 9270 * \frac{BW}{\frac{BW}{HT^2} * S_{sex} + C_{sex}} \tag{3}$$

where HT corresponds to height in meters and S_{sex} (m^2/kg) and C_{sex} (dimensionless) are constants that change upon sex. S_{sex} takes values of 216 and 244 and C_{sex} 6680 and 8780, for males and females respectively.

Continuous covariates were evaluated using both a linear and a power function (equation 4 and 5):

$$P_i = P_{pop} + (Cov_i - Cov_{typical}) * l \tag{4}$$

$$P_i = P_{pop} * (\frac{Cov_i}{Cov_{typical})}^{p} \tag{5}$$

where Cov_i is the covariate value for the i^{th} individual, and $Cov_{typical}$ is the typical or median value for the covariate in the population. The estimated parameters are l and p for the linear and power function, respectively.

Binary categorical covariates were tested by using equation 6:

$$P_i = P_{pop} * (1 + P_{cov}) \tag{6}$$

where P_{cov} is the estimated proportional factor with which the parameter changes for a specific covariate value. To implement body size descriptors on PK parameters, initially, standard allometric scaling was applied using equation 4 with p fixed at 0.75 (BW/FFM) or 1 (BSA) for clearances and 1 for distribution volumes (BW/FFM/BSA). Alternative body size measures (FFM, BSA) were tested as a replacement of BW. Empirical estimation of the exponents was tested for the optimal body size descriptor and was only preferred, if this resulted in a relevant improvement of the model fit and when the estimated parameters were markedly different from the theoretical values.
Renal function was evaluated as a covariate, since fludarabine is predominantly eliminated renally[15, 10]. As creatinine levels were not measured daily, the mean value of available individual creatinine values between day -7 and day 0 prior to infusion were used. Subsequently, renal function (as eGFR) was calculated using the Cockroft and Gault equation, which takes age into account[21]. eGFR for patients below the age of 17 for women and 14 for men was calculated using the Schwartz equation.[22] To prevent physiologically implausible high eGFR values, these were capped to a maximum. Maximum eGFR was set at 140 mL/min/1.73m^2, but was assumed to increase to this value from birth till 1.5 years of age starting at 35 ml/min/1.73 m^2 (25% of maximum value) and to decline by 8 mL/min/1.73m^2 per decade after the age of 30, as suggested earlier[23].

Pharmacokinetic model evaluation

The structural and covariate model with corresponding estimates had to be scientifically and biologically plausible. To investigate parameter-covariate relationships, covariates were plotted versus empirical Bayes estimates of IIV. Trends in these plots indicated potential relationships.
Addition of a parameter had to result in a significant improvement in model fit. This was evaluated using the objective function value (OFV), equal to minus twice the log-likelihood, which is assumed to follow a chi-squared

distribution. In hierarchical models, an OFV change (ΔOFV) of -3.84 corresponds with a p-value of 0.05 for addition of one parameter (i.e. 1 degree of freedom). Covariates were evaluated for significance using forward inclusion and backward elimination[11]. A significance level of $p < 0.005$ (-7.9 ΔOFV) was used for the forward inclusion, and $p < 0.001$ (-10.8 ΔOFV) for backward elimination. In addition, inclusion of a covariate had to result in a substantial decline in unexplained IIV[24].
A visual inspection of model performance was done through standard goodness-of-fit plots. Examples of these plots are observed concentrations plotted versus individual and population predicted concentrations, and conditional weighted residuals (CWRES) versus time and observed concentrations[25]. Particular emphasis was given to goodness-of-fit plots stratified for age to assess potential age-related misspecification.
Furthermore, relative standard errors (RSEs) (as estimated from the $COVARIANCE step) of all parameters and shrinkage of random and residual error components were assessed[11]. Values below 30% for shrinkage (IIV and residual error) and RSE (all parameters) were deemed acceptable. Finally, the condition number was calculated after each addition of a parameter, to check for over-parameterization, where a value below 1000 was accepted[26].
Several evaluation techniques were performed, all in accordance with EMA and FDA guidelines for population PK analyses[33, 27]. A non-parametric bootstrap evaluation (1000 samples) was performed to assess parameter precision. In addition, the normalized prediction distribution error (NPDE) was evaluated. Discrepancies between the final model and 1000 simulations of the model were evaluated, taking into account the correlation between observations in the same individual and the predictive distribution[28].
To assess the simulation properties, prediction-corrected visual predictive checks (VPCs) were created to assess the predictive performance of the final model as compared to the observed concentrations. The prediction-corrected VPC allows for variability in dosing[29]. In this analysis, the observed concentration data and its median and 95% confidence intervals were compared to 95% confidence intervals of the predicted mean, 2.5^{th} and 95^{th} percentile, derived from 1000 model simulations.

7.4 Results

Patients and Samples

A total of 258 patients were included with a median age of 18 years. Of these patients 197 received a cumulative fludarabine dose of 160 mg/m_2 and 61

received a cumulative dose of 40 mg/m_2 in combination with clofarabine. Of the obtained 2605 samples, none were below the lower limit of quantification (LLOQ). The concentration-time data was divided over 596 administered doses, of which 116 (19%) contained peak samples (<3 hours) and 117 (20%) contained trough samples (>8 hours). Detailed patient characteristics are shown in table 1.

Table 1: Patient characteristics

	Children and adolescents (≤20 years)	**Adults (>20 years)**	**Total**
Available data			
No. of patients (no.(%))[a]	134 (52%)	124 (48%)	258 (100%)
Total no. of pharmacokinetic samples (no.(%))$_a$	1384 (53%)	1221 (47%)	2605 (100%)
No. of samples per patient[b]	10 (3–19)	10 (3 - 16)	10 (3 - 19)
Patient characteristics			
Female sex (no.(%))[c]	52 (39%)	46 (37%)	98 (38%)
Age at transplantation (years)[d]	9.6 (0.2–20, 2.7-14)	54 (22–74, 40-62)	18 (0.2 – 74, 40-62)
Actual bodyweight (kg)[d]	29 (4.3–96, 14-52)	78 (47-130, 65-89)	60 (4.3-130, 28-78)
Renal function[d] (eGFR: ml/min/1.73m2)d	140 (40–140, 110-140)	110 (25-140, 93-130)	120 (25-140, 100-140)
Indication for transplantation (no.(%))[c]			
Benign-disorder	65 (49%)	4 (3.2%)	69 (27%)
Leukemia	69 (51%)	48 (39%)	117 (45%)
Lymphoma	0 (0%)	17 (14%)	17 (6.6%)
Myelodysplastic syndrome	0 (0%)	32 (26%)	32 (12%)
Plasma cell disorder	0 (0%)	23 (19%)	23 (8.9%)
Transplant cell source (no.(%))[c]			
Peripheral blood	2 (2%)	109 (88%)	111 (43%)
Cord blood	94 (70%)	12 (10%)	106 (41%)
Bone marrow	38 (28%)	3 (2%)	41 (16%)

eGFR = estimated glomerular filtration rate

[a]Percentage of full data

[b]Displayed as median (range)

[c]Percentage of sub-group data within selected population (adults/children/total)

[d]Displayed as median (range, interquartile range)

[e]Calculated as described in methods section 7.3

Structural and Sochastic model

BW was a priori included as a covariate using a power function (equation 4) on all (inter-compartmental) clearances and volumes of distribution during model development. The exponents for body weight on volume of distribution and clearance were fixed to 1.0 and 0.75, respectively, prior to covariate analyses.

A 3-compartment model best described the data. In addition, both the VPC and goodness-of-fit plots showed substantial misspecification for the 2-compartment model, which was absent in the 3-compartment model. The 3-compartment model was parameterized in terms of volume of distribution of the central (V_1), peripheral (V_2) and second peripheral (V_3) compartment and clearance from the central compartment as well as inter-compartmental clearance between V_1 and V_2 (Q_2) and V_1 and V_3 (Q_3).
IIV was added on V_1 as well as clearance from the central compartment and inclusion of IOV was significant for both these parameters as well. Inclusion of IOV and IIV on Q_2, Q_3, V_2, and V_3, led to improved model fit, however, this model was highly over-parameterized (condition number >1000) and unstable (sensitive to initial estimates). Upon visual inspection of the random effects and estimation of the covariance matrices, it was shown that the random effects on volume (V_1, V_2, V_3) and clearances (Cl, Q_2, Q_3) were highly correlated. Therefore, single random effects (both IIV and IOV) were estimated for V_1, V_2 and V_3 and for clearance, Q_2, Q_3, respectively. This approximation was adequate to describe the observed variability and provided stable and reproducible parameter estimates. (table 2).

Covariate model

Figure 1A-B depict the variability in observed fludarabine concentrations over time (figure 1A) and total exposure (observed $AUC_{t_0-\infty}$, figure 1B) stratified BSA-adjusted dose (10 and 40 mg/m^2). In both the low (40 mg/m^2) and high (160 mg/m^2) dose groups, concentrations over time after dose were highly variable, resulting in a large range for $AUC_{t_0-\infty}$ (2.7-12, 10-66 mg*h/L for 10 and 40 mg/m^2 respectively). With a median $AUC_{t_0-\infty}$ of 21 mg*h/l (range: 5.7-42) and 26 mg*h/l (range: 13-65) at a 40 mg/m^2 dose for children and adults respectively, these values are slightly higher than those reported for children by Ivaturi et al (median 18 mg*h/L) and adults by Long-Boyle et al (median 20 mg*h/L) at the same cumulative dose of 160 mg/m^2.

Figure 1C depicts the exposures at different weight quartiles. Low weights correlate to low exposures and high weights to high exposures, indicating that BSA is not a sufficient body-size descriptor for fludarabine clearance. Figure 1D depicts exposures at subgroups stratified for eGFR values according to the classification of the National Kidney Foundation: healthy renal function (eGFR >90 ml/min/1.73 m^2), mild (eGFR 60-90 ml/min/1.83 m^2: n=37) and moderate (eGFR 30-60 ml/min/1.83 m^2:

Table 2: Parameter estimates of the pharmacokinetic model

	Estimate (95% confidence intervala)	**Shrinkage**	**Bootstrap outcomesb**	
			Median	2.5th and 97.5th percentiles
Structural model				
$Cl = (Cl_{70kg-non-renal} + eGFR * Slope_{pop}) * \frac{BW}{70kg}^{0.75}$				
$Cl_{70kg-non-renal}$ (L/h)	3.2 (1.6-4.9)		3.2	2.0-4.3
$Slope_pop$	0.78 (0.57-1.0)		0.79	0.65-0.95
$V_{1,pop} = V_{1,70kg} * \frac{BW}{70kg}^{1}$				
$V_{1,70kg}$ (L)	39 (33-45)		39	33-46
$V_{2,pop} = V_{2,70kg} * \frac{BW}{70kg}^{1}$				
$V_{2,70kg}$ (L)	20 (17-24)		21	16-28
$V_{3,pop} = V_{3,70kg} * \frac{BW}{70kg}^{1}$				
$V_{3,70kg}$ (L)	50 (41-58)		51	43-64
$Q_{2,pop} = Q_{2,70kg} * \frac{BW}{70kg}^{0.75}$				
$Q_{3,70kg}$ (L)	8.6 (6.8-10)		8.8	7.3-11
$Q_{3,pop} = Q_{3,70kg} * \frac{BW}{70kg}^{0.75}$				
$Q_{3,70kg}$ (L)	3.8 (2.9-4.6)		3.7	2.4-5.0
Random variability				
Inter-individual variability on CL,Q_2,Q_3 (%)	23 (15-31)	7%	23	20-26
Inter-individual variability on V_1,V_2,V_3 (%)	48 (36-60)	15%	47	37-57
Inter-occasion variability on CL,Q_2,Q_3 (%)	12 (9.6-14)	28%	12	11-14
Inter-occasion variability on V_1,V_2,V_3 (%)	31 (18-44)	37%	32	25-38
Proportional residual error (%)	6.3 (4.3-8.3)	20%	6.2	5.4-7.2

CI=confidence interval, pop=population, i=individual, BW=actual body weight, V_1=central volume of distribution, V_2=peripheral volume of distribution, V_3=second peripheral volume of distribution, Q_2=intercompartmental clearance V_1–V_2, Q_3=intercompartmental clearance V_1–V_3

Population estimates $V_{1,70kg}$, $V_{2,70kg}$, $V_{3,70kg}$, Cl_{70kg}, $Q_{2,70kg}$, $Q_{3,70kg}$ correspond to a subject weighing 70 kg and are adjusted to an individual value, according to the corresponding parameter formula in the table

[b]Displayed as median (range)

[c]Percentage of sub-group data within selected population (adults/children/total)

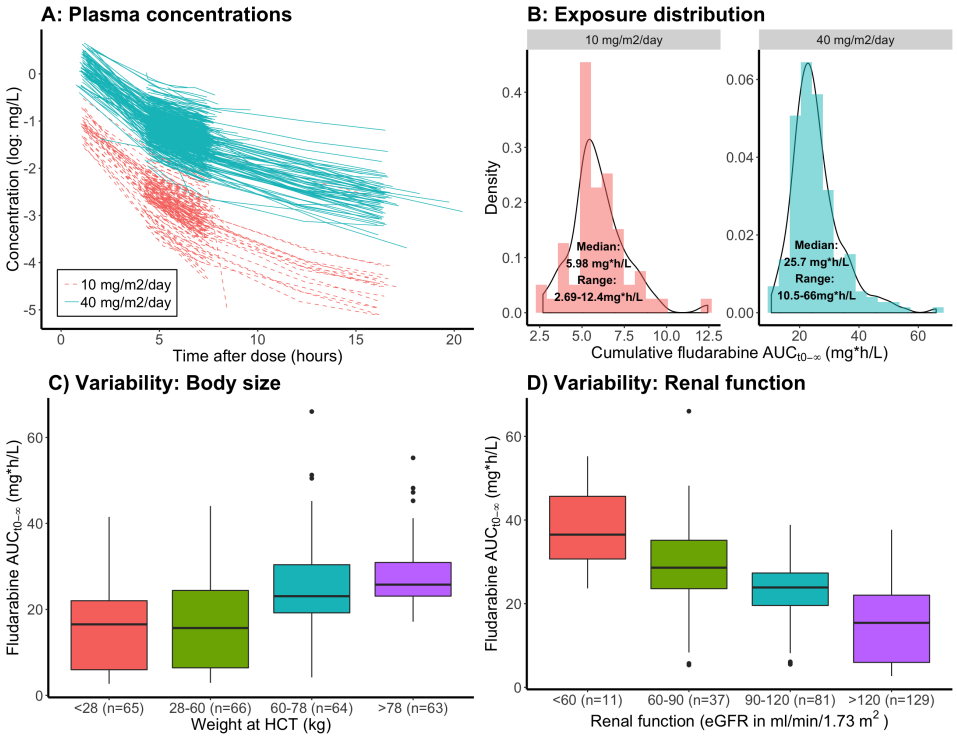

Figure 1: Exposure variability and covariates predicting variability A) Fludarabine plasma concentrations versus time after last dose on a logarithmic scale. Each line corresponds to a single dose, stratified to dose. B) Histogram (grey area) and density plot (black solid line) of the observed $AUC_{t_0-\infty}$. $AUC_{t_0-\infty}$ of patients receiving low dose (40 mg/m^2) were normalized to 160 mg/m^2. C) Boxplots of $AUC_{t_0-\infty}$ per weight quartile of observed $AUC_{t_0-\infty}$. D) Boxplots of observed $AUC_{t_0-\infty}$ stratified for renal function.

n=11) renal impairment,[6]. Healthy renal function was further subdivided in above (n=129) and below (n=81) the median (eGFR 120 ml/min/1.73 m^2). Herein it seems that eGFR and fludarabine clearance are correlated. Therefore, non-renal clearance was differentiated from renal clearance by adding eGFR with an estimated slope. BW, IIV and IOV were implemented on the total clearance[30], as illustrated in equation 7:

$$Cl = (Cl_{70kg-non-renal} + eGFR * Slope_{pop}) * \frac{BW}{70kg}^{0.75} * e^{IIV+IOV} \quad (7)$$

where IIV represents the IIV for clearance, Q_2 and Q_3; eGFR is based on individual creatinine levels and used in this equation as l/h/kg; and $Slope_{pop}$ is a unit-less estimate, representing the fraction of eGFR accounting for renal clearance of fludarabine. Addition of eGFR resulted in a statistically significant improvement in fit (OFV(ΔOFV) -172, 1 degree of freedom, $p < 0.001$) and IIV on clearance reduced from 34 to 23%. The $Slope_{pop}$ was estimated at 0.78 (RSE 11%:) indicating that there is limited renal resorption of fludarabine.
Use of alternative body size descriptors FFM and BSA, did not improve the model fit (ΔOFV of +60 and +68 respectively). Estimation of the allometric exponents for volume or clearance resulted in values very close to 0.75 and 1.0, respectively, and did not result in a relevant improvement in model fit. Therefore, the fixed exponents were kept in the model. After inclusion of BW and eGFR, no trends were visible in the plots of empirical Bayes estimates of IIV's versus eGFR, BW.
No other covariates, such as co-adminstration of clofarabine or rATG, could be identified. Importantly, no effect of age could be identified on any of the PK parameters.

Model evaluation

The final estimates and the results of the bootstrap analysis are shown in table 2. The median, 2.5^{th} and 97.5^{th} percentiles of bootstrap estimates are in line with those of the original data.
Age- and renal function-stratified goodness-of-fit plots (figure 2) generally demonstrated accurate population and individual predictions. Population predictions for children with a renal function below 60 ml/min/1.73m^2 seemed a bit off, though this group consisted of only 4 patients.

No trends were observed for CWRES versus time after dose (figure 3A), predicted concentrations (figure 3B), or covariates renal function (figure 3C)

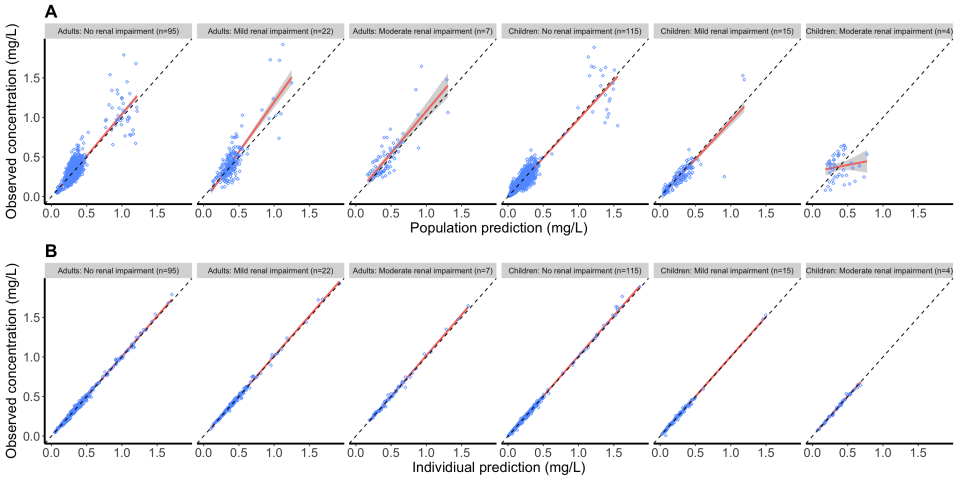

Figure 2: Goodness of fit plots for the final pharmacokinetic model stratified by age and renal function. Panel A and B depict the population- and individual predictions, respectively, versus observed concentrations, stratified for age (</≥ 20 years) and renal function (eGFR >90, 90-60, <60 ml/min/1.73 m_2). Blue open circles represent the observations and the solid red line is a local regression fit of these values. Dashed lines depict the line of unity.

and weight (figure 3D). The NPDE-analysis showed a normal distribution, and no trends were observed in the NPDEs versus time or predictions (data not shown).

The VPC showed that the median and 95% confidence interval of the observed data were in line with those from the simulation-based predictions from the model for all strata (figure 4), but the de children with moderate renal impairment (<60 ml/min/1.73m^2). The median of the observations is slightly higher than predicted indicating an over-prediction of clearance for this group. All 4 children in this group were at an age <0.5 years, indicating that maturation for very young children possibly is not well accounted for in the model.

We performed a sensitivity analysis to investigate potential bias caused by the imbalanced sampling design. This analysis showed that bias in clearance was negligible (data not shown).

7.5 Discussion

To our knowledge, this is the first PK model developed on the basis of a large and diverse dataset (n=258) including both adult and pediatric patients. In addition, this study covered the vast majority of HCT indications: acute

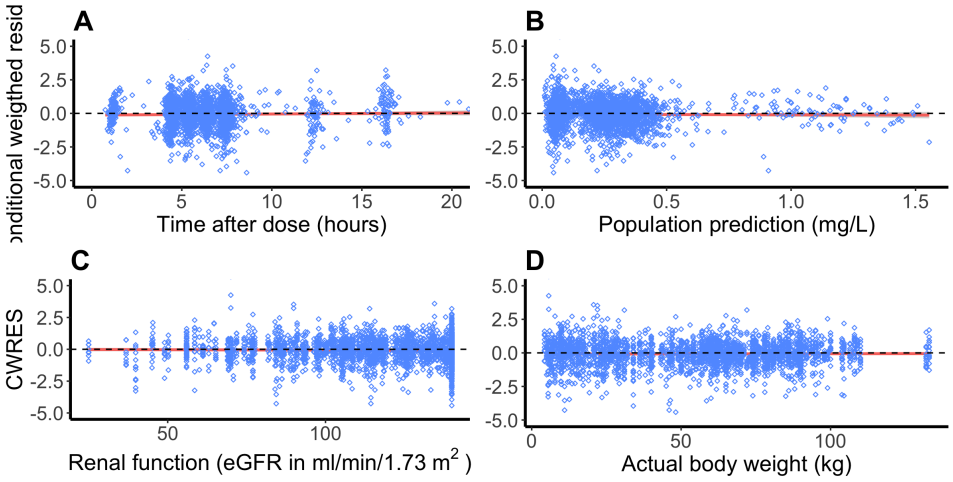

Figure 3: Conditional weighted residuals versus time after dose, population predictions, and covariates. Panel A, B, C and D depict the CWRES versus time after dose, the population predictions, renal function (eGFR) and actual bodyweight, respectively. Black open circles represent the CWRES values and the solid grey line is a local regression fit of these values. Dashed lines depict the zero-line.

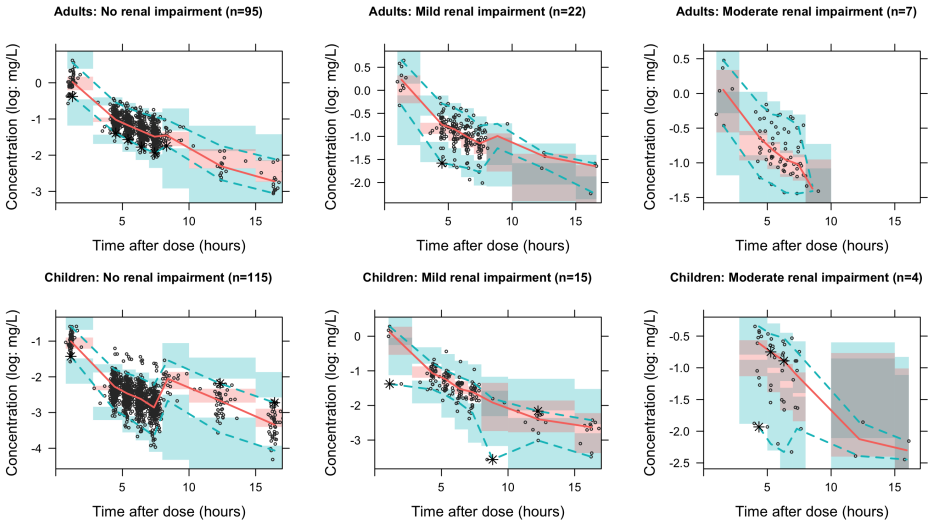

Figure 4: Stratified prediction-corrected visual predictive check. Black lines depict the observed median (solid) and 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles (dashed) concentrations. Dark- and light-grey areas represent 95% prediction intervals of the simulated mean and the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles, respectively. Round dots represent observations. Asterisks highlight observed percentiles outside the prediction area. Increased median concentrations after 8 hours are caused by the prediction correction. Subjects of whom samples for these bins were available had higher actual body-weights.

leukemia, lymphoma, plasma cell disorders, myelodysplastic syndrome and a variety of benign disorders (autoimmune diseases, immune deficiencies, bone-marrow failure, and metabolic diseases). This allowed for a unique platform to quantify fludarabine PK for all HCT populations.

Fludarabine PK was best described with a 3-compartment model. This is in line with the data of the phase I study[10]. In contrast, fludarabine plasma PK in two other population PK studies was described with a 2-compartment model[9,13]. However, these analyses had smaller sample sizes (n=54 / n=133) and included only children. In addition, no formal testing of 3-compartmental kinetics was mentioned.

Allometric scaling of all parameters using BW was found to best account for differences in body-size and, after inclusion of eGFR, no body-size-independent effect of age could be identified. Other studies did not compare body-size descriptors but rather implemented either allometric scaling to BW[9] or BSA-adjusted PK[13,14] *a priori*. Given the evidence supporting allometric theory over BSA-adjusted PK[31,32] allometry-based adjustment is preferred. In addition, we found that fludarabine BSA-based dosing led to major under- and over-prediction of exposures at high and low BW, respectively.

Similar to Ivaturi et al, eGFR was included using a body size adjusted method. This method has the advantage of reflecting solely renal function, while absolute eGFR strongly correlates to body-size. Using the BW-, rather than BSA-adjusted[9] eGFR, allowed separation of a renal and non-renal fraction of clearance. In addition, the estimated slope now represents the fraction of eGFR accounting for renal clearance. This slope was estimated at 0.78 implicating that renal resorption of fludarabine occurs. However, this is based on the assumption that eGFR accurately reflects actual glomerular filtration rate. Furthermore, the use of the Schwartz- and Cockroft-Gault-equation as well as the age-dependent capping of eGFR, may have impact on the relationship between actual glomerular filtration rate and eGFR. As seen in the VPC the maturation of renal function might not be properly accounted for in the current model, thus careful monitoring might be advised for the very young children.

Interestingly, the fludarabine label[33] only indicates a dose-reduction (up to 50%) when eGFR is below 70 ml/min/1.73 m^2. Though this 50% dose reduction is supported by our findings, an eGFR below 120 ml/min/1.73m^2 is already associated with a substantial decrease in fludarabine clearance (>25%) and concomitant high exposures. Furthermore, the decrease in clearance is a gradual process. Therefore, a dosing algorithm which takes eGFR into account, similar to the clinically applied algorithm for carbo-

platin, may be more appropriate for fludarabine. Such an algorithm could be an equation directly derived from the model predicted clearance.
Applicability of such an algorithm depends on the diversity of the underlying dataset, which was sufficient in our study regarding age and indication, but less regarding co-conditioning agents (busulfan, +/- clofarabine). This has the disadvantage of not being able to quantify possible PK interactions with a variety of conditioning agents. Given the predominant renal elimination, such interactions are unlikely and were indeed found to be absent with clofarabine in these data. PK extrapolation of dosing algorithms may therefore be well translatable to other conditioning regimens, though the target exposure might differ. The advantage of this homogeneously treated cohort is the possibility to find an optimal PK exposure for fludarabine in the widely applied busulfan/fludarabine regimen.
In conclusion, given the observed high variability in exposure, the current BSA-based dosing regimen without taking eGFR into account, may not be appropriate. The current analysis provides a rational base for a harmonized optimal dosing regimen for all age groups in HCT.

References

1. Gyurkocza, B. & Sandmaier, B. M. Conditioning regimens for hematopoietic cell transplantation: one size does not fit all. *Blood* **124,** 344–53 (2014).
2. Ben-Barouch, S., Cohen, O., Vidal, L., Avivi, I. & Ram, R. Busulfan fludarabine vs busulfan cyclophosphamide as a preparative regimen before allogeneic hematopoietic cell transplantation: systematic review and meta-analysis. *Bone Marrow Transplant* **51,** 232–40 (2016).
3. Rambaldi, A., Grassi, A., Masciulli, A., Boschini, C., Mico, M. C., Busca, A., *et al.* Busulfan plus cyclophosphamide versus busulfan plus fludarabine as a preparative regimen for allogeneic haemopoietic stem-cell transplantation in patients with acute myeloid leukaemia: an open-label, multicentre, randomised, phase 3 trial. *Lancet Oncol* **16,** 1525–1536 (2015).
4. McCune, J. S., Mager, D. E., Bemer, M. J., Sandmaier, B. M., Storer, B. E. & Heimfeld, S. Association of fludarabine pharmacokinetic/dynamic biomarkers with donor chimerism in nonmyeloablative HCT recipients. *Cancer Chemother Pharmacol* **76,** 85–96 (2015).
5. Plunkett, W., Huang, P. & Gandhi, V. Metabolism and action of fludarabine phosphate. *Semin Oncol* **17,** 3–17 (1990).
6. Consoli, U., El-Tounsi, I., Sandoval, A., Snell, V., Kleine, H. D., Brown, W., *et al.* Differential induction of apoptosis by fludarabine monophosphate in leukemic B and normal T cells in chronic lymphocytic leukemia. *Blood* **91,** 1742–8 (1998).
7. McCune, J. S., Vicini, P., Salinger, D. H., O'Donnell, P. V., Sandmaier, B. M., Anasetti, C., *et al.* Population pharmacokinetic/dynamic model of lymphosuppression after fludarabine administration. *Cancer Chemother Pharmacol* **75,** 67–75 (2015).
8. Woodahl, E. L., Wang, J., Heimfeld, S., Sandmaier, B. M., O'Donnell, P. V., Phillips, B., *et al.* A novel phenotypic method to determine fludarabine triphosphate accumulation in T-lymphocytes from hematopoietic cell transplantation patients. *Cancer Chemother Pharmacol* **63,** 391–401 (2009).
9. Ivaturi, V., Dvorak, C. C., Chan, D., Liu, T., Cowan, M. J., Wahlstrom, J., *et al.* Pharmacokinetics and Model-Based Dosing to Optimize Fludarabine Therapy in Pediatric Hematopoietic Cell Transplant Recipients. *Biol Blood Marrow Transplant* **23,** 1701–1713 (2017).

10. Malspeis, L., Grever, M. R., Staubus, A. E. & Young, D. Pharmacokinetics of 2-F-ara-A (9-beta-D-arabinofuranosyl-2-fluoroadenine) in cancer patients during the phase I clinical investigation of fludarabine phosphate. *Semin Oncol* **17,** 18–32 (1990).

11. Bonin, M., Pursche, S., Bergeman, T., Leopold, T., Illmer, T., Ehninger, G., *et al.* F-ara-A pharmacokinetics during reduced-intensity conditioning therapy with fludarabine and busulfan. *Bone Marrow Transplant* **39,** 201–6 (2007).

12. Long-Boyle, J. R., Green, K. G., Brunstein, C. G., Cao, Q., Rogosheske, J., Weisdorf, D. J., *et al.* High fludarabine exposure and relationship with treatment-related mortality after nonmyeloablative hematopoietic cell transplantation. *Bone Marrow Transplant* **46,** 20–6 (2011).

13. Mohanan, E., Panetta, J. C., Lakshmi, K. M., Edison, E. S., Korula, A., Fouzia, N. A., *et al.* Population pharmacokinetics of fludarabine in patients with aplastic anemia and Fanconi anemia undergoing allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplant* **52,** 977–983 (2017).

14. Salinger, D. H., Blough, D. K., Vicini, P., Anasetti, C., O'Donnell, P. V., Sandmaier, B. M., *et al.* A limited sampling schedule to estimate individual pharmacokinetic parameters of fludarabine in hematopoietic cell transplant patients. *Clin Cancer Res* **15,** 5280–7 (2009).

15. Punt, A. M., Langenhorst, J. B., Egas, A. C., Boelens, J. J., van Kesteren, C. & van Maarseveen, E. M. Simultaneous quantification of busulfan, clofarabine and F-ara-A using isotope labelled standards and standard addition in plasma by LC-MS/MS for exposure monitoring in hematopoietic cell transplantation conditioning. *J Chromatogr B Analyt Technol Biomed Life Sci* **1055-1056,** 81–85 (2017).

16. Sheiner, L. B., Rosenberg, B. & Marathe, V. V. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *J Pharmacokinet Biopharm* **5,** 445–79 (1977).

17. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

18. Keizer, R. J., van Benten, M., Beijnen, J. H., Schellens, J. H. & Huitema, A. D. Pirana and PCluster: a modeling environment and cluster infrastructure for NONMEM. *Comput Methods Programs Biomed* **101,** 72–9 (2011).

19. McCune, J. S., Bemer, M. J., Barrett, J. S., Scott Baker, K., Gamis, A. S. & Holford, N. H. Busulfan in infant to adult hematopoietic cell transplant recipients: a population pharmacokinetic model for initial and Bayesian dose personalization. *Clin Cancer Res* **20,** 754–63 (2014).

20. Du Bois, D. & Du Bois, E. F. The measurement of the surface area of man. *Arch Intern Med* **15,** 868–881 (1915).

21. Cockcroft, D. W. & Gault, M. H. Prediction of creatinine clearance from serum creatinine. *Nephron* **16,** 31–41 (1976).

22. Schwartz, G. J., Haycock, G. B., Edelmann C. M., J. & Spitzer, A. A simple estimate of glomerular filtration rate in children derived from body length and plasma creatinine. *Pediatrics* **58,** 259–63 (1976).

23. Cohen, E., Nardi, Y., Krause, I., Goldberg, E., Milo, G., Garty, M., *et al.* A longitudinal assessment of the natural rate of decline in renal function with age. *J Nephrol* **27,** 635–41 (2014).

24. Krekels, E. H., van Hasselt, J. G., Tibboel, D., Danhof, M. & Knibbe, C. A. Systematic evaluation of the descriptive and predictive performance of paediatric morphine population models. *Pharm Res* **28,** 797–811 (2011).

25. Hooker, A. C., Staatz, C. E. & Karlsson, M. O. Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. *Pharm Res* **24,** 2187–97 (2007).

26. Byon, W., Smith, M. K., Chan, P., Tortorici, M. A., Riley, S., Dai, H., *et al.* Establishing best practices and guidance in population modeling: an experience with an internal population pharmacokinetic analysis guidance. *CPT Pharmacometrics Syst Pharmacol* **2,** e51 (2013).

27. EMEA. Guideline on reporting the results of population pharmacokinetic analyses (2007).

28. Comets, E., Brendel, K. & Mentre, F. Computing normalised prediction distribution errors to evaluate nonlinear mixed-effect models: the npde add-on package for R. *Comput Methods Programs Biomed* **90,** 154–66 (2008).

29. Bergstrand, M., Hooker, A. C., Wallin, J. E. & Karlsson, M. O. Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. *AAPS J* **13,** 143–51 (2011).

30. Rhodin, M. M., Anderson, B. J., Peters, A. M., Coulthard, M. G., Wilkins, B., Cole, M., *et al.* Human renal function maturation: a quantitative description using weight and postmenstrual age. *Pediatr Nephrol* **24,** 67–76 (2009).

31. Anderson, B. J. & Holford, N. H. Mechanistic basis of using body size and maturation to predict clearance in humans. *Drug Metab Pharmacokinet* **24,** 25–36 (2009).

32. Germovsek, E., Barker, C. I., Sharland, M. & Standing, J. F. Scaling clearance in paediatric pharmacokinetics: All models are wrong, which are useful? *Br J Clin Pharmacol* **83,** 777–790 (2017).

33. *National kidney foundation Glomerular Filtration Rate* `https://www.kidney.org/atoz/content/gfr`. Accessed: 2018-07-23.

8 Fludarabine exposure in the conditioning prior to allogeneic hematopoietic cell transplantation predicts outcomes

Jurgen B Langenhorst, Charlotte van Kesteren, Erik M van Maarseveen, Thomas PC Dorlo, Stefan Nierkens, Caroline A Lindemans, Moniek A de Witte, Anna van Rhenen, Reinier Raijmakers, Marc Bierings, Jürgen Kuball*, Alwin DR Huitema*, Jaap Jan Boelens*

*contributed equally

Submitted (December 2018).

8.1 Abstract

Fludarabine is the most frequently used agent in conditioning regimens for allogeneic hematopoietic cell transplantation (HCT). Body-surface-area-based dosing leads to highly variable fludarabine plasma exposure, possibly leading to unpredictable HCT outcome. We studied the relation between fludarabine exposure and clinical outcomes after HCT.

A retrospective, pharmacokinetic-pharmacodynamic analysis was conducted with data from patients undergoing HCT with fludarabine (160 mg/m^2) as part of a myeloablative conditioning (busulfan targeted to an area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($AUC_{t_0-\infty}$) of 90 mg*h/L and rabbit anti-thymocyte globulin (6-10 mg/kg; from day -9/-12) between 2010 and 2016. Fludarabine exposure as $AUC_{t_0-\infty}$ was calculated for each patient using a previously published population pharmacokinetic model and related to 2-year-event-free survival (EFS) by means of (parametric)-time-to-event-models. Relapse, non-relapse mortality (NRM), and graft failure were considered events.

192 patients were included (68 benign and 124 malignant disorders). The optimal fludarabine exposure was determined as an $AUC_{t_0-\infty}$ of 20 mg*h/L. Estimated 2-year EFS in the group with optimal exposure was 67%. In the over-exposed group, EFS was lower (hazard-ratio (HR) 2.0, 95%-confidence interval (CI) 1.1-3.4, $p = 0.02$), due to higher NRM (HR 3.5, 95%-CI 1.7-7.2, $p < 0.001$) associated with impaired immune-reconstitution (HR: 0.43, 95% CI 0.26-0.70, $p < 0.001$). The risks of NRM and graft failure were increased in the under-exposed group (HR 3.4, 95% CI 1.2-9.5, $p = 0.02$; HR 5.7, 95% CI 1.3-25, $p = 0.02$; respectively). No relationship with relapse was found.

Fludarabine-exposure is a strong predictor of survival after HCT, stressing the importance of optimum Fludarabine dosing. Individualized dosing, based on weight and renal function or therapeutic drug monitoring to achieve optimal fludarabine exposure might improve survival.

8.2 Introduction

Fludarabine (Flu) is the most frequently used agent in conditioning regimens for allogeneic hematopoietic cell transplantation (HCT). HCT is a potentially curative, but high-risk, treatment for a variety of malignant and benign hematological disorders. Besides disease relapse (20-50%), transplant related mortality (10-40%) is of major concern[1]. Therefore, optimization of this procedure, which leads to improved safety of the therapy itself without affecting disease control, is urgently needed.
It has previously been shown that optimization of pharmacokinetic exposure of agents used in the conditioning regimen prior to HCT can be used to achieve this[2–5]. For busulfan, this has led to the introduction of therapeutic drug monitoring aiming at an optimal target exposure, which has been proven superior over fixed dosing in a randomized clinical trial[6]. Also for rabbit anti-thymocyte globulin (rATG), widely used as serotherapy in HCT, an optimal exposure and resulting optimized dosing regimen have been developed by us[2,3,7–9].
Flu is currently dosed based on body-surface-area (BSA) and intravenously administered as a monophosphate prodrug (F-ara-AMP). It is very rapidly fully converted to the circulating metabolite F-ara-A, which is mainly cleared by the kidney. Recently, we found more than a six-fold variability in F-ara-A plasma exposure, with current BSA-based dosing[10]. We hypothesize that such variability in chemotherapy exposure will also lead to a variable and unpredictable treatment outcome, as previously shown for other agents used in conditioning: e.g. busulfan and rATG. Therefore, we conducted a retrospective cohort analysis, where different Flu exposure measures were related to various clinical outcomes of HCT, such as graft-failure, non-relapse mortality (NRM), relapse, CD4+ T-cell reconstitution (IR), and survival.

8.3 Methods

Study design and patients

A retrospective, pharmacokinetic-pharmacodynamic analysis was performed with data from patients who received myeloablative conditioning before HCT, between May 2010 and January 2017 at the UMC Utrecht, the Netherlands. The conditioning regimen consisted of Flu-phosphate, busulfan (Busilvex, Pierre Fabre) and plus/minus rATG (Thymoglobulin, Sanofi Genzyme). No restrictions were applied for co-morbidities, age and indication for HCT. Combined Haploidentical+Cord grafts[11] were excluded and all other cell sources were accepted. Clinical data were collected prospectively and pa-

tients were included after written informed consent was acquired. Ethical approval by the institutional medical ethics committee of the UMC Utrecht was obtained under protocol number 11/063.

Table 1: Patient characteristics[a]

Age at transplantation (years)	36.2 (14-58, 0.23-74)
Sex	
Male	115 (60%)
Female	77 (40%)
Indication	
Benign	68 (35%)
Leukemia/Lymphoma	71 (37%)
MDS	30 (16%)
Plasma cell disorder	23 (12%)
Cell source	
PB: Full graft	24 (12%)
Cord blood	65 (34%)
Bone marrow	16 (8.3%)
PB: α-β depleted	87 (45%)
Human leukocyte antigen disparity[b]	
Matched	150 (78%)
Mismatched	42 (22%)
Conditioning	
Samples per patient	10 (8-12, 3-19)
Fludarabine $AUC_{t0-\infty}$ (mg*h/L)	24 (20-29, 10-66)
Busulfan $AUC_{t0-\infty}$ (mg*h/L)	96.1 (90-100, 59[c]-120)
Serotherapy	
Patients without ATG	20 (10%)
Patients with ATG	172 (90%)
$AUC_{t_x-\infty}$ of patients with ATG	14.5 (2.5-38, 0-270)
Creatinine clearance (ml/min/1.73m^2)	114 (94-130, 25-140)
Median Follow-up (days)[d]	639 (482-758)

[a]Categorical variables are displayed as N (%).Continuous variables are displayed as median (inter-quartile range, range) or [d]95% confidence interval

[b]In peripheral blood HCT, 90 were 10/10, 09/10, 16 were 8/8, 3 were 6/6). In cord blood HCT, 4 were 10/10, 4 were 9/10, 3 were 8/10, 21 were 6/6, 24 were 5/6, 9 were 4/6. In bone marrow HCT all were matched 10/10.

[c]Minima exclude Fanconi Anemia patients (n=3): for these patients indication-specific myeloablation was achieved with $AUC_{t0-\infty}$ of 31, 31, 25 mg*h/L

Procedures

Patients received a conditioning regimen containing Flu-phosphate at a cumulative dose of 160 mg/m^2 (4 days of 40 mg/m^2/day) and intravenous busulfan, targeted to a myeloablative cumulative exposure of 90 mg*h/L or 30 mg*h/L for Fanconi anemia patients (expressed as area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($\mathrm{AUC}_{t_0-\infty}$)). Flu-phosphate was given daily from day -5 to -2 relative

to transplantation as a 1-hour infusion before the daily busulfan infusion. rATG was added in the unrelated donor HCT setting: 4 consecutive days from day -9 (10 mg/kg <30 kg, 7.5 mg/kg >30 kg) for children and day -12 (6 mg/kg) for adults.
For patients receiving rATG, clemastine, paracetamol, and 2 mg/kg prednisolone (with max of 100 mg) were given intravenously prior to rATG infusion. Graft-versus-host disease (GvHD) prophylaxis in adults consisted of mycophenolate mofetil monotherapy, or cyclosporine A (CsA) with mycophenolate mofetil, in patients with- or without $\alpha - \beta$ T-cell depleted peripheral blood grafts respectively.
In pediatric patients, CsA and methotrexate 10 mg/m^2 at day +1, +3 and +6 were given in bone marrow recipients and CsA + prednisone 1 mg/kg/day in cord blood recipients until day +28, which was tapered in 2 weeks.
Target trough levels for CsA were 200–350 mg/L. Patients were prophylactically treated with acyclovir. All patients received partial gut decontamination with ciprofloxacin and fluconazole and *Pneumocystis jiroveci* pneumonia prophylaxis using co-trimoxazol. All cord blood transplantation patients received 10 mg/kg granulocyte colony-stimulating factor (Neupogen®) from day +7 after HCT until neutrophils were above 2000 cells/μL. Thorough immune monitoring was performed. After reaching a white blood cell count of $\geq 0.4 \cdot 10^9/L$, an extended T, B and NK lymphocyte subset analysis was performed. This was done at least every other week (2-4 times a month) up to twelve weeks post-HCT and monthly thereafter up to six months.

Outcomes

The main outcome of interest was 2-year event-free survival (EFS). Events considered were graft failure, relapse, and NRM. Relapse was defined as disease recurrence and NRM as death while in complete remission. Both graft rejection and non-engraftment were considered graft failure, where in case of non-engraftment, the time was set at day +50 day or time of follow-up/death, whichever occurred first.
Other outcomes of interest were overall survival (OS), IR, and overall mortality subdivided by cause of death. IR was defined as the first of 2 consecutive measurements above a value of $> 50/\mu L$ CD4+ T-cells within 100 days after HCT as described previously[3,12]. Causes of death were defined to be infection (either viral or bacterial), GvHD, multi-organ-failure (MOF), disease recurrence, and other disease-unrelated causes.

Fludarabine pharmacokinetics and optimum

A previously developed Flu pharmacokinetic model[10] was used to estimate measures of Flu exposure: the $AUC_{t_0-\infty}$ and the AUC from time of transplantation until infinity ($AUC_{t_x-\infty}$). These exposure measures were respectively chosen for having a previously found relationship with overall toxicity[5,13]/efficacy[14] and for having a hypothesized capability of in-vivo depletion of T-cells as previously shown for rATG[3,15–17]. For the most predictive exposure measure, the value corresponding to the highest EFS probability was selected as the pharmacokinetic target.

Statistical analysis

Duration of follow-up was defined as the time from HCT to last contact or death. Patients were censored at the date of last contact. Median time to follow-up was calculated using the reverse Kaplan-Meier method[18].
Factors, other than Flu exposure, considered to influence outcome included patient variables (age at transplantation, previous allogeneic-HCT, cytomegalovirus serostatus), conditioning variables (cumulative busulfan-exposure, rATG-exposure after transplantation), donor variables (graft-source: cord blood, bone marrow, peripheral blood; human-leukocyte-antigen-disparity; donor cytomegalovirus serostatus) and disease variables (plasma cell disorders, acute leukemia, lymphoma, benign hematological disorders).
To find the most predictive Flu exposure measure, $AUC_{t_0-\infty}$ and $AUC_{t_x-\infty}$ were quantitatively linked to the primary outcome measure (EFS) using a parametric time-to-event model. For this, an optimal hazard function was selected and the pharmacokinetic exposure measure with the strongest predictive power was identified based on the Akaike information criterion[19]. Using the same procedure, the optimal Flu exposure measure was compared to estimated glomerular filtration rate (eGFR) (as calculated previously[10]), as these variables were highly correlated and could not be independently introduced in the model.
Models for separate events were constructed similarly to the EFS model. The previously found most predictive exposure measure was included and models were expanded with other relevant covariates. Flu exposure was only kept in models, if this was proven to be a significant predictor ($p < 0.05$) by backward deletion.
To find the 2-year probability of any event (1-EFS) according to Flu exposure, the cumulative hazard for each event (2-years post-HCT) was estimated and the sum of these cumulative hazards was exponentiated to

compute EFS probability. A target exposure window was created by taking the exposure congruent with minimal 1-EFS probability and expanding it to ±25%, thus defining, an optimal, below-optimal, and above-optimal exposure group. Baseline characteristics of patients in these exposure groups were compared using the Kruskall-Wallis test for categorical covariates and one-way analysis of variance for continuous variables. Unadjusted probability of OS and EFS were computed with use of the Kaplan–Meier method and p-values were calculated using a 2-sided log-rank test. Unadjusted probability of, relapse, NRM, graft failure, IR, and mortality subdivided by cause were calculated using cumulative-incidence estimates and p-values were calculated using Gray's test.
Adjusted estimates for EFS and OS were computed using Cox-regression models. The adjusted incidence of relapse, NRM, graft failure, and IR was calculated using Fine-Gray models. Though hazard-ratio (HR) is not the proper term for measure of effect size in Fine-Gray models, the term was kept for readability purposes. Adjusted values correspond to the estimated probability given an equal distribution of model-included-covariates in all groups. P-values for categorical covariates in the regression models were calculated using Wald's test and for continuous covariates using the likelihood-ratio test.
Statistical analyses were performed using R version 3.2.4 with packages, flexsurvreg, cmprsk, survival and rms.

8.4 Results

Patients

A total of 192 patients (119 adults, 73 children) were available for the survival analysis after exclusion of 5 patients for receiving a combined haploidentical+cord graft. Patient characteristics are shown in table 1.

Fludarabine exposure and outcome

For the main outcome of interest (EFS), cumulative exposure to Flu from start of conditioning ($AUC_{t_0-\infty}$) was shown to be the best predictor, compared to both $AUC_{t_x-\infty}$ and eGFR. On the high end of Flu exposure, the incidence of NRM increased ($p < 0.001$, Figure 1A) and at lower exposures more graft failures were observed ($p = 0.04$, Figure 1B). Flu exposure had no significant influence on relapse ($p = 0.57$, Figure 1C). This resulted in a minimal event probability at a cumulative Flu exposure of 20 mg*h/L, with an estimated 2-year-EFS advantage of 43% and 11% compared to supra-

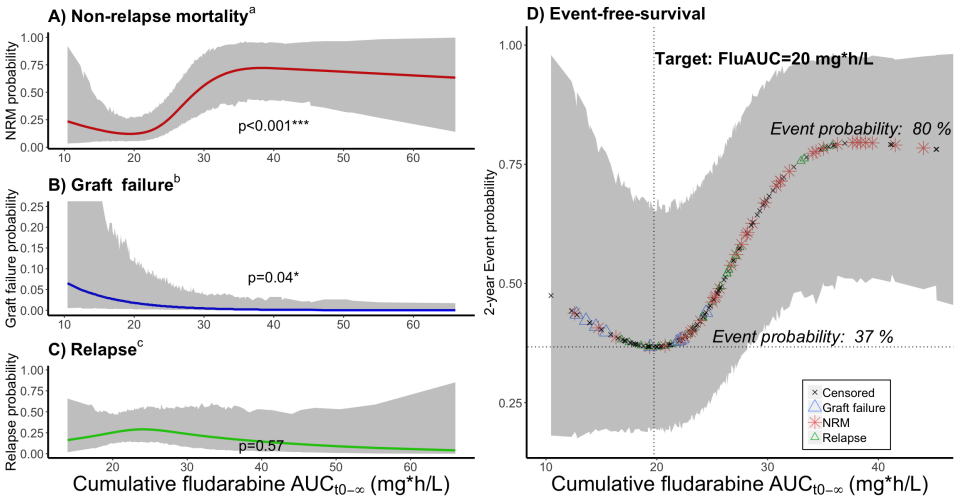

Figure 1: Fludarabine exposure-response relationship. Figures A-C: lines depict the estimated event probability (y-axis) at the given fludarabine AUC (x-axis), for A) non-relapse mortality B) graft failure C) relapse. Figure D: symbols correspond to the estimated event probability (y-axis) of sequential patients at their cumulative fludarabine AUC (x-axis). Red stars indicate the occurrence of non-relapse mortality, blue triangles indicate graft failure, green triangles indicate relapses and the black stars correspond to patients without events. The shaded area's depict the 95% confidence intervals. Displayed event probabilities correspond to a patient at the median age of 35 years, diagnosed with Leukemia/Lymphoma and no prior HCT. P-values (figure A-C) are calculated by likelihood ratio test using backwards deletion from the full regression model.

[a] Adjusted for age (polynomial spline, 3 degrees of freedom) and prior allogeneic transplants (yes/no)
[b] Adjusted for indication (Malignant/Benign)
[c] Adjusted for age (polynomial spline, 3 degrees of freedom), indication (Lymphoma/Leukemia, Benign, myelodysplastic syndrome, plasma cell disorder) and prior allogeneic transplants (yes/no)
*Significant at level $p < 0.05$
**Significant at level $p < 0.01$
***Significant at level $p < 0.001$

and sub-optimal exposures, respectively (figure 1D). This results in a target window of 15 mg*h/L to 25 mg*h/L. Outcome probabilities were adjusted for disease and other baseline characteristics (see footnotes figure 1), but the exposure target itself was found the same among different ages and indications, indicated by absence of any statistical interaction between Flu exposure and age or diagnosis in the survival models.
With the current conditioning of targeted busulfan and 160 mg/m^2 Flu-phosphate, an overall mortality of 43% was observed (Figure 2A). Death through disease recurrence occurred in 14% of patients, leaving most deaths attributable to the transplantation (28%). For malignancies, disease recurrence was higher at 23%, while the main causes of death were still disease-unrelated (29%). Figures 2B-D depict the overall mortality and causes of death, stratified for target attainment. Overall mortality was lowest in the optimally exposed group (31%), compared to the under- (43%) and over-exposed group (64%). This increase in overall mortality was mainly attributable to infections (over –and under exposure), MOF (over-exposure) and GvHD (over-exposure). Interestingly, overall GvHD (grade 2-4 or 3-4) incidence was similar among different exposure groups: acute GvHD grade 2-4 was 21% in optimally exposed patients compared to 18% below optimum, and 27% above optimum ($p = 0.68$). The trend was similar for grade 3-4 with incidences of 6%, 8% and 13% respectively ($p = 0.50$). In the under-exposed-group no relapse related death occurred, though only one patient was transplanted for a malignancy in this group (Table 2).

In the adjusted regression models, it was found that the optimal exposure group had a significantly higher EFS compared to the above-optimal-exposure group (HR 2.0, 95% confidence interval (CI) 1.1-3.5, $p = 0.01$; figure 3A) and (non-significantly) higher than the below-optimal group (HR: 1.8, 95% CI: 0.72-4.5, $p = 0.21$). The lower EFS in the above-optimal-exposure group was primarily caused by a higher incidence of NRM (HR: 3.4, 95% CI: 1.6-6.9, $p < 0.001$; figure 3C) and no difference in relapse (HR 0.85, 95% CI: 0.35-2.0, $p = 0.71$, figure 3E). In addition, the risk for graft failure and NRM were increased in the below-optimal-exposure group (HR 4.8, 95% CI 1.2-19, $p = 0.02$; figure 3D and HR 3.3, 95% CI 1.2-9.4, $p = 0.02$; figure 3C respectively). No graft-failures were observed in the above-optimal-exposure group. IR was significantly lower in patients exposed above optimum (Figure 3F, HR: 0.43, 95% CI 0.26-0.70, $p < 0.001$). Due to the high percentage of graft failures, IR could not adequately be assessed for under-exposed patients. In those patients reaching IR (n=101), early reconstitution (lower quartile) associated with increased EFS (84%

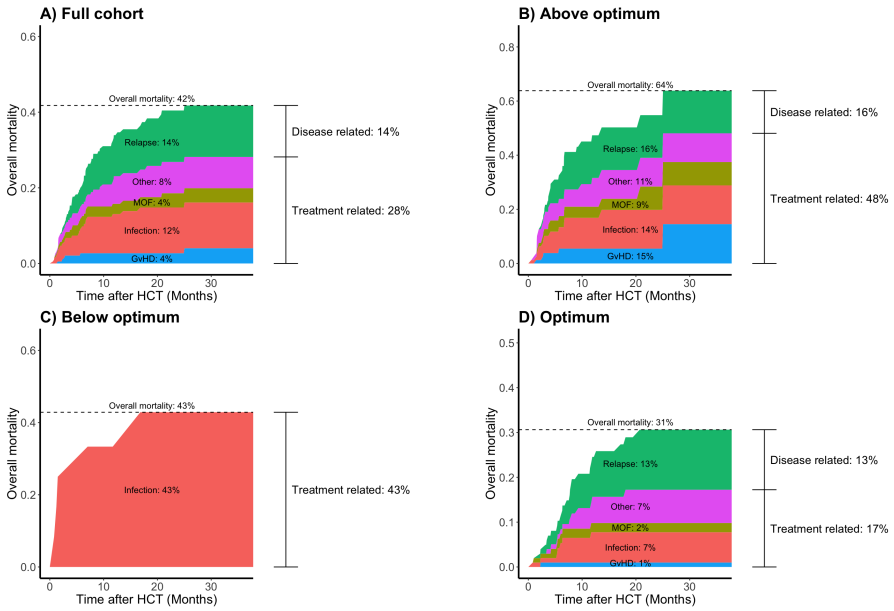

Figure 2: Causes of death according to fludarabine exposure. Stacked cumulative incidence plots of overall mortality subdivided by cause for A) the full cohort, B) patients exposed above optimum, C) patients exposed below optimum, and D) patients with an exposure within the optimal range. For each cause, the cumulative incidence was computed throughout the follow-up period. At every time-point the incidence of all causes was summed up to get the overall mortality incidence..

Table 2: Patient characteristics stratified on fludarabine exposure groups[a]

Variable	Within optimal exposure range (N=101)	Above optimal exposure range (N=79)	Below optimal exposure range (N=12)
Age at transplantation (years)***	13.7 (5.1-16, 1.9-32)	23.5 (7.1-47, 0.23-70)	56.7 (39-63, 0.29-74)
Sex			
Male	9 (75%)	58 (57%)	48 (61%)
Female	3 (25%)	43 (43%)	31 (39%)
Indication***			
Benign	11 (92%)	44 (44%)	13 (16%)
Leukemia/Lymphoma		41 (41%)	30 (38%)
MDS	1 (8.3%)	11 (11%)	18 (23%)
Plasma cell disorder		5 (5%)	18 (23%)
Cell source**			
PB: Full graft		10 (9.9%)	14 (18%)
Cord blood	11 (92%)	39 (39%)	15 (19%)
Bone marrow		12 (12%)	4 (5.1%)
PB: α-β depleted	1 (8.3%)	40 (40%)	46 (58%)
Human leukocyte antigen disparity[b]***			
Matched	79 (78%)	66 (84%)	5 (42%)
Mismatched	22 (22%)	13 (16%	7 (58%)
Conditioning			
Samples per patient	10.5 (8-12, 6-16)	10 (8-12, 3-19)	9 (8-12, 3-15)
Fludarabine $AUC_{t0-\infty}$ (mg*h/L)***	13.8 (13-15, 10-15)	21.3 (19-23, 16-25)	30.8 (27-36, 26-66)
Busulfan $AUC_{t0-\infty}$ (mg*h/L)	95.9 (90-100, 82-120)	95.3 (89-100, 59[c]-120)	97.8 (91-100, 65-120)
Serotherapy			
Patients without ATG		13 (13%)	7 (8.9%)
Patients with ATG	12 (100%)	88 (87%)	72 (91%)
$AUC_{t_x-\infty}$ of patients with ATG	3.82 (2.8-9.1, 1.5-86)	12 (1.6-35, 0-130)	18 (4.4-43, 0-270)
Creatinine clearance (ml/min/1.73m^2)***	140 (140-140, 110-140)	121 (110-140, 40-140)	100 (79-110, 25-140)
Median Follow-up (days)[d]	884 (779-NA)	688 (591-913)	369 (341-493)

[a]Categorical variables are displayed as N (%).Continuous variables are displayed as median (interquartile range, range) or [d]95% confidence interval

[b]In peripheral blood HCT, 90 were 10/10, 09/10, 16 were 8/8, 3 were 6/6). In cord blood HCT, 4 were 10/10, 4 were 9/10, 3 were 8/10, 21 were 6/6, 24 were 5/6, 9 were 4/6. In bone marrow HCT all were matched 10/10.

[c] Minima exclude Fanconi Anemia patients (n=3): for these patients indication-specific myeloablation was achieved with $AUC_{t0-\infty}$ of 31, 31, 25 mg*h/L

**Values between groups are significantly different at level $p < 0.01$

***Values between groups are significantly different at level $p < 0.001$

NA=Not available

vs 60%). Both EFS and IR were found similar in the adjusted regression models for the different cell sources, including $\alpha - \beta-$depleted grafts. In addition, subgroup analyses were performed for EFS according to optimal fludarabine exposure in the upper quartile of age and stratified for indication (malignant/benign). For patients with an age within the upper quartile, we still found superior survival in patients with optimal exposures (n=11, EFS=91%) versus those who were over-exposed (n=37, EFS=39%, $p = 0.04$). In malignancies, the optimal exposure also associated with superior EFS compared to over-exposure (55% vs 19.5%, p=0.0047). Also, for benign disorders optimal exposure associated with the best outcome (EFS: 73%) compared to below optimal (41%, $p = 0.047$) and above optimal exposure (48%, $p = 0.11$).

8.5 Discussion

To our knowledge, this is the first study describing the effect of Flu exposure on outcomes in a large HCT cohort with homogeneous pharmacological treatment. This analysis clearly shows that exposure to Flu in the conditioning prior to allogeneic HCT is a major predictor for EFS. Though differences in baseline characteristics and stem cell products were present, multivariate regression models and subset analyses showed consistency of the Flu exposure-outcome relationship. This suggests that optimizing Flu exposure, may have a relevant impact on survival in HCT. Highest EFS was found in patients with a cumulative $AUC_{t_0-\infty}$ of 20 mg*h/L (± 5) of Flu. This optimal Flu exposure was found to be consistent for all ages and indications. The risk for NRM increased at higher exposures and risk of graft failure increased at lower exposures. Interestingly, no influence on relapse incidence was noted. A limitation of this study is however that the establishment of the optimal exposure and the following survival analysis were conducted on the same dataset. Although the described cohort was a substantially large cohort (n$\sim$200), the low number of events did not allow for data splitting. The optimal exposure range has to be confirmed in an independent cohort.

Interestingly, we observed that impaired renal function in older patients and concurrent higher Flu exposures better predict worse outcomes than older age itself. Previously, others found direct relationships between older age and NRM[16,20,21]. Part of NRM in these older patients might be caused by unfavourable high Flu exposures. This also suggests that the current dosing method (per m^2) is not optimal, as it results in highly variable exposures. As weight and renal function were the only predictors for exposure found in

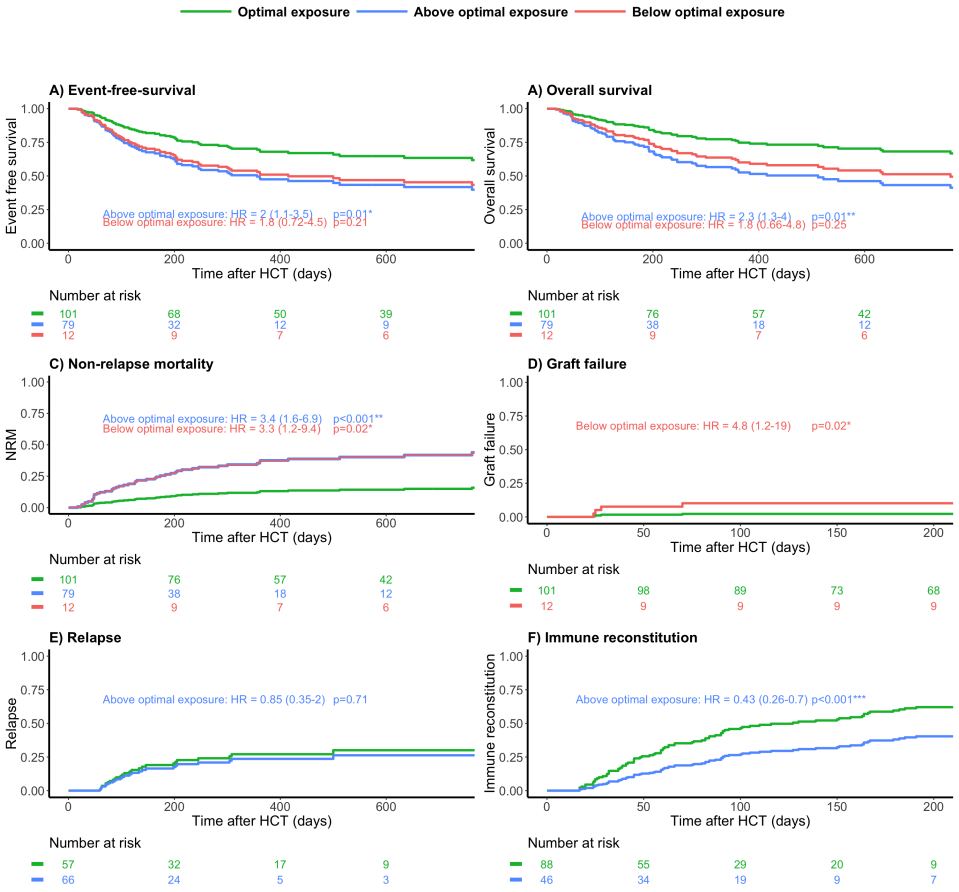

Figure 3: HCT outcome according to fludarabine exposure. Adjusted Kaplan-Meier plots and cumulative incidence plots of A) event-free survival, B) overall survival, C) Non-relapse mortality, D) Graft failure, E) Relapse, and F) immune reconstitution. Patients are stratified for within optimal exposure (Optimum ±25%), above optimal exposure and below optimal exposure. P-values are derived from the Wald's test in the full regression models.

[a]Adjusted for age (polynomial spline, 3 degrees of freedom), indication (Lymphoma/Leukemia, Benign, myelodysplastic syndrome, plasma cell disorder) and prior allogeneic transplants (yes/no)
[b]Adjusted for age (polynomial spline, 3 degrees of freedom) and prior allogeneic transplants (yes/no)
[c]Adjusted for indication (Malignant/Benign)
[d]Adjusted for age (polynomial spline, 4 degrees of freedom)
*Significant at level $p < 0.05$
**Significant at level $p < 0.01$
***Significant at level $p < 0.001$

major pharmacokinetic analyses[10,14], dosing based on these variables would likely result in better predictable, optimal exposure and subsequently result in better outcomes, including survival chances.

With respect to the identified optimal exposure range it should be noted that for the outcome graft-failure, the event frequency was low. The lower limit of the defined Flu target should, therefore, be interpreted with caution. Nevertheless, as Flu is considered an effective and important part of the conditioning regimen prior to HCT, it is very likely that a minimal target exposure should be obtained. In addition, the identified lower limit of Flu exposure is in line with results described by Ivaturi et al[14]. With regards to the upper limit, over-exposure was found to be strongly related to impaired IR and NRM, which is in line with studies described by others[5,13].

The strong effect of high Flu exposure before HCT on IR after HCT is however intriguing: 'How can exposure before HCT influence IR after HCT?'. We postulate two hypotheses for this finding. Firstly, due to over-exposure of Flu an unfavourable host environment is created, impacting CD4+ T-cell homeostatic expansion through a negative influence on thymus size. That chemotherapy can do this was described previously[22,23], although the role of the thymus at older age is not completely understood. Alternatively, it could be that cytotoxic F-ara-A still resides in the bone marrow after transplantation, when plasma levels have already declined, given the excellent penetration of Flu in the bone marrow[24,25]. Delayed IR has also been associated with increased relapse risk[3]. In this study, however, no relationship between high Flu exposure and relapse was found, despite the observed impaired IR. An explanation for this may be that in in vitro studies, synergism between Flu and busulfan was observed[26], resulting in increased apoptosis of residual malignant cells. The higher exposure, in over-exposed group may have counterbalanced the negative effect of delayed IR on the relapse probability.

In addition to being the most frequently used agent in the conditioning prior to HCT, Flu is also used in the treatment of chronic lymphocytic leukemia (CLL)[27] as well as in conditioning regimens before "chimeric antigen receptor T-cell" (CART) therapy and other cellular T-cell therapies[28–31]. In CLL, profound Flu-therapy-related lymphopenia, which might indicate high Flu-exposure, has been associated with better survival[32]. In conditioning regimens prior to cellular therapy (such as CART), Flu has been found necessary for cell persistence[33], while also adverse events related to Flu conditioning have been observed[34]. This suggests that a Flu exposure-response relationship in these settings exists as well.

In conclusion, optimal cumulative exposure to Flu during the conditioning

phase predicts superior survival following HCT. With renal function and weight being the predictors for Flu exposure, dosing based on these parameters would make more sense, instead of dosing based on body-surface area. Such an algorithm could be readily derived from the predicted clearance using weight, eGFR and the previously published model[10]. Alternatively, implementing 'therapeutic drug monitoring' would give the most precise exposures. As such, individualized dosing could aid in improving predictability of, and survival after, Flu-containing conditioning regimens.

References

1. Hahn, T., McCarthy P. L., J., Hassebroek, A., Bredeson, C., Gajewski, J. L., Hale, G. A., *et al.* Significant improvement in survival after allogeneic hematopoietic cell transplantation during a period of significantly increased use, older recipient age, and use of unrelated donors. *J Clin Oncol* **31,** 2437–49 (2013).
2. Admiraal, R., de Koning, C., Lindemans, C. A., Bierings, M. B., Wensing, A. M., Versluys, A. B., *et al.* Viral Reactivations and Associated Outcomes in Context of Immune Reconstitution after Pediatric Hematopoietic Cell Transplantation. *J Allergy Clin Immunol* (2017).
3. Admiraal, R., van Kesteren, C., Jol-van der Zijde, C. M., Lankester, A. C., Bierings, M. B., Egberts, T. C., *et al.* Association between anti-thymocyte globulin exposure and CD4+ immune reconstitution in paediatric haemopoietic cell transplantation: a multicentre, retrospective pharmacodynamic cohort analysis. *Lancet Haematol* **2,** e194–203 (2015).
4. Bartelink, I. H., Lalmohamed, A., van Reij, E. M., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *Lancet Haematol* **3,** e526–e536 (2016).
5. Sanghavi, K., Wiseman, A., Kirstein, M. N., Cao, Q., Brundage, R., Jensen, K., *et al.* Personalized fludarabine dosing to reduce nonrelapse mortality in hematopoietic stem-cell transplant recipients receiving reduced intensity conditioning. *Transl Res* **175,** 103–115 e4 (2016).
6. Andersson, B. S., Thall, P. F., Valdez, B. C., Milton, D. R., Al-Atrash, G., Chen, J., *et al.* Fludarabine with pharmacokinetically guided IV busulfan is superior to fixed-dose delivery in pretransplant conditioning of AML/MDS patients. *Bone Marrow Transplant* **52,** 580–587 (2017).
7. Admiraal, R. & Boelens, J. J. Individualized conditioning regimes in cord blood transplantation: Towards improved and predictable safety and efficacy. *Expert Opinion on Biological Therapy* **16,** 801–813 (2016).
8. Admiraal, R., van Kesteren, C., Jol-van der Zijde, C. M., van Tol, M. J., Bartelink, I. H., Bredius, R. G., *et al.* Population pharmacokinetic modeling of Thymoglobulin((R)) in children receiving allogeneic-hematopoietic cell transplantation: towards improved survival through individualized dosing. *Clin Pharmacokinet* **54,** 435–46 (2015).

9. Gyurkocza, B. & Sandmaier, B. M. Conditioning regimens for hematopoietic cell transplantation: one size does not fit all. *Blood* **124,** 344–53 (2014).

10. Langenhorst, J. B., Dorlo, T. P. C., van Maarseveen, E. M., Nierkens, S., Kuball, J., Boelens, J. J., *et al.* Population Pharmacokinetics of Fludarabine in Children and Adults during Conditioning Prior to Allogeneic Hematopoietic Cell Transplantation. *Clinical pharmacokinetics.* Epub ahead of print (Oct. 2018).

11. Lindemans, C. A., Te Boome, L. C., Admiraal, R., Jol-van der Zijde, E. C., Wensing, A. M., Versluijs, A. B., *et al.* Sufficient Immunosuppression with Thymoglobulin Is Essential for a Successful Haplo-Myeloid Bridge in Haploidentical-Cord Blood Transplantation. *Biol Blood Marrow Transplant* **21,** 1839–45 (2015).

12. Bartelink, I. H., Belitser, S. V., Knibbe, C. A., Danhof, M., de Pagter, A. J., Egberts, T. C., *et al.* Immune reconstitution kinetics as an early predictor for mortality using various hematopoietic stem cell sources in children. *Biol Blood Marrow Transplant* **19,** 305–13 (2013).

13. Long-Boyle, J. R., Green, K. G., Brunstein, C. G., Cao, Q., Rogosheske, J., Weisdorf, D. J., *et al.* High fludarabine exposure and relationship with treatment-related mortality after nonmyeloablative hematopoietic cell transplantation. *Bone Marrow Transplant* **46,** 20–6 (2011).

14. Ivaturi, V., Dvorak, C. C., Chan, D., Liu, T., Cowan, M. J., Wahlstrom, J., *et al.* Pharmacokinetics and Model-Based Dosing to Optimize Fludarabine Therapy in Pediatric Hematopoietic Cell Transplant Recipients. *Biol Blood Marrow Transplant* **23,** 1701–1713 (2017).

15. Admiraal, R., Lindemans, C. A., van Kesteren, C., Bierings, M. B., Versluijs, A. B., Nierkens, S., *et al.* Excellent T-cell reconstitution and survival depend on low ATG exposure after pediatric cord blood transplantation. *Blood* **128,** 2734–2741 (2016).

16. Admiraal, R., Nierkens, S., de Witte, M. A., Petersen, E. J., Fleurke, G. J., Verrest, L., *et al.* Association between anti-thymocyte globulin exposure and survival outcomes in adult unrelated haemopoietic cell transplantation: a multicentre, retrospective, pharmacodynamic cohort analysis. *Lancet Haematol* **4,** e183–e191 (2017).

17. De Koning, C., Gabelich, J. A., Langenhorst, J., Admiraal, R., Kuball, J., Boelens, J. J., *et al.* Filgrastim enhances T-cell clearance by antithymocyte globulin exposure after unrelated cord blood transplantation. *Blood Adv* **2,** 565–574 (2018).

18. Altman, D. G., De Stavola, B. L., Love, S. B. & Stepniewska, K. A. Review of survival analyses published in cancer journals. *Br J Cancer* **72,** 511–8 (1995).

19. Sun, G. W., Shook, T. L. & Kay, G. L. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* **49,** 907–16 (1996).

20. Gilleece, M. H., Labopin, M., Yakoub-Agha, I., Volin, L., Socie, G., Ljungman, P., *et al.* Measurable residual disease, conditioning regimen intensity, and age predict outcome of allogeneic hematopoietic cell transplantation for acute myeloid leukemia in first remission: A registry analysis of 2292 patients by the Acute Leukemia Working Party European Society of Blood and Marrow Transplantation. *Am J Hematol* **93,** 1142–1152 (2018).

21. Wais, V., Bunjes, D., Kuchenbauer, F. & Sorror, M. L. Comorbidities, age, and other patient-related predictors of allogeneic hematopoietic cell transplantation outcomes. *Expert Rev Hematol* **11,** 805–816 (2018).

22. Mackall, C. L., Fleisher, T. A., Brown, M. R., Andrich, M. P., Chen, C. C., Feuerstein, I. M., *et al.* Age, thymopoiesis, and CD4+ T-lymphocyte regeneration after intensive chemotherapy. *N Engl J Med* **332,** 143–9 (1995).

23. Sfikakis, P. P., Gourgoulis, G. M., Moulopoulos, L. A., Kouvatseas, G., Theofilopoulos, A. N. & Dimopoulos, M. A. Age-related thymic activity in adults following chemotherapy-induced lymphopenia. *Eur J Clin Invest* **35,** 380–7 (2005).

24. Dhilly, M., Guillouet, S., Patin, D., Fillesoye, F., Abbas, A., Gourand, F., *et al.* 2-[18F]fludarabine, a novel positron emission tomography (PET) tracer for imaging lymphoma: a micro-PET study in murine models. *Mol Imaging Biol* **16,** 118–26 (2014).

25. Lindemalm, S., Liliemark, J., Larsson, B. S. & Albertioni, F. Distribution of 2-chloro-2'-deoxyadenosine, 2-chloro-2'-arabino-fluoro-2'-deoxyadenosine, fludarabine and cytarabine in mice: a whole-body autoradiography study. *Med Oncol* **16,** 239–44 (1999).

26. Valdez, B. C., Li, Y., Murray, D., Champlin, R. E. & Andersson, B. S. The synergistic cytotoxicity of clofarabine, fludarabine and busulfan in AML cells involves ATM pathway activation and chromatin remodeling. *Biochem Pharmacol* **81,** 222–32 (2011).

27. Skarbnik, A. P. & Faderl, S. The role of combined fludarabine, cyclophosphamide and rituximab chemoimmunotherapy in chronic lymphocytic leukemia: current evidence and controversies. *Ther Adv Hematol* **8,** 99–105 (2017).

28. Heczey, A., Louis, C. U., Savoldo, B., Dakhova, O., Durett, A., Grilley, B., *et al.* CAR T Cells Administered in Combination with Lymphodepletion and PD-1 Inhibition to Patients with Neuroblastoma. *Mol Ther* **25,** 2214–2224 (2017).

29. Maude, S. L., Frey, N., Shaw, P. A., Aplenc, R., Barrett, D. M., Bunin, N. J., *et al.* Chimeric antigen receptor T cells for sustained remissions in leukemia. *N Engl J Med* **371,** 1507–17 (2014).

30. Neelapu, S. S., Locke, F. L., Bartlett, N. L., Lekakis, L. J., Miklos, D. B., Jacobson, C. A., *et al.* Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. *N Engl J Med* **377,** 2531–2544 (2017).

31. Turtle, C. J., Hay, K. A., Hanafi, L. A., Li, D., Cherian, S., Chen, X., *et al.* Durable Molecular Remissions in Chronic Lymphocytic Leukemia Treated With CD19-Specific Chimeric Antigen Receptor-Modified T Cells After Failure of Ibrutinib. *J Clin Oncol* **35,** 3010–3020 (2017).

32. McCune, J. S., Vicini, P., Salinger, D. H., O'Donnell, P. V., Sandmaier, B. M., Anasetti, C., *et al.* Population pharmacokinetic/dynamic model of lymphosuppression after fludarabine administration. *Cancer Chemother Pharmacol* **75,** 67–75 (2015).

33. Wallen, H., Thompson, J. A., Reilly, J. Z., Rodmyre, R. M., Cao, J. & Yee, C. Fludarabine modulates immune response and extends in vivo survival of adoptively transferred CD8 T cells in patients with metastatic melanoma. *PLoS One* **4,** e4749 (2009).

34. Lowe, K. L., Mackall, C. L., Norry, E., Amado, R., Jakobsen, B. K. & Binder, G. Fludarabine and neurotoxicity in engineered T-cell therapy. *Gene Ther* **25,** 176–191 (2018).

9 A novel simulation framework for rational design of trials evaluating optimal fludarabine dosing in allogeneic hematopoietic cell transplantation

Jurgen B Langenhorst, Thomas PC Dorlo, Charlotte van Kesteren, Erik M van Maarseveen, Stefan Nierkens, Moniek A de Witte, Jaap Jan Boelens, Alwin DR Huitema

Manuscript in preparation

9.1 Abstract

Fludarabine exposure is strongly related to treatment outcome in allogeneic hematopoietic cell transplantation (HCT), offering potential to individualize dosing for treatment optimization. We aimed to find optimal trial design to evaluate potential superiority of novel dosing strategies. For this, previously developed cause-specific hazard (CSH)-models for events: non-relapse mortality (NRM), relapse, and graft failure were expanded by Markovian elements to model the transition from relapse/graft failure to death. Randomized controlled trials were simulated, comparing conventional dosing (160 mg/m^2) to either model-based or therapeutic drug monitoring (TDM)-guided-dosing. The number of patients required to achieve 80% power was calculated for NRM and overall survival (OS). The optimal trial had TDM-guided-dosing as intervention and NRM as primary outcome, with 82% power at with n=70 per dosing arm, compared to n=160 with TDM-guided-dosing + OS and n=120 for model-based-dosing + NRM. The expected survival gain of the optimal trial was 13%. We conclude that a superiority trial is feasible.

9.2 Introduction

Allogeneic hematopoietic cell transplantation (HCT) is a potentially curative treatment for a variety of malignant and benign hematological disorders. Unfortunately, non-relapse mortality (NRM) (10-40%) and disease relapse (20-50%) remain major causes of therapy failure[1], thus further treatment optimization is potentially life-saving.

The conditioning regimen prior to HCT consists of a combination of cytotoxic agents (chemo- and serotherapy) administered to eradicate recipient's bone marrow and immune system. Reducing the toxicity while maintaining the efficacy of such regimens is one of the key strategies to reduce NRM[2]. Fludarabine combined with busulfan and rabbit anti-thymocyte globulin (rATG) is a frequently used conditioning regimen for HCT. Variability in exposure has been shown to predict variable outcome for all these agents[3–5]. Given that exposure-targeted busulfan through therapeutic drug monitoring (TDM) is widely applied and has been proven superior over fixed dosing in a randomized controlled trial (RCT)[6], the focus for dose optimization has now shifted to fludarabine. We previously described highly variable fludarabine plasma exposures following conventional body-surface-area (BSA)-based dosing[7]. Ivaturi et al. showed an association between low fludarabine exposure and graft failure with subsequent lower survival[5]. Furthermore, other studies demonstrated that over-exposure[8,9] (or lower predicted fludarabine clearance[9]) was associated with NRM. In follow-up to these findings an optimal fludarabine exposure was identified at an area under the concentration time curve of the full conditioning regimen (area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($AUC_{t_0-\infty}$)) of 20 mg*h/L[10]. This exposure appeared to be associated with the minimal probability of having any of the negative events generally used to determine HCT outcome (NRM, relapse, graft failure). Individualized dosing, guided by TDM or based on the pharmacokinetic (PK) model of fludarabine[7], is likely to improve PK target attainment compared to conventional dosing, and possibly increase survival.

To implement and evaluate the superiority and feasibility of the suggested individualized fludarabine dosing algorithms in clinical practice, a RCT would be required to confirm the suggested advantage in terms of reduced NRM or increased overall survival (OS) of individualized over conventional dosing. However, the design of such a trial is complicated by various choices and practical circumstances/limitations, such as choice of endpoint, limited number of patients available, uncertainty regarding which dosing strategy is optimal and the corresponding effect sizes.

To aid rational decision making regarding optimal trial design, we constructed a framework connecting dosing to 1) plasma exposure ($AUC_{t_0-\infty}$) 2) events (NRM, relapse, graft failure) and 3) survival by using the PK model, the previously developed cause-specific hazard (CSH)-models, and newly developed Markovian transition elements. This framework was subsequently used to perform a clinical trial simulation (CTS) of RCT's comparing conventional to alternative dosing strategies. CTS were used to assess the feasibility of an RCT evaluating the suggested individualized regimens, identify the most optimal endpoint for such a trial, and predict the possible gain of individualized fludarabine dosing.

9.3 Methods

General framework

In figure 1, the full simulation framework is depicted. The framework consists of a PK part before HCT and a time-to-event part following HCT. For the PK components, the previously published PK model[7] was used. The time-to-event part can be further subdivided in the time until first event and, if applicable, a subsequent Markovian post-event survival transition.
For the first part of the time-to-event framework, the previously used dataset and CSH-models[10] were developed and refined further (methods section 9.3) to include only predictive covariates. Previously, covariate inclusion was on an *a priori* basis. Applying a different approach this time was necessary due to the switch from an etiological (influence of fludarabine exposure) to a prognostic setup (predicting the outcome resulting from dosing alteration and other predictive covariates).
For the second part of the time-to-event framework, new models were constructed (methods section 9.3).

Data

For estimation purposes, the same dataset as in the previous publication was used[10].
The trial-populations used for simulation were sampled from a population based on an in-house database of HCT recipients (collected during 2005-2016: n=194). Selection was based on indication (myelodysplastic syndrome (MDS), leukemia, lymphoma), age (>20 years at HCT), and availability of creatinine values within 1 month of transplantation. These criteria were chosen to exclude plasma cell disorders (debatable eligibility for HCT due to poor outcome[11]), select the population with a high risk of over-exposure

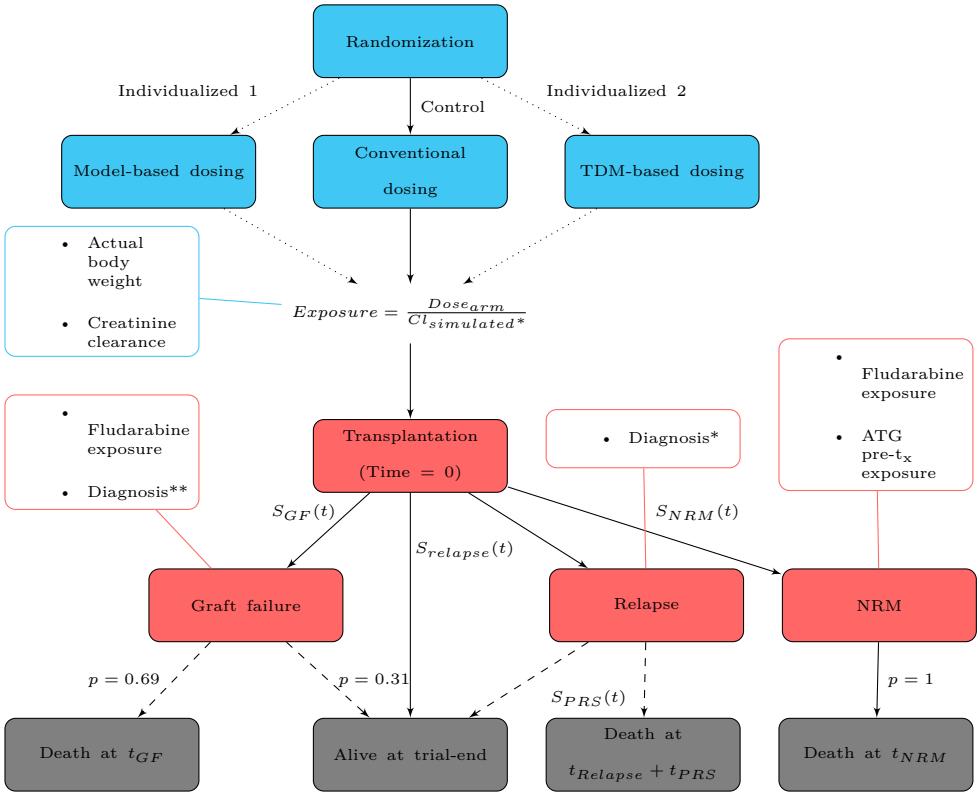

Figure 1: Simulation framework. Pharmacokinetic simulations are depicted by blue boxes, time-to-event simulations by the red boxes, and survival status by the gray boxes. Dotted lines are optional directions (i.e. individualized dosing is either TDM or model-based, but the control is always conventional dosing). Dashed lines depict the Markovian elements: transition from event to survival/death. Colored lines connect the covariates boxes to the corresponding model.

GF= Graft failure, rATG = rabbit anti-thymocyte globulin, t_x = Time of transpantation, Cl = clearance, PRS = post-relapse survival
*Diagnosis here is subdivided in: Leukemia/Lymphoma, Myelodysplastic Syndrome, plasma cell disorders, and benign disorders.
**Diagnosis here is subdivided in: Benign and Malignant disorders.

with conventional dosing (adults)[7,10], or omit imputation of creatinine values, respectively.

For the largest proportion of the in-house dataset covariates were well documented. To allow a complete set of input covariates necessary to predict PK and event probabilities, missing covariate values for remaining patients were imputed as follows. Body weights were sampled per sex from the known body weight distributions in each quintile (20% quantile) of age. To impute heights that were considered plausible for the corresponding weight, body-

mass index (BMI) was sampled from a distribution per 5% age quantile. The heights were then calculated from the body weight and imputed BMI. RATG exposures were not documented for all patients in the simulation cohort and were, therefore, simulated as described in section 9.3.

Table 1: Patient characteristics

	Population: estimation	Population: simulation
Age at HCT (years)	36 (0.23-74)	46 (20-74)
Weight at HCT (kg)	65 (4.3-130)	74 (36-130)
Dose (mg/m^2)	160	Conventional: 160 TDM: 119 (73.2-183)* Model-based: 123 (82.6-142)*
Fludarabine $AUC_{T0-\infty}$ (mg*h/L)	24.7 (13.9-66)	Conventional: 27.4 (17.2-45.6)* TDM: 20.2 (17.3-24.1)* Model-based*: 20.3 (13.4-30)*
Creatinine clearance (ml/min/1.73m^2)	110 (25-140)	110 (42-140)
ATG pre-HCT exposure (AU*day/ml)	54 (4.4-210)	77 (28-226)*
Diagnosis (N, %)		
Leukemia/Lymphoma	71 (37%)	76%
MDS	30 (16%)	24%
Plasma cell disorder	23 (12%)	Excluded
Benign	68 (35%)	Excluded

Displayed values are: median (range) unless specified otherwise
*2.5th and 97.5th percentiles
MDS = myelodysplastic syndrome
ATG = anti-thymocyte globulin

Simulating pharmacokinetics

The input fludarabine PK database was constructed, assuming conventional dosing of 160 mg/m^2 divided over 4 days for each patient. The previously published PK model[7] was used to simulate 25 sets of individual clearance values per input set of covariates, to take into account inter-patient variability. Based on the simulated typical individual clearance, clearance values per day were simulated, taking into account inter-occasion variability (IOV). Subsequently, five fludarabine plasma concentrations (also taking into account residual variability) were simulated on the first day of conditioning at times centered around 1, 5, 6, 7, and 8 hours after the start of infusion, where

the distributions around these sample times were taken from the previous study[7]. These samples were simulated in order to perform TDM. TDM samples were randomly excluded according to the previously observed distribution of missing samples (~8% randomly missing data).
For rATG, dosing and simulations were performed according to the respective regimen ($400 + 350 \cdot ALC$ starting at day -9) and PK model, described by Admiraal et al[3]. Where ALC is the absolute number of lymphocytes a patient has preceding rATG administration in 10^9 cells per liter of blood.
Henceforth, rATG exposure was calculated by integration of the plasma-concentration-time curve as AUC from the first dose until time of transplantation ($\mathrm{AUC}_{t_0 - t_x}$) and $\mathrm{AUC}_{t_0 - t_x}$.

Time-to-event model (re-)estimation

Baseline hazards for the CSH-models were derived from the previous PK-pharmacodynamic (PD) analysis[10]. To include predictive covariates to be used in the simulations, a selection based on forward addition ($p < 0.1$) and backwards deletion ($p > 0.05$) was performed. Evaluated covariates were based on reports from literature, biological and/or physiological plausibility: age at HCT[3,12], rATG exposure pre- and post-transplantation[3,13], cumulative busulfan exposure[4], graft source[14,15], cytomegalovirus serostatus (patient and donor)[16], diagnosis[17], prior HCT[18–20], and human leukocyte antigen disparity[21].
Continuous covariates were tested as covariate on the baseline hazard using a linear, polynomial (1st, 2nd, 3rd degree) and a polynomial spline (3, 4, 5 degrees of freedom). The relationship resulting in the lowest Akaike information criterion in the univariate analysis was selected as the most relevant form and was tested throughout the procedure.
For categorical covariates, the category with the highest number of patients was chosen as a reference. Other categories were chosen to have its own proportional relationship with the baseline hazard relative to the reference.
The probability of transition from relapse to death was modelled using a parametric survival model. Patients included were those who experienced a relapse and were followed-up from the time of relapse until death or last follow-up, whichever occurred first. Due to the small number of graft failures, the probability of death could not be estimated with reasonable precision and was taken from literature: 0.69[22].

Framework evaluation

The different elements of the simulation platform were evaluated using visual predictive checks. For this, events, survival and corresponding event-times were simulated 1000 times for each subject. The cumulative incidence of events and overall survival were computed for each simulation. The mean and 95% prediction interval (PI) of 1000 simulations were compared to the observed cumulative incidence and Kaplan-Meier curves for the events and OS, respectively. Results were stratified for indication (leukemia, lymphoma, MDS) and/or fludarabine exposure (tertiles per indication), depending on event-type. The model should adequately reflect the observed survival differences as a result from exposure variation in all indications, to justify extrapolation of the exposure response relationship to the proposed clinical trial setting.

Dosing regimens

The fludarabine $\mathrm{AUC}_{t_0-\infty}$ proved to be predictive for outcome[5,8–10] and was calculated for each dosing regimen (conventional dosing, model-based dosing, and TDM-based dosing) using equation 1.

$$AUC_{T0-\infty} = \sum_{i=1}^{4} \frac{Dose_{day_i}}{Clearance_{day_i}} \tag{1}$$

All doses were administered over 4 days at day -5 to -2 relative to HCT. Conventional dosing was defined as a cumulative dose of 160 mg/m^2, and model-based dosing was calculated using equation 2, with the total calculated dose for both dosing regimens being equally divided over the 4-day period.

$$Cumulative\,dose = 20mg \cdot h \cdot L^{-1} \cdot (0.782 \cdot eGFR \cdot + 3.24L/h \cdot \frac{BW}{70}^{0.75}) \tag{2}$$

where the cumulative dose is in mg of the fludarabine base (administered as phosphate), eGFR is estimated glomerular filtration rate in l/h/70kg and BW is in kg.

TDM-based dosing was performed by first predicting the clearance from the simulated fludarabine concentration-time data (section 9.3) on the first of 4 days of conditioning. This was done with Bayesian forecasting, where the final estimates from the PK model (fixed and random effects) were used as

priors and PK output was generated using the posthoc step in NONMEM. Based on this the exposure after unchanged (conventional) dosing could be estimated. Henceforth, the dose adjustments for the subsequent days necessary to achieve the desired target exposure (20 mg*h/L) were calculated, taking into account the already administered exposure of the first conventional dose.

Simulating events and survival

Using baseline characteristics and the simulated fludarabine $AUC_{t_0-\infty}$ per dosing regimen, daily event probabilities were simulated with the CSH-models for each subject. A trial was simulated, by randomly assigning patients, stratified on diagnosis, to receive either conventional (160 mg/m^2) or individualized dosing (TDM or model-based). Events and OS were simulated using the survival probabilities corresponding to the assigned dosing regimen. Trials were simulated until one year after HCT.
To calculate the power of a trial design, the proposed trial was simulated 1000 times and for each trial the p-value was calculated using either Gray's test (separate events) or a 2-sided log-rank test (OS). The power was defined as the fraction of trials reaching $p < 0.05$. The designs were evaluated by performing randomization with a varying number {50, 60, ...,220} of subjects per arm (n). The trial format with the lowest n to achieve 80% power was deemed optimal.

Trial uncertainty

A sensitivity analysis was performed in the optimal trial to test uncertainties in the assumptions. Firstly, an alteration in drug effect favoring high exposure for all events was assumed. The advantage of individualized dosing was, thereby, reduced as these doses and associated exposures were generally lower (table 1). For this, the probability of graft failure at exposures below the optimum was adjusted by a relative risk of 1.1 (10% higher). In a separate simulation, the NRM probability at exposures above the optimum was adjusted by a relative risk of 0.9 (10% lower). A third analysis was conducted where a previously unobserved relationship with relapse was introduced. Here, the relapse probability increased exponentially below the optimum up to a relative risk of 2.0 for each 10 mg*h/L decrease of fludarabine exposure.
Furthermore, the possibility that TDM is unsuccessful in 10% of the subjects, defined as a loss of all samples for these patients, was tested. For this,

randomly selected subjects received model-based dosing instead of TDM-based dosing.
For all situations, power was recalculated by simulating the trial 1000 times, implementing aforementioned assumptions.

9.4 Results

Population characteristics

Characteristics of the populations used for estimation[10] and trial sampling are listed in table 1. The characteristics used as input for the simulations were well within the range observed in the original population, where the models were based on.

Time-to-event model (re-)estimation

Baseline hazard distributions for NRM, relapse and graft failure were kept as previously described[10]. Predictive covariates for NRM were: rATG pre-transplant exposure (polynomial spline, $p = 0.04$), fludarabine $AUC_{t_0-\infty}$ (polynomial spline, $p < 0.001$); for relapse: diagnosis (benign, leukemia/ lymphoma, plasma cell disorder, MDS, $p < 0.001$); for graft failure: fludarabine $AUC_{t_0-\infty}$ (linear, $p = 0.04$) and diagnosis (Benign, Malignant, $p = 0.03$). For post-relapse survival, a Gompertz distribution was found optimal.
The visual predictive check (figure 2) shows adequate description of overall survival with the current platform. The same holds true for cumulative incidence of separate events (data not shown).

Exposure simulation

In figure 3 the expected exposures for the different dosing regimens are depicted. In general, the individualized dosing regimens resulted in lower exposures with a less wide distribution. This is emphasized by the high proportion (59%) of patients that exhibited an exposure >25% higher than the defined optimum after conventional dosing. In contrast, model-based dosing led to more patients (11% opposed to 2%) with exposures that were >25% lower than the optimum. TDM performed best, with 97% of patients within ±25% of the optimum.

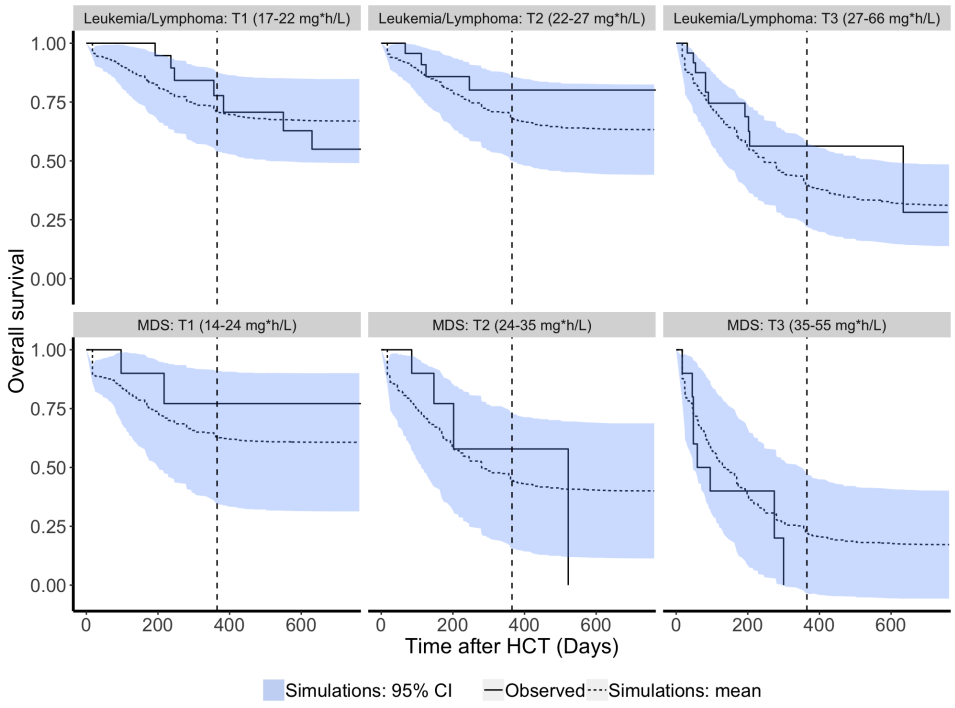

Figure 2: Visual predictive checks. Overall survival stratified for diagnosis and fludarabine exposure tertiles within each diagnosis. Solid lines indicate the observed Kaplan-Meier estimates and the dotted line shows the mean of simulations. The blue area represents the 95% confidence interval of simulations.

Power calculations

In all simulated scenario's the power to detect a significant difference in relapse or graft failure remained low (figure 4: <15%). When conventional dosing was compared to model-based dosing, a power of 80% is reached with a trial size of 120 patients per arm for NRM, while this power threshold was not reached for OS up until a trial size of 220 per arm. In scenario's where TDM-based dosing was the intervention, power increased more rapidly with 70 and 160 patients per arm necessary to reach 80% power for NRM and OS respectively. The scenario where TDM-based dosing was compared to conventional dosing with NRM as a primary end-point was deemed the optimal trial format.

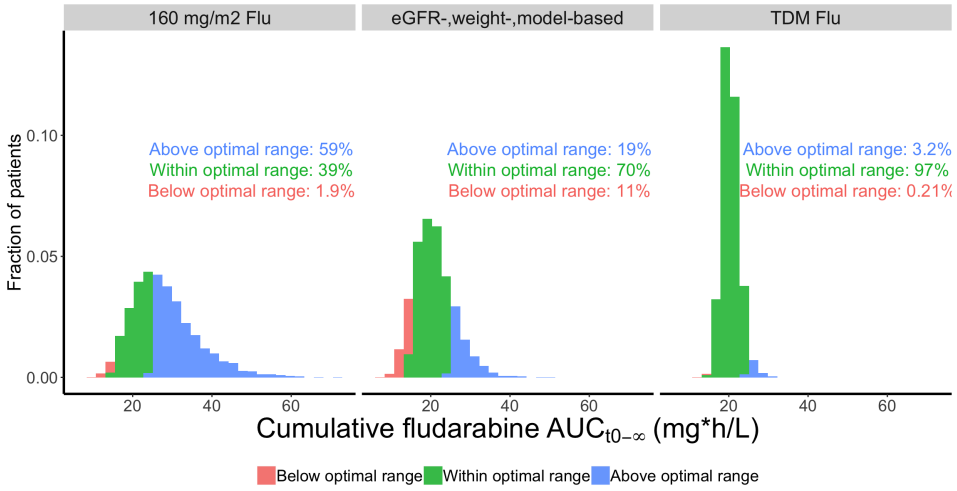

Figure 3: Fludarabine exposures for different dosing regimens. Histograms of the simulated fludarabine exposure following the different dosing regimens. The optimal range is defined as the optimum ± 25% (15-25 mg*h/L)

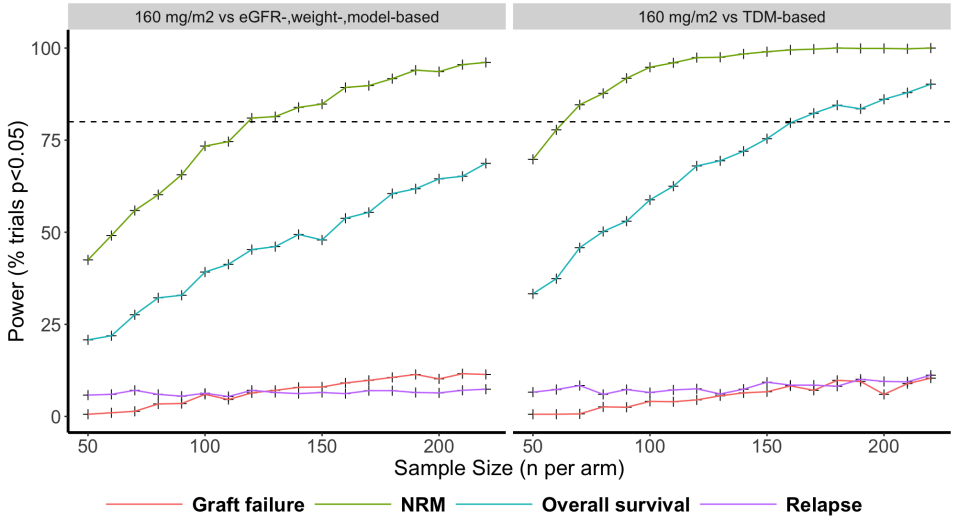

Figure 4: Power calculations for simulated trials. Each point depicts the percentage of 1000 trials with $p < 0.05$ at the corresponding number of patients per arm. The horizontal dashed line is the threshold of 80% power.

Optimal trial results

The difference in median overall mortality was 13% in favor of TDM-based dosing (figure 5). This was due to a decrease in NRM of 16%, comparable

graft failure and an increase in relapse of 3%. The increased relapse incidence is caused by more patients being at risk for this later occurring event, due to a decreased early NRM.

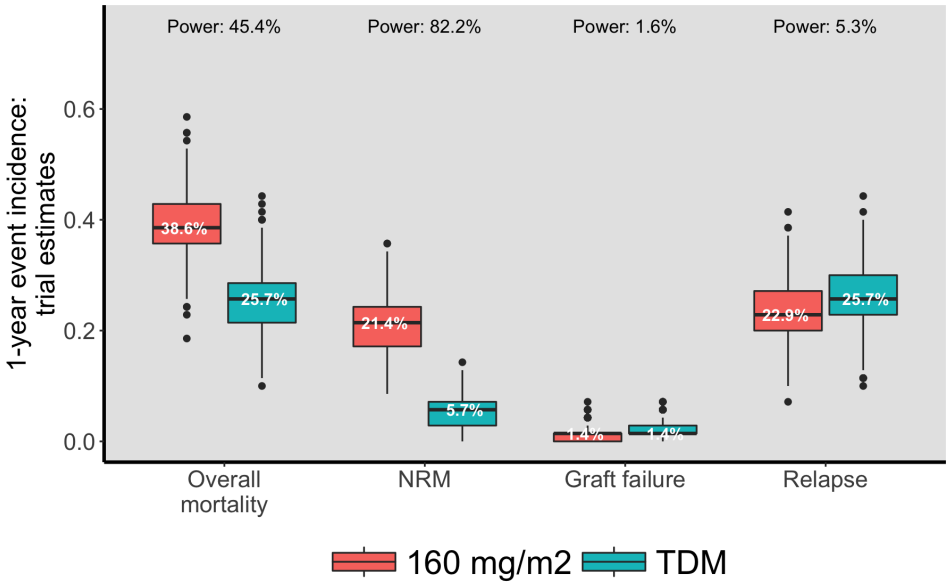

Figure 5: Simulated outcomes with optimal trial design. Boxplots of simulated trial results stratified by TDM versus conventional dosing. Values of overall mortality represent 1-Kaplan-Meier estimate (overall mortality) or the cumulative incidence (Graft failure, NRM, Relapse). The horizontal center line and percentages correspond to the median estimates. Boxes go from the 25^{th} to 75^{th} percentile and whiskers are twice the interquartile range.

Trial uncertainty

The sensitivity analyses showed that increasing the event probability for relapse and graft failure at low exposures ($AUC_{t0-\infty} < 20mg * h/L$), as described in section 9.3, had a marginal effect on trial power (table 2). However, reducing NRM probability at $AUC_{t0-\infty} > 20mg * h/L$ by 10% did reduce power below the 80% threshold. Also, unsuccessful TDM had an effect on power, but to a lesser extent (adjusted power: 79.8%).

9.5 Discussion

In the CTS, we showed that individualized fludarabine dosing through TDM might improve OS after HCT in adults with malignancies by approximately

Table 2: Sensitivity analysis

Trial assumptions (n=70, NRM primary end-point)	NRM (power,%)
Original assumptions	82.2
P_{NRM} reduced from $AUC_{T0-\infty} > 20mg*h/L$: RR=0.9	73.8
P_{GF} increased from $AUC_{T0-\infty} < 20mg*h/L$: RR=1.1	82.0
$P_{Relapse}$ increased from $AUC_{T0-\infty} < 20mg*h/L$: RR=2.0/10mg*h/L	82.1
TDM fails in 10% of cases	79.8

P = Event probability
NRM = Non-relapse mortality
GF = Graft failure
TDM = Therapeutic drug monitoring
RR = Relative risk

13%. This is attributable to reduced NRM, without increasing risk of graft failure. The described simulations were based on the exposure-response relationship for fludarabine[10], which was established on a dataset of 192 HCT recipients of all ages. In order to confirm superiority of TDM to conventional 160 mg/m^2 dosing in a RCT with NRM as endpoint, 70 patients per arm would be required for a power of 80% under ideal circumstances.
The predicted 16% absolute reduction of NRM resulting in 13% superior OS by using TDM is substantial. This also explains the relatively low number of patients required in a confirmatory setting. To put aforementioned numbers in perspective, a trial similar to the setting simulated here was performed for busulfan in HCT conditioning[6]. In that trial, TDM- to BSA-based dosing has been compared and an approximate survival gain of 20% has been found. Statistical significance was achieved ($p = 0.04$) with 110 patients per arm, which is more than the 70 found in this study. However, their primary end-point was OS, which was also associated with lower power in our simulations. Even though, for busulfan this end-point was necessary as the highly prevalent events of relapse and NRM were expected to be reduced through omission of under- and over-exposure respectively. In our study, we simulated NRM to be the end-point with most discriminative power for fludarabine, because graft failure, which mainly drives the lower limit, is not of much concern in malignancies[23] and an influence of fludarabine exposure on relapse was not identified. This lack of relationship between fludarabine

exposure and relapse was explained by the fact that on one hand high fludarabine exposures might result in more cytotoxic activity but on the other hand might reduce the graft-versus-disease effect essential for tumor control after HCT[10]. In addition, the proposed CTS-framework could also be used in other HCT settings, where graft failure plays a more prominent role, such as transplantation for benign disorders and/or a comparison with the also used lower BSA-based fludarabine dosing of 120-150 mg/m^2[24–26].

The results of a CTS as performed here are fully dependent on the underlying assumptions. Our CTS framework consisted of a PK model and survival models. Here, assumptions in the PK model were of less concern than in the survival models, given the difference in parameter uncertainty between these two models[7,10] and the well quantified PK inter-patient variability[7]. The predicted fludarabine exposures are, therefore, expected to be representative for the real clinical situation. However, the survival models have more uncertainty compared to the PK model, which makes extrapolation beyond characteristics of the dataset used for survival modeling more difficult. Importantly, in the dataset used for model development, lower exposures were underrepresented in adult malignancies. Therefore, sensitivity analyses were conducted to evaluate the impact of uncertainty in the survival models. To test this, a previously unobserved association between low exposure and relapse was introduced. Also, the probability of graft failure at low exposures was increased. These unfavorable effect of low exposures (more relapse and graft failure) had little effect on study power. A dampened relationship with NRM at high $AUC_{t_0-\infty}$ did introduce loss of power, but uncertainty of the survival models was generally lower for these exposures, which makes it less likely that the assumed deviation holds true in the real situation. Reasonable deviations from the currently used exposure-response relationship are therefore not expected to have an extreme influence on study outcome.

Furthermore, we tested how unsuccessful TDM (10%) would affect power, which was negligible. Several other possible protocol deviations, such as the possibility of missing samples, were already implemented in the simulations and were based on our local TDM experience. Nevertheless, one could also expect an iatrogenic form of non-adherence to the TDM-based dosing recommendation, especially when the suggested adjustment is marginal or, by contrast, extremely large. Omitting extreme adjustments will vastly reduce power, and should thus be avoided at all cost. This can be done by clearly stating the dosing alteration procedure in the trial protocol.

In summary, CTS indicate the possibility of a substantial improvement in HCT survival by using TDM-based dosing. Furthermore, the proposed CTS framework can be used for prediction of efficacy for other treatment alter-

ations in HCT. Based on the findings in this study, prospective evaluation studies can be designed to provide definite proof for the added value of individualized fludarabine dosing.

References

1. Hahn, T., McCarthy P. L., J., Hassebroek, A., Bredeson, C., Gajewski, J. L., Hale, G. A., *et al.* Significant improvement in survival after allogeneic hematopoietic cell transplantation during a period of significantly increased use, older recipient age, and use of unrelated donors. *J Clin Oncol* **31,** 2437–49 (2013).
2. Gooley, T. A., Chien, J. W., Pergam, S. A., Hingorani, S., Sorror, M. L., Boeckh, M., *et al.* Reduced mortality after allogeneic hematopoietic-cell transplantation. *N Engl J Med* **363,** 2091–101 (2010).
3. Admiraal, R., Nierkens, S., de Witte, M. A., Petersen, E. J., Fleurke, G. J., Verrest, L., *et al.* Association between anti-thymocyte globulin exposure and survival outcomes in adult unrelated haemopoietic cell transplantation: a multicentre, retrospective, pharmacodynamic cohort analysis. *Lancet Haematol* **4,** e183–e191 (2017).
4. Bartelink, I. H., Lalmohamed, A., van Reij, E. M., Dvorak, C. C., Savic, R. M., Zwaveling, J., *et al.* Association of busulfan exposure with survival and toxicity after haemopoietic cell transplantation in children and young adults: a multicentre, retrospective cohort analysis. *Lancet Haematol* **3,** e526–e536 (2016).
5. Ivaturi, V., Dvorak, C. C., Chan, D., Liu, T., Cowan, M. J., Wahlstrom, J., *et al.* Pharmacokinetics and Model-Based Dosing to Optimize Fludarabine Therapy in Pediatric Hematopoietic Cell Transplant Recipients. *Biol Blood Marrow Transplant* **23,** 1701–1713 (2017).
6. Andersson, B. S., Thall, P. F., Valdez, B. C., Milton, D. R., Al-Atrash, G., Chen, J., *et al.* Fludarabine with pharmacokinetically guided IV busulfan is superior to fixed-dose delivery in pretransplant conditioning of AML/MDS patients. *Bone Marrow Transplant* **52,** 580–587 (2017).
7. Langenhorst, J. B., Dorlo, T. P. C., van Maarseveen, E. M., Nierkens, S., Kuball, J., Boelens, J. J., *et al.* Population Pharmacokinetics of Fludarabine in Children and Adults during Conditioning Prior to Allogeneic Hematopoietic Cell Transplantation. *Clinical pharmacokinetics.* Epub ahead of print (Oct. 2018).
8. Long-Boyle, J. R., Green, K. G., Brunstein, C. G., Cao, Q., Rogosheske, J., Weisdorf, D. J., *et al.* High fludarabine exposure and relationship with treatment-related mortality after nonmyeloablative hematopoietic cell transplantation. *Bone Marrow Transplant* **46,** 20–6 (2011).

9. Sanghavi, K., Wiseman, A., Kirstein, M. N., Cao, Q., Brundage, R., Jensen, K., *et al.* Personalized fludarabine dosing to reduce nonrelapse mortality in hematopoietic stem-cell transplant recipients receiving reduced intensity conditioning. *Transl Res* **175,** 103–115 e4 (2016).

10. Langenhorst, J. B., van Kesteren, C., van Maarseveen, E. M., Dorlo, T. P. C., Nierkens, S., Lindemans, C. A., *et al. Fludarabine exposure in the conditioning prior to allogeneic hematopoietic cell transplantation predicts outcomes* Submitted. Dec. 2018.

11. Dhakal, B., Vesole, D. H. & Hari, P. N. Allogeneic stem cell transplantation for multiple myeloma: is there a future? *Bone marrow transplantation* **51,** 492–500 (4 Apr. 2016).

12. Yanada, M., Emi, N., Naoe, T., Sakamaki, H., Iseki, T., Hirabayashi, N., *et al.* Allogeneic myeloablative transplantation for patients aged 50 years and over. *Bone Marrow Transplant* **34,** 29–35 (2004).

13. Admiraal, R., van Kesteren, C., Jol-van der Zijde, C. M., Lankester, A. C., Bierings, M. B., Egberts, T. C., *et al.* Association between anti-thymocyte globulin exposure and CD4+ immune reconstitution in paediatric haemopoietic cell transplantation: a multicentre, retrospective pharmacodynamic cohort analysis. *Lancet Haematol* **2,** e194–203 (2015).

14. Milano, F., Gooley, T., Wood, B., Woolfrey, A., Flowers, M. E., Doney, K., *et al.* Cord-Blood Transplantation in Patients with Minimal Residual Disease. *N Engl J Med* **375,** 944–53 (2016).

15. Murata, M., Nishida, T., Taniguchi, S., Ohashi, K., Ogawa, H., Fukuda, T., *et al.* Allogeneic transplantation for primary myelofibrosis with BM, peripheral blood or umbilical cord blood: an analysis of the JSHCT. *Bone Marrow Transplant* **49,** 355–60 (2014).

16. Ljungman, P. The role of cytomegalovirus serostatus on outcome of hematopoietic stem cell transplantation. *Curr Opin Hematol* **21,** 466–9 (2014).

17. Wahid, S. F. A. Indications and outcomes of reduced-toxicity hematopoietic stem cell transplantation in adult patients with hematological malignancies. *International journal of hematology* **97,** 581–598 (5 May 2013).

18. Fielding, A. K., Richards, S. M., Chopra, R., Lazarus, H. M., Litzow, M. R., Buck, G., *et al.* Outcome of 609 adults after relapse of acute lymphoblastic leukemia (ALL); an MRC UKALL12/ECOG 2993 study. *Blood* **109,** 944–50 (2007).

19. Poon, L. M., Bassett R., J., Rondon, G., Hamdi, A., Qazilbash, M., Hosing, C., *et al.* Outcomes of second allogeneic hematopoietic stem cell transplantation for patients with acute lymphoblastic leukemia. *Bone Marrow Transplant* **48,** 666–70 (2013).

20. Ruutu, T., de Wreede, L. C., van Biezen, A., Brand, R., Mohty, M., Dreger, P., *et al.* Second allogeneic transplantation for relapse of malignant disease: retrospective analysis of outcome and predictive factors by the EBMT. *Bone Marrow Transplant* **50,** 1542–50 (2015).

21. Huo, M. R., Xu, L. P., Li, D., Liu, D. H., Liu, K. Y., Chen, H., *et al.* The effect of HLA disparity on clinical outcome after HLA-haploidentical blood and marrow transplantation. *Clin Transplant* **26,** 284–91 (2012).

22. Rondon, G., Saliba, R. M., Khouri, I., Giralt, S., Chan, K., Jabbour, E., *et al.* Long-term follow-up of patients who experienced graft failure postallogeneic progenitor cell transplantation. Results of a single institution analysis. *Biol Blood Marrow Transplant* **14,** 859–66 (2008).

23. Olsson, R., Remberger, M., Schaffer, M., Berggren, D. M., Svahn, B. M., Mattsson, J., *et al.* Graft failure in the modern era of allogeneic hematopoietic SCT. *Bone Marrow Transplant* **48,** 537–43 (2013).

24. Fedele, R., Messina, G., Martinello, T., Gallo, G. A., Pontari, A., Moscato, T., *et al.* Tolerability and efficacy of busulfan and fludarabine as allogeneic pretransplant conditioning therapy in acute myeloid leukemia: comparison with busulfan and cyclophosphamide regimen. *Clin Lymphoma Myeloma Leuk* **14,** 493–500 (2014).

25. Lee, J. H., Joo, Y. D., Kim, H., Ryoo, H. M., Kim, M. K., Lee, G. W., *et al.* Randomized trial of myeloablative conditioning regimens: busulfan plus cyclophosphamide versus busulfan plus fludarabine. *J Clin Oncol* **31,** 701–9 (2013).

26. Liu, H., Zhai, X., Song, Z., Sun, J., Xiao, Y., Nie, D., *et al.* Busulfan plus fludarabine as a myeloablative conditioning regimen compared with busulfan plus cyclophosphamide for acute myeloid leukemia in first complete remission undergoing allogeneic hematopoietic stem cell transplantation: a prospective and multicenter study. *J Hematol Oncol* **6,** 15 (2013).

General discussion

10.1 Aim

The aim of this thesis was to quantify predictable and unpredictable sources accounting for variable outcome after allogeneic hematopoietic cell transplantation (HCT). Donor selection, conditioning, engraftment and immune reconstitution were highlighted as the crucial parts of the procedure (figure 1). We hypothesized that for each component of HCT opposing risks need to balanced. In this following chapter, the results of the several studies described in this thesis will be discussed and put in context. The text is further subdivided in an immunological and a pharmacological part, that are respectively introduced in section 10.2 and 10.3. After each introduction the parts are elaborated on in the corresponding subsections 10.2.1-10.2.1 (immunology) and 10.3.1-10.3.4 (pharmacology).

10.2 **Donor selection & Immune reconstitution:** Monitor and regulate immunogenicity after transplantation

It has been shown extensively that the dynamics of immune reconstitution after HCT are strongly related to treatment outcome. Furthermore, in malignant diseases, the donor graft is a form of immunotherapy itself, called the graft-versus-leukemia effect. Both these crucial HCT components are interrelated and highly dependent on donor selection.
Historically, matched sibling donors (MSD) have been the preferred cell source, but unfortunately such donor is unavailable for the majority of patients. Alternatives, such as cord blood (CB) and matched unrelated donors (MUD), have mostly been used in absence of an available MSD. The major concern of the alternative donor sources has been increased probability of graft-versus-host disease (GvHD) and subsequent non-relapse mortality (NRM). However, improved GvHD-prophylaxis and conditioning have led to reduction of GvHD and NRM in the last decades, resulting in better MUD and CB transplantation outcomes. Considering that GvHD and graft-versus-leukemia effect are usually opposing probabilities, improved GvHD control might result in more favorable outcomes in the more immunologically active CB transplants. When baseline NRM probability is substantially low, the difference between cell sources in associated absolute NRM risks becomes less problematic and disease persistence or relapse develops into the dominant driver of HCT survival. Therefore, the donor source that reduces the risk for relapse the most is preferable.
Besides creating an optimal balance between GvHD and graft-versus-leukemia

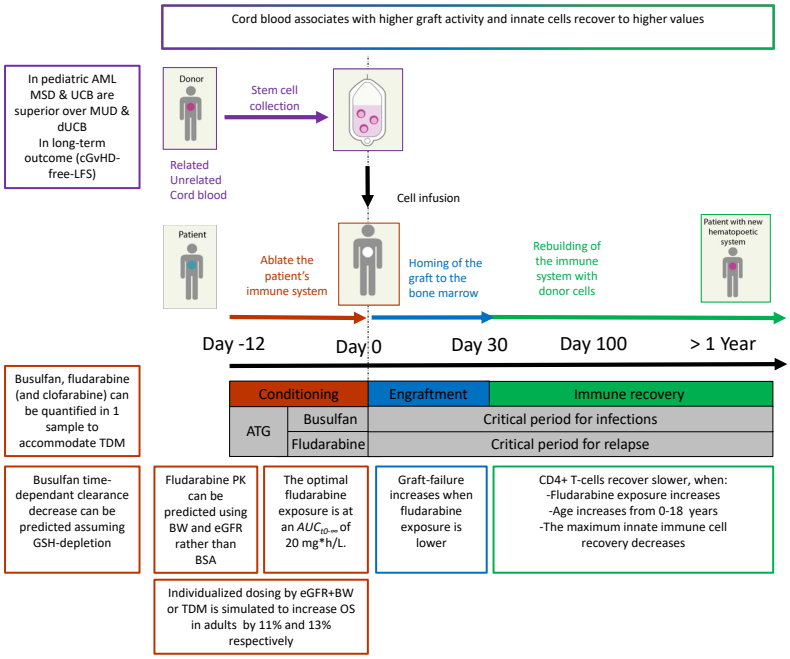

Figure 1: General scheme of UMC-U myeloablative hematopoietic cell transplantation setting. Different colours indicate the crucial parts of the procedure. Text boxes highlight the important findings of this thesis.
ATG = Anti-thymocyte globulin.

effect, the choice of cell source may also influence the dynamics of immune reconstitution. An optimal reconstitution is important as during this period, the body is relatively unprotected against pathogens. Furthermore, in malignant settings, the graft-versus-leukemia effect has to take place early as possible to prevent relapse. Though all immune cells play a role in protection against pathogens, primarily donor T-lymphocytes have been found crucial in exhibiting graft-versus-leukemia effect. These cells, however, are part of the adaptive immune system that reconstitutes relatively late after HCT (1 month to 1 year). In contrast, cells of the innate immune system usually grow out earlier after transplantation and might already predict the dynamics of adaptive recovery.

10.2.1 Cell source & HCT outcome

In chapter 2 the outcome of HCT was compared between cell sources in children and young adults transplanted for acute myeloid leukemia. The cell sources were subdivided in the golden standard MSD and alternative donors MUD and CB. Patients preferably received CB from a single donor, but for several young adults a double CB donor was necessary to achieve sufficient cell numbers.
The study showed that both MSD and single CB, had the best long-term outcome, as quantified by a composite endpoint consisting of chronic GvHD, leukemia relapse, graft failure and death (cGvHD-LFS). Inferior outcome in double CB grafts was mainly caused by increased NRM, while after MUD transplantation generally more chronic GvHD cases were seen.
Zooming in on the alternative donors in this study, CB and MUD, one can speculate that the difference in immunologic activity causes the difference in outcome. For example, MUD was associated with lower acute GvHD rates and subsequent less frequent NRM, compared to CB. However, relapse probability was highest after MUD transplant, leading to similar leukemia-free survival (composed of relapse and NRM) as compared with the single CB transplantation. In the double CB, the increased probability of NRM exceeded the decreased relapse rate, causing leukemia-free survival to be inferior in double CB. Possibly the introduction of two different donors, with the double CB, causes excessive immunological activity. These findings imply that the immunological activity increases from MUD, to single onto double CB. Increased immunological activity causes acute GvHD but also superior graft-versus-leukemia effect. Double CB leads to more NRM, implying that the balance of GvHD versus graft-versus-leukemia is not optimal. In the typical patient, leukemia-free survival is similar between CB and MUD, but higher risk patients, such as those with minimal residual disease at the moment of transplantation, would benefit even more from the superior graft-versus-leukemia effect of CB. Furthermore, when the early acute phase of GvHD is overcome in CB, this generally does not translate into chronic GvHD.
In the era of reduced early NRM due to improved supportive care, CB should be the first alternative to MSD. When cell doses are insufficient from a single CB, a MUD is preferred over double CB.

10.2.2 Cell source & immune reconstitution

In chapter 3 the relationship between dynamics of innate immune reconstitution of different cell types and graft sources was explored. Subsequently, innate immune reconstitution was related to adaptive immune reconstitution. A model was developed where innate cell recovery (NK-cells, neutrophils, monocytes) was described using a single latent variable, representing overall immune reconstitution. Using this model, it was found that the innate cell counts after CB transplantation recovered more quickly compared to transplantation with bone marrow sources (MSD & MUD). In a separate analysis, this was found to be mainly attributable to increased proliferation of CB grafts early after HCT. The higher innate cell counts, in turn, were predictive for higher probability of early CD4+ T-cell reconstitution.
Overall, these results suggest that the probability of adequate adaptive immune recovery can be predicted in an earlier stage, by monitoring the innate immune cells. Innate recovery can be enhanced *a priori* by using a CB graft, thereby indirectly influencing adaptive recovery. The modeling method, using a single latent variable describing the underlying immunological process to describe multiple cell types, can be applied to adaptive recovery as well. Until now, CD4+ T-cells have been used as the main predictor for survival in HCT, while other adaptive cell types might contribute to the positive effects attributed to CD4+ T-cells as well. Perhaps multiple lineages of the adaptive recovery can be described by a similar latent variable, which can be related to outcomes in order to explore the overall contribution of adaptive recovery to HCT survival.

10.3 **Conditioning:** Optimizing the busulfan-fludarabine regimen

Busulfan-fludarabine with or without rATG is one of the most frequently used conditioning regimens in HCT nowadays. Conditioning intensity should be sufficient to ensure engraftment and minimize the probability of disease relapse, however, with acceptable risk of toxicity. It has previously been demonstrated that an optimized targeting of the exposure to the agents used in conditioning regimens, is of crucial importance for treatment outcome. For both rATG and busulfan optimal dosing regimens have been described in this setting, thus the focus for this thesis has been to strengthen and widen the current platform of individualized dosing.

10.3.1 The basis for busulfan-fludarabine conditioning in HCT

In chapter 4 it was shown that busulfan-fludarabine seems to be the superior regimen compared to (otherwise frequently used) regimens combining cyclophosphamide with either busulfan or total body irradiation (TBI). This superiority is dependent on rational dosing, based on clinical pharmacology of these agents. Busulfan has a narrow therapeutic window, thus dosing needs to guide towards the optimal plasma exposure to have maximal efficacy with acceptable toxicity. Indeed, busulfan-cyclophosphamide with the unpredictable exposure after oral administration of busulfan (mainly used before introduction of an intravenous form in 2000) was shown to be inferior to cyclophosphamide-TBI, while intravenous busulfan-cyclophosphamide was found equal or superior. Maximal efficacy was achieved using therapeutic drug monitoring (TDM), with 20% survival gain compared to dosing based on body-weight, mainly through reduced toxicity. Fludarabine has synergystic anti-leukemic activity when co-administered with busulfan, thus these agents need to be dosed on the same day. As such, busulfan-fludarabine has been shown to associate with similar relapse rates compared to busulfan-cyclophosphamide, but toxicity and NRM were lower. Therefore, busulfan-fludarabine may be the preferred regimen to be used in conditioning prior to HCT.

10.3.2 Bio-analysis of busulfan, fludarabine and clofarabine for TDM

In chapter 5, a liquid chromatography + mass spectometry (LCMS) method to quantify busulfan, fludarabine (and clofarabine) in a single plasma sample is described. The main purpose of this method was to support 1) quantification of the dose-exposure-response relationships of the full regimen and 2) implementation of these findings in the clinic using multi-agent TDM. To make this method efficient, a chromatographical challenge had to be overcome, due to the large structural differences of busulfan compared to purine analogues fludarabine and clofarabine. This was solved by using a gradient from 95% water to 90% acetonitrile after elution of fludarabine and clofarabine, to accommodate timely elution of busulfan. Furthermore, standard addition was used to adjust for matrix effects, as stable isotopically labeled internal standards were not readily available for fludarabine and clofarabine. This resulted in a fast and efficient method, validated according to European Medicine Agency guidelines, to quantify fludarabine, clofarabine and

busulfan in a single plasma sample.

10.3.3 Busulfan

A semi-mechanistic model to aid in busulfan TDM was developed in chapter 6. Previously is has been shown that clearance of busulfan decreases during the 4-day treatment. The hypothesis that this reduction in clearance was due to glutathione depletion which was implemented in a semi-mechanistic model. As a result, patients with a high initial busulfan clearance showed a more pronounced decrease compared to patients with a lower initial clearance. This effect increased after an age of 40, most likely caused by an age dependent decrease in glutathione concentration.
The constructed semi-mechanistic model was shown to already increase target attainment compared to a previously used empirical model in a TDM simulation. Though the improvement in target attainment was small (1-5%), this is relevant to patients with an exposure outside the target area, since they have an increase in probability of therapy failure (relapse, graft failure) or severe toxicity.
Despite the possible clinical improvement, from a biological perspective the presented model should mainly be seen as a starting point in elucidating the mechanism behind the decrease in busulfan clearance. As these data strongly suggest glutathione depletion to be a viable hypothesis, further research should focus on this mechanism. Firstly, actual glutathione levels should be measured and implemented into the developed model. Secondly, as paracetamol exposure is expected to influence GSH, this effect should be added as well. Henceforth, new simulations can be used to see if glutathione levels and/or paracetamol usage can be used to further improve TDM towards a point where all patients achieve an exposure within the target area.

10.3.4 Fludarabine

In chapter 7 we showed highly variable fludarabine plasma exposures after standard dosing with 160 mg/m^2. Body weight and estimated glomerular filtration rate (eGFR) were found to better predict fludarabine clearance than body surface area, the current basis for fludarabine dosing.
This observed effect is important for the HCT setting, but may also be applicable to other settings where fludarabine is being used, such as chimeric antigen receptor T-cell therapy, gene-transduced autologous HCT, and chronic lymphocytic leukemia. Especially in chronic lymphocytic leukemia, which

is mainly a disease in elderly patients, accounting for the decline in renal function may help in reducing fludarabine associated toxicity such as prolonged lymphopenia.
In HCT, the variability in fludarabine plasma exposures is further complicated by previous findings that under-exposure associates with graft failure, and over-exposure with NRM, potentially resulting in decreased survival. Indeed, in chapter 8 these effects were confirmed: an optimal fludarabine exposure was found at an $AUC_{t_0-\infty}$ of 20 mg*h/L. The increased NRM at high exposures appeared to be mainly caused by impaired immune reconstitution resulting in death due to infections.
An unanswered question is by what mechanism high exposures to fludarabine prior to HCT affects the reconstitution of infused T-cells at the time of HCT. We hypothesized that this could be related to an overly damaged thymus, due to fludarabine exposure, that can no longer fulfill its role in supporting homeostatic expansion of CD4+ T-cells. This has previously been proposed as a mechanism for GvHD-related delay in immune reconstitution, and could be explored for fludarabine in the future. In addition, cytotoxic fludarabine moieties could persist in tissue after HCT. For the latter hypothesis, sensitive magnetic resonance imaging techniques could be used to quantify the distribution of fluorine containing fludarabine in tissues other than blood. The strong relationship between fludarabine systemic exposure and treatment outcome combined with the finding that fludarabine clearance can be predicted by eGFR and body weight allows for possible improvement through individualized dosing. Nevertheless, the target and concurrently targeted individualized dosing should be prospectively confirmed in an independent cohort.
To aid in design of such a study, a framework for clinical trial simulations (CTS) was developed in chapter 9. Cause-specific-hazard models were used to relate fludarabine exposure and other relevant predictors to the main HCT end-points: graft failure, relapse and NRM. Markovian elements were added to model the transition from the events graft failure and relapse to either death or survival. Next, the framework was used to aid in rational design of a randomized controlled trial comparing individualized dosing regimens to BSA-based dosing in an adult cohort. We found that an increase in overall survival of 13% could be achieved by using TDM-based dosing and that 70 patients per arm would be necessary to achieve 80% power with NRM as a primary end-point. If superiority of individualized fludarabine dosing is confirmed in adults, the CTS platform can also be used to predict the best strategies in other settings, such as pediatric HCT.

10.4 General conclusions & Perspectives

Optimizing treatment outcome in HCT is challenging. The low number of patients with a variety of indications and ages complicates the use of the more standard head-to-head comparisons as patient groups are hardly homogeneous.
Using modeling approaches, the process of optimizing the conditioning regimen has resulted in substantial improvements over the past 10 years. This has led to the solid basis of exposure guided dosing for busulfan and rATG as it is today. Combined with the findings in this thesis, the optimal use of the busulfan-fludarabine regimen as such has been defined. The dosing for both agents should be guided by their respective optimal exposures of 90 and 20 mg*h/L. The most precise way of doing so is using TDM, aided by the proposed LCMS method and the, in this thesis constructed, pharmacokinetic models.
The immunologic aspects of HCT, such as donor selection, immune reconstitution, graft-versus-leukemia effect, and GvHD, are more complex and full understanding will take more time and research. Especially, since all these processes intertwine, modeling approaches will need to be more sophisticated to take into account such interactions. First steps have been taken with the latent variable model in this thesis. In addition, the solid background of predictable conditioning will aid in elucidating these subtle, variable but important immunological effects.
In conclusion, 'model-based therapy in allogeneic hematopoietic cell transplantation' ensures safe yet effective application of HCT.

Appendices

Appendix A English summary

Allogeneic hematopoietic cell transplantation (HCT) is a high-risk treatment for a variety of malignant and benign disorders with a curative intention. Malignancies treated with HCT are relapsed cancers with an origin in immune cells and/or the bone marrow, such as leukemia (acute/chronic myeloid/lymphocytic), lymphoma, plasma cell disorders and myelodysplastic syndrome. While these relapsed malignancies are prevalent in both children and adults, the benign disorders are mostly seen in children and include bone marrow failures, immune deficiencies, metabolic-, and autoimmune diseases. Albeit the curative intention for these otherwise often infaust diagnoses, the procedure comes at a high cost, with an overall mortality of around 40%. In all settings, a substantial part of deaths are attributable to the treatment.

The HCT procedure globally consists of four steps: donor selection, conditioning, transplantation, and immune reconstitution. Donor selection, transplantation and immune reconstitution together form the immunotherapeutic element of HCT. The treatment namely consists of two components: the diseased bone marrow is replaced by the donor cells and the donor cells themselves yield an immune response against any residual disease. The latter is mainly important to prevent leukemia relapse. To maximally exhibit the immunotherapeutic effect and to minimize the period of immune deprivation, it is imperative that the immune system reconstitutes quickly and in a controlled fashion. The latter is necessary due to a downside of the immunological effects of HCT, where potentially life threatening graft-versus-host disease (GvHD) counterbalances the positive immunological effects. GvHD arises when the immune cells from the infused graft recognize host-cells as hostile and start to form an immunological response. Finding such an optimal balance is the focus of the first part of this thesis.

Conditioning takes place between donor selection and immune reconstitution. Conditioning is a process, where high dose chemotherapy, sometimes combined with irradiation, is administered to ablate the bone marrow and immune system of the patient. Due to the high dose of toxic compounds, adequate pharmacological knowledge regarding the mechanism of action and optimal dose is essential. The bone marrow and immune system need to be sufficiently ablated, but unnecessary toxicity needs to be prevented. Here again, an optimal balance needs to be founds, to generate sufficient efficacy, while omitting serious adverse events. In the second part of this thesis this optimal balance in conditioning is studied.

A 1 Immunology of transplantation

When HCT is necessary for children with acute myeloid leukemia there remains debate about the best stem cell source. Post-HCT relapse is a common cause of mortality and complications such as chronic GvHD are debilitating and also life-threatening. To compare post-HCT outcomes of different donor sources, in chapter 2 a retrospective analysis was performed of consecutive transplants conducted in several international centers from 2005-2015. A total of 317 patients were studied, of whom 19% matched sibling donor, 23% matched unrelated donor, 39% cord blood and 19% double cord blood recipients. The median age at transplant was 10 years (range 0.42-21) and median follow-up was 4.74 years (range 4.02-5.39). As baseline characteristics differed amongst the cell sources, comparisons were made while controlling for these patient, transplant, and disease characteristics. There were no differences in relapse, leukemia-free survival, or non-relapse mortality. Double cord blood recipients had inferior survival compared to matched sibling donor, but all other comparisons showed similar overall survival. Despite the majority of cord blood transplants being human leukocyte antigen mismatched, usually a risk factor for the immunological downsides of HCT, the frequencies of chronic GvHD with this cell source were low compared to those transplanted with a well matched unrelated donor (hazard-ratio (HR)= 0.3; 0.14-0.67; $p = 0.02$). The composite measure of chronic GvHD and leukemia free survival, which represents both the quality of life and risk of mortality, was significantly better after cord blood compared to the matched unrelated donor transplantation (HR=0.56; 0.34-1; $p = 0.03$). In conclusion, the use of cord blood is an excellent donor choice for pediatric acute myeloid leukemia patients when a matched sibling cannot be identified. The optimal balance between immunotherapeutic effect (low relapse) and limited immunological downsides (low chronic GvHD) is seen with cord blood. This could ultimately lead to long-term survival with superior quality of life.

With these apparent immonological advantages of CB it would be of interest to know if these advantages are related to the timing of immune reconstitution. Especially predicting early CD4+ T-cell reconstitution is of clinical interest, as this parameter directly associates with survival after HCT. However, innate, not adaptive immune cells are the first to recover after transplantation. Reports of innate immune cell recovery and their relation to adaptive recovery after HCT, are largely lacking. Therefore, in chapter 3 it was evaluated whether innate recovery relates to CD4+ T-cell

reconstitution probability, and differences between innate recovery after cord blood transplantation and bone marrow transplantation were investigated. For this, a multivariate, non-linear mixed-effects model was developed describing monocyte, neutrophil and natural killer-cell recovery after transplantation using a single latent variable. In total, 205 patients undergoing a first HCT between 2007-2016, were included of whom 76 received a bone marrow transplantation and 129 a cord blood transplantation. Patients had a median age of 7.3 years (range 0.16-23) . Innate recovery was highly associated with CD4+ T-cell reconstitution probability. Monocyte, neutrophil, and natural killer-cell recovery reached higher levels during the first 200 days after cord blood transplantation compared to bone marrow transplantation. The higher innate recovery after cord blood transplantation may be explained by increased proliferation capacity, measured by Ki-67 expression, of innate cells inside cord blood grafts compared to bone marrow grafts ($p = 0.041$), and of innate cells *in vivo* after cord blood transplantation compared to bone marrow transplantation ($p = 0.048$). Patients with increased innate recovery after either cord blood transplantation or bone marrow transplantation (BMT) had received grafts with higher proliferating innate cells (CB; $p = 0.004$, BM; $p = 0.01$, respectively).
These findings implicate the possibility to use early innate immune monitoring for predicting the probability of CD4+ T-cell reconstitution after HCT. In addition, cord blood transplantation associates with higher innate recovery due to increased proliferative capacity.

A 2 Optimizing the conditioning

Busulfan and fludarabine are increasingly combined in conditioning regimens prior to HCT. Therefore, in chapter 4 the suitability, clinical pharmacology and the optimal use of the busulfan-fludarabine regimen in HCT is summarized. The busulfan-fludarabine regimen has been proven a less toxic alternative to the historically used regimens cyclophosphamide combined with total body irradiation (TBI) and the busulfan-cyclophosphamide regimen. Owing to the synergistic anti-leukemic effect of busulfan and fludarabine, while preserving the immunotherapeutic effect of HCT, the relapse rates remain similar compared to the busulfan-cyclophosphamide regimen and cyclophosphamide-TBI. Toxicity and consequently non-relapse mortality are substantially lower for the busulfan-fludarabine regimen. As is common for high-dose chemotherapeutics, the therapeutic windows for both drugs are narrow. A minimal exposure of both agents is essential to prevent graft failure (busulfan and fludarabine) and relapse (busulfan). Over-exposure to

either agent leads to an unnecessary increase in toxicity and even non-relapse mortality. Busulfan target attainment can be achieved in increasing order by intravenous dosing based on body weight, dosing based on population-pharmacokinetic models and therapeutic drug monitoring (TDM). Current dosing for fludarabine has been shown to yield unfavorable exposures in a large part of the population, thus, similarly to busulfan, new dosing algorithms and TDM are needed.

To evaluate and possibly adjust the dose of all agents in conditioning, a simultaneous analysis of busulfan, fludarabine (F-ara-A, circulating metabolite of fludarabine as the prodrug F-ara-AMP) and clofarabine in one analytical run would be of great interest. However, this is a chromatographical challenge due to the large structural differences of busulfan compared to fludarabine and clofarabine. Furthermore, stable isotope labelled standards (SILS), that would ideally be used to adjust for matrix effect, are not readily available for fludarabine and clofarabine .
In chapter 5 a fast analytical method, using standard addition instead of SILS, for the simultaneous analysing of busulfan, clofarabine and fludarabine with liquid chromatography-tandem mass spectrometry (LCMS) is described. The analytical method was validated in accordance with European Medicines Agency guidelines.
In conclusion, an efficient method for the simultaneous quantification of busulfan, clofarabine and fludarabine in plasma was developed. Thus, a suitable method is available to be applied in multi-agent TDM, to further optimize conditioning with high efficiency.

For busulfan it is known for a long time, that the therapeutic window is narrow. TDM is necessary for precise dosing, but is complicated by a previously described intra-individual reduction in busulfan clearance over time. In addition, this reduction in clearance is associated with extensive unpredictable inter-individual variability.
Therefore, in chapter 6 a semi-mechanistic model was developed to predict this reduction in busulfan clearance and aid in more precise dosing. As conjugation to glutathione is the rate-limiting step for busulfan metabolism, depletion of glutathione was hypothesized to cause the decrease in clearance and included in the model. In this model, the amount of glutathione was assumed to decrease proportional to busulfan metabolism and busulfan clearance was assumed to be proportional to the amount of glutathione. As older age has been related to lower glutathione levels, this was tested as a covariate in the semi-mechanistic model. Lastly, a simulation was performed

comparing TDM based on the semi-mechanistic model to TDM based on a previously developed model, that emperically described the decrease in busulfan clearance.
In both models, a similar clearance decrease of 7% (range: -82% to 44%) was found with a proportionality to busulfan metabolism. After 40 years of age, the time-effect increased proportionally each year, causing the effect to increase more than a 2-fold over the observed age-range (0-73 years). Compared to the empirical model, the final semi-mechanistic model increased target attainment in the TDM-simulation from 74% to 76%, mainly through better predictions in adult patients.
These data suggest that the time-dependent decrease in busulfan clearance may be related to gluthathione depletion. This effect is more clear above the age of 40 years and is proportional to busulfan metabolism. The newly constructed semi-mechanistic model could be used to further improve TDM-guided exposure target attainment of busulfan in patients undergoing HCT.

In contrast to busulfan, not much is known about the optimal fludarabine dosing for the wide age- and weight range in HCT. Currently, linear body-surface-area-based dosing is still used. To provide more rationale to fludarabine dosing, in chapter 7 the population pharmacokinetics of fludarabine in HCT recipients of all ages was described.
From 258 HCT recipients aged 0.3 to 74 years, 2605 samples were acquired on day 1 (42%), day 2 (17%), day 3 (4%) and day 4 (37%) of conditioning. Herein, the circulating metabolite of fludarabine was quantified and derived concentration-time data were used to build a population pharmacokinetic model using non-linear mixed effects modelling.
Variability was extensive, where area under the plasma-concentration-time curve (AUC) from the first dose until infinity ($AUC_{t_0-\infty}$) ranged from 10-66 mg*h/L. Not body-surface-area, but actual body weight with standard allometric scaling was found to be the best body-size descriptor for all pharmacokinetic parameters. Estimated glomerular filtration rate (eGFR) was included as a descriptor of renal function. Thus, clearance was differentiated into a non-renal (3.24±20% L/h/70kg) and renal (eGFR*0.782±11% L/h/70kg) component. Actual body weight and eGFR proved to be important predictors of fludarabine pharmacokinetic. Therefore, current linear body-surface-area-based dosing leads to highly variable exposure.

This extensive variability with body-surface-area-based dosing potentially leads to unpredictable HCT outcome. Therefore, in chapter 8 the relation between fludarabine exposure and clinical outcomes after HCT was studied.

A retrospective, pharmacokinetic-pharmacodynamic analysis was conducted with data from patients undergoing HCT with 160 mg/m^2 fludarabine as part of a myeloablative conditioning with targeted busulfan and rabbit anti-thymocyte globulin between 2010 and 2016. Fludarabine exposure (AUC) was calculated for each patient using aforementioned population pharmacokinetic model and related to 2-year-event-free survival by means of (parametric)-time-to-event-models. Relapse, non-relapse mortality, and graft failure were events of interest.
In total, 192 patients were included (68 benign and 124 malignant disorders). The optimal fludarabine exposure was determined as a $AUC_{t_0-\infty}$ of 20 mg*h/L. Estimated 2-year event-free survival in the group with optimal exposure was 67%. In the over-exposed group, event-free survival was lower, due to approximately doubled risk of non-relapse mortality, which also associated with strongly impaired immune-reconstitution. The risks of non-relapse mortality and graft failure were highly increased (2- and 5-fold respectively) in the under-exposed group. No relationship with relapse was found.
Fludarabine-exposure is a strong predictor of survival after HCT, stressing the importance of precise fludarabine dosing. Individualized dosing, based on weight and eGFR or TDM aimed at achieving the optimal exposure might improve HCT survival.

Knowing that the current fludarabine dosing strategy leads to highly variable exposures associated with unfavourable outcome, this offers the potential to individualize dosing for further treatment optimization. As all results were obtained retrospectively, a randomized controlled trial would provide definite proof for superiority of individualized fludarabine dosing. It is imperative to rationally design such a trial, to expose a minimum number of patients to potentially inferior outcomes of either dosing regimens, but still provide a decisive answer.
Therefore, in chapter 9 a study was performed aiming to find optimal trial design to evaluate potential superiority of novel dosing strategies. For this, the previously developed pharmacokinetic (chapter 7) and cause-specific hazard-models (chapter 8) were expanded by Markovian elements to model the transition from relapse/graft failure to death, thereby allowing simulation of the full procedure. Randomized controlled trials were simulated comparing conventional dosing (160 mg/m^2) to either dosing based on eGFR and actual body weight or TDM-guided-dosing. The number of patients required to achieve 80% power was calculated for non-relapse mortality and overall survival. The optimal trial had TDM-guided-dosing as individualized

arm and non-relapse mortality as primary outcome. A power of 82% was achieved with n=70 per dosing arm. This number of patients was markedly lower compared to n=160 with TDM-guided-dosing with overall survival as an outcome and n=120 for actual body weight and eGFR based dosing and non-relapse mortality as primary outcome. The expected absolute overall survival gain for TDM compared to conventional dosing was 13%, which was caused by a marked decrease in non-relapse mortality from 21% to 5.7%.

A 3 Conclusions

Throughout this thesis, several aspects of HCT have been investigated, both immonological and pharmacological.
The immunologic aspects of HCT, such as donor selection, transplantation, immune reconstitution, graft-versus-leukemia effect, and GvHD, are more complex and full understanding will take more time and research. In this work, different cell sources were compared regarding outcome and immune reconstitution. In both cases cord blood was shown to compare favourably to bone marrow derived sources such as a matched unrelated donor. In addition, a relationship between innate and adaptive recovery was identified, allowing to earlier predict adequate immune reconsitution.
For the pharmacological part, previous work regarding individualized busulfan dosing combined with the findings in this thesis, have allowed to define the optimal use of the busulfan-fludarabine regimen as such. The dosing for both agents should be guided by their respective optimal exposures of 90 and 20 mg*h/L. The most precise way of doing so is using TDM, aided by the proposed LCMS method and the, in this thesis constructed, pharmacokinetic models.
By identifying the optimal use of the conditioning regimen and describing as well as predicting the immunological effects of HCT, 'model-based therapy in allogeneic hematopoietic cell transplantation' ensures safe yet effective application of HCT.

Appendix B Nederlandse samenvatting

Allogene stamceltransplantatie (SCT) is een potentieel curatieve behandeling voor zowel kwaadaardige als goedaardige bloedaandoeningen. De kwaadaardige aandoeningen waar SCT als behandeling voor wordt gebruikt, bestaan met name uit recidieven van leukemie, die niet meer met conventionele chemotherapie genezen kunnen worden. De goedaardige SCT indicaties zijn veelal aangeboren afwijkingen, zoals bloedarmoedes, immuundeficiënties, auto-immuun-, en metabole ziektes.
De SCT procedure bestaat uit vier stappen: donorselectie, conditionering, de transplantatie en immuun herstel. De donorselectie en het immuunherstel vormen samen het immunotherapeutische onderdeel van de behandeling. Het therapeutisch effect van SCT bestaat namelijk uit twee componenten: het zieke beenmerg en/of immuunsysteem wordt vervangen en vervolgens richt dit nieuwe immuunsysteem zich ook tegen eventueel resterende ziekte. Dit is met name belangrijk bij het voorkomen van een nieuw recidief bij leukemie. Om deze reden en om zo kort mogelijk zonder afweer te zijn, is het essentieel dat de juiste cellen gekozen worden en dat deze snel en gecontroleerd uitgroeien tot een nieuw en volledig functionerend immuunsysteem. Het gecontroleerd uitgroeien is met name van belang voor het voorkomen van een potentieel dodelijke complicatie: transplantatieziekte, waarbij de donor immuuncellen het lichaam van de patiënt aanvallen. Hierdoor is het dus nodig een optimale balans te vinden tussen de gewenste activiteit van het nieuwe immuunsysteem tegen resterende ziekte en de ongewenste activiteit tegen het lichaam. In het eerste deel van het proefschrift is de aandacht uitgegaan naar het vinden van deze optimale balans.
De conditionering vindt plaats tussen donorselectie en immuunherstel. Dit is het afbreken van het beenmerg en het immuunsysteem van de patiënt met hoge dosis chemotherapie soms in combinatie met bestraling. Vanwege de hoge dosis van zeer toxische chemotherapie, is een goede farmacologische kennis over de juiste werking en dosis van belang. Het beenmerg en immuunsysteem van de patiënt moet voldoende afgebroken worden, maar ernstige toxiciteit moet worden voorkomen. Ook dit vraagt een optimale balans tussen voldoende hoog doseren voor optimale effectiviteit en niet te hoog te doseren om ernstige toxiciteit te vermijden. In het tweede deel van het proefschrift ligt de nadruk op het vinden van deze balans.

B 1 Immunologie van transplantatie

Een geschikte verwante donor, met een imuunsysteem identiek aan dat van de patiënt, wordt beschouwd als de gouden standaard van donor opties voor SCT, maar is niet altijd beschikbaar. Om die reden zijn alternatieve donoren bronnen van onderzoek. Een vergelijkbare mogelijkheid is een geschikte onverwante donor. Andere bronnen zijn navelstrengbloed en (als een gewone navelstrengbloed donor onvoldoende cellen oplevert) een dubbel navelstrengbloed donor, waar voor allebei deze donoren geldt dat het imuunsysteem van de donor minder goed hoeft te passen bij dat van de patiënt ten opzichte van zowel een geschikte onverwante donor als een geschikte verwante donor.
In hoofdstuk 2 is onderzocht wat de uitkomsten waren na SCT met verschillende donor bronnen in 317 kinderen met acute myeloide leukemie. De inclusie voor de studie vond plaats tussen 2005 en 2015 in acht verschillende centra wereldwijd. De belangrijkste determinant was de SCT bron. Omdat de karakteristieken tussen groepen getransplanteerd met de verschillende donoren niet hetzelfde waren, werden bij vergelijkingen deze verschillen in patiënt-, transplantatie- en ziekte karakteristieken meegenomen.
De mediane leeftijd op moment van SCT was 10 jaar (0.4 tot 21) met een mediane opvolging van 4.74 jaar. In dit cohort was 19% getransplanteerd met een geschikte verwante donor, 23% met een geschikte onverwante donor, 39% met navelstrengbloed donor en 19% met een dubbele navelstreng donor. De uitkomst van de SCT was gelijk tussen de verschillende bronnen, wat betreft de kans op recidief van acute myeloide leukemie alsmede de kans op overlijden niet geassocicieerd met ziekte recidief (NRM). Hierdoor was ook de leukemie-vrije overleving gelijkwaardig tussen de verschillende stamcelbronnen. De algehele overleving was slechter met een dubbel navelstrengbloed donor ten opzichte van een geschikte verwante donor. Ondanks het feit dat de navelstrengbloed donor en dubbel navelstrengbloed donor qua immuunsysteem minder goed bij de patiënt pasten ten opzichte van de andere donoren, was de incidentie van chronische transplantatieziekte laag, zeker vergeleken met de geschikte onverwante donor transplantaties. Omdat de overleving gelijkwaardig was voor een navelstrengbloed donor, was de overleving vrij van chronische transplantatieziekte hoger met een navelstrengbloed donor ten opzichte van geschikte onverwante donor.
Deze uitkomsten wijzen erop dat een navelstrengbloed donor een goede keuze is als stamcelbron, wanneer een geschikte verwante donor niet beschikbaar is. Hoewel de leukemie-vrije overleving vergelijkbaar is met geschikte onverwante donor, is de kans op de ernstige complicatie chronische transplantatieziekte lager, wat mogelijk leidt tot een betere kwaliteit van leven op de

lange termijn.

Ook is het voor te stellen dat de keuze van transplantatie bron invloed heeft op immuunherstel. De vorming van het nieuwe immuunsysteem kan worden verdeeld in twee fases. In de eerste fase nestelen de stamcellen zich in het lege beenmerg en verschijnen de eerste aspecifieke afweercellen in het bloed, zoals neutrofielen, 'natural killer' of NK-cellen, en monocyten. In de tweede fase komen de adaptieve immuuncellen opzetten, zoals T- en B-cellen. Van deze cellen zijn met name de CD4+ T-cellen van belang voor de overleving, door zowel bescherming tegen NRM en recidieven. Omdat deze cellen relatief laat opkomen, is het waardevol de snelheid van dit herstel te voorspellen aan de hand van de eerder zichtbare aspecifieke cellen. Daarom is in hoofdstuk 3 een verband gezocht tussen aspecifiek en adaptief immuunherstel. Ook is het aspecifeke herstel gerelateerd aan de stamcelbron: navelstrengbloed donor of beenmerg (bestaande uit geschikte verwante donor & geschikte onverwante donor).
Uit deze studie kwam naar voren dat het aantal aspecifieke immuuncellen in het bloed na initieel herstel, de snelheid van CD4+ T-cel herstel kan voorspellen. Ook werd hier gevonden dat na een SCT met navelstrengbloed donor het aantal aspecifieke cellen hoger was dan met donor beenmerg als bron. Dit werd verklaard door aan te tonen dat navelstrengbloed donorcellen zowel voor als na (in de patiënt) transplantatie sneller delen dan beenmergcellen.
Het is dus mogelijk vroeger te voorspellen of het adaptieve herstel adequaat zal zijn met behulp van monitoren van de eerder opkomende aspecifieke cellen. Ook kan de kans op een sneller herstel van beide soorten cellen verhoogd worden met navelstrengbloed donor als bron.

B 2 Conditionering optimalizeren

In Nederland en veel andere landen, is de combinatie van fludarabine en busulfan het meest gebruikte conditioneringsschema. In hoofdstuk 4 wordt een overzicht gegeven van gebruikte schema's en de resultaten hiervan als onderbouwing voor het gebruik van fludarabine en busulfan.
Het belangrijkste alternatief voor busulfan is bestraling en zowel busulfan als bestraling worden vaak gecombineerd met cyclofosfamide. In de vroege dagen van SCT (1980-2000) was er van busulfan alleen nog een orale formulering beschikbaar. De opname in het lichaam van deze vorm was erg onvoorspelbaar, waardoor met een gelijke dosis de blootstelling tussen patiënten

en zelfs binnen een patiënt erg variabel was. Dit is problematisch omdat de therapeutische breedte zeer nauw is voor busulfan, waar een lage blootstelling de kans op zowel falen van de SCT als een recidief verhoogt en te hoge blootstelling associeert met toxiciteit en NRM. In 2000 kwam een intraveneuze formulering van busulfan op de markt die veel beter voorspelbare blootstellingen gaf. Daaropvolgend kwam ook de optie busulfan op basis van individuele bloedspiegels te doseren beschikbaar, waardoor de blootstelling nog beter gecontroleerd kon worden. Dit zorgde voor een gelijkwaardige tot betere overleving ten opzichte van bestraling, met minder lange termijn complicaties.
Fludarabine werd ingebracht als alternatief voor cyclofosfamide om de toxiciteit in combinatie met busulfan nog verder terug te dringen. Inderdaad werd in verschillende studies bewezen dat NRM lager was met busulfan-fludarabine dan met busulfan-cyclofosfamide. Omdat de kans op een recidief gelijkwaardig was, zorgde dit voor een betere algehele overleving. Ook voor fludarabine zijn aanwijzingen dat de dosering nauw komt, waardoor meer precieze methodes dan het op dit moment gebruikte doseren op basis van lichaamsoppervlakte de uitkomsten nog verder zouden kunnen verbeteren.

Om verder onderzoek te kunnen doen naar het blootstellings-gericht doseren van busulfan en fludarabine is een efficiente meetmethode, om deze middelen in bloed van patiënten te bepalen, nodig. In hoofdstuk 5 werd daarom een methode ontwikkeld waarmee beide middelen en een derde middel (clofarabine) in één bloedmonster gemeten kunnen worden.
Dit bleek echter een chromatografische uitdaging, vanwege de extreme structuur verschillen tussen busulfan aan de ene kant en fludarabine/clofarabine aan de andere kant. Ook zijn stabiele isotoop gelabelde standaarden, normaliter gebruikt voor bio-analyse, voor fludarabine en clofarabine slecht beschikbaar. Ter vervanging van een dergelijke isotoop, werd standaard additie toegepast. De methode werd gevalideerd volgens European Medicine Agency (EMA) richtlijnen. Er is dus een efficiënte methode ontwikkeld om busulfan, fludarabine en clofarabine in één monster te meten. Dit kan gebruikt worden voor zowel toekomstig onderzoek als voor klinische toepassing van bloedspiegel gestuurd doseren van alle drie de middelen.

Van busulfan is bekend dat de therapeutische breedte zeer nauw is en dat daarom doseren op basis van bloedspiegels geïndiceerd is. Een bijkomend probleem is echter dat de eliminatie van busulfan afneemt gedurende de tijd die de conditionering in beslag neemt. Zeker omdat dit effect zeer variabel is tussen patiënten, is het moeilijk een goede voorspelling te doen van de

ideale dosis, op basis van de bloedspiegels die gemeten zijn op de eerste dag. In hoofdstuk 6 is daarom een semi-mechanistisch model ontwikkeld om dit effect beter te begrijpen en mogelijk te kunnen voorspellen.
De snelheidsbeperkende stap in de afbraak van busulfan is de conjugatie aan het lichaamseigen molecuul glutathion. De hypothese die onderzocht werd in dit hoofdstuk is dat depletie van glutathion door afbraak van busulfan zorgt dat diezelfde afbraak minder snel wordt. Hiervoor is gebruik gemaakt van bloedspiegelwaardes voortkomend uit routine patiëntenzorg bij zowel kinderen als volwassenen, die gemeten zijn tijdens de conditionering ten behoeve van dosis optimalisatie. Deze data zijn gebruikt om 1) een eerder ontwikkeld model, waarin het tijds-effect empirisch verwerkt was, opnieuw op te evalueren; 2) dit model uit te breiden met theoretische glutathion depletie; 3) de twee modellen te vergelijken in hun vermogen de juiste dosis te voorspellen.
Voor beide modellen werd een afname in klaring geschat van 7% (-82% tot 44%). Het bleek mogelijk het effect van een afnemende klaring te beschrijven aan de hand van glutathion depletie. Hiermee werd het geobserveerde effect dat mensen die busulfan snel afbraken ook een grotere afname in afbraak hadden goed beschreven. Vanaf een leeftijd van 40 jaar en ouder, werd het tijds-effect sterker per levensjaar, waardoor het effect meer dan een tweevoud toenam over de leeftijdsspanne (0-73 jaar). Dit werd verklaard door een afname in glutathion voorraad bij hogere leeftijd. Het nieuwe model bleek beter in het voorspellen van de juiste dosis: van 74% naar 76% blootstelling binnen het therapeutisch venster in de simulatie-studie. Dit kwam met name door een betere voorspelling bij volwassen patiënten.
Deze studie laat zien dat de depletie van glutathion een plausibele verklaring is voor de afname in busulfan klaring, waardoor deze afname proportioneel is aan de omzetting van busulfan. Verder is een toename van dit effect zichtbaar bij een leeftijd van boven de 40 jaar. Het model kan gebruikt worden om nog nauwkeuriger busulfan te doseren op basis van bloedspiegels op de eerste dag.

Hoewel voor busulfan al veel bekend is over de relatie tussen dosis, bloedspiegels en SCT-uitkomsten, is dit voor fludarabine minder duidelijk. Een belangrijk begin hierin is gemaakt in hoofdstuk 7 door een farmacokinetisch model te ontwikkelen voor fludarabine. Het doel hiermee was het verbinden van dosis aan bloedspiegels en ook het vinden van voorspellers die deze relatie het best beschrijven.
Hiervoor is gebruik gemaakt van data van zowel volwassenen als kinderen die een SCT kregen tussen 2010 en 2016. In totaal waren er data van 258

patiënten met een leeftijd van 0.3 tot 74 jaar, van wie 2605 bloedmonsters beschikbaar waren. Hierin werd de circulerende metaboliet van fludarabine gekwantificeerd middels de methode uit hoofdstuk 5 en de verworven plasmaconcentratie-tijd data zijn gebruikt voor het ontwikkelen van het farmacokinetisch model.
Hoewel op dit moment fludarabine gedoseerd wordt op basis van lichaamsoppervlakte, bleken lichaamsgewicht en nierfunctie (als geschatte glomerulaire filtratie (GF)) een veel betere voorspeller voor de fludarabine klaring. Daarom werd de klaring opgedeeld in een niet renale (3.24±20% L/uur/70kg) en renale (GF*0.782±11% L/uur/70kg) fractie. Omdat deze componenten niet meegenomen zijn in het huidige doseerregime leidde dit tot hoge variabiliteit in de blootstelling: de oppervlakte onder de concentratie-tijd-curve met extrapolatie naar oneindig ($AUC_{t0-\infty}$) varieerde van 10 tot 66 mg*uur/L bij een dosis van 160 mg/m^2.
Concluderend leidt het huidige doseerregime voor fludarabine op basis van lichaamsoppervlakte tot grote variabiliteit in blootstelling. Een doseerregime gebaseerd op lichaamsgewicht en nierfunctie (als GF) leidt mogelijk tot minder variabiliteit in blootstelling, waardoor het mogelijk is preciezer te doseren.

De voorgenoemde variabiliteit in blootstelling is mogelijk deels de oorzaak van de sterk varierende uitkomst van SCT. Om de therapeutische breedte van fludarabine vast te stellen, is in hoofdstuk 8 de relatie tussen fludarabine bloedspiegels en SCT uitkomsten onderzocht.
Hiervoor is gebruikt gemaakt van de data uit hoofdstuk 7, maar dan alleen van patiënten die 160 mg/m^2 fludarabine kregen als onderdeel van myeloablatieve conditionering met verder: busulfan gericht op een $AUC_{t0-\infty}$ van 90 mg*uur/L en anti-thymocyte globuline (6-10mg/kg; van dag -9/-12). Vervolgens is de fludarabine blootstelling, als $AUC_{t0-\infty}$, gerelateerd aan NRM, ziekte recidief en transplantatie falen, middels parametrische tijd-tot-event (TTE) modellen. De optimale blootstelling aan fludarabine werd gezocht waarbij de negatieve effecten van een te lage blootstelling en de negatieve effecten van een te hoge blootstelling beperkt werden. Nadien werd gekeken wat de uitkomsten waren van patiënten met een $AUC_{t0-\infty}$ nabij het optimum ten opzichte van patiënten met een potentieel minder gunstige blootstelling.
Uiteindelijk waren 192 patiënten, met een leeftijd tussen de 0.3 en 74 jaar en van wie 68 getransplanteerd werden voor een goedaardige aandoening en 124 voor een maligniteit, beschikbaar voor de analyse. Een $AUC_{t0-\infty}$ van 20 mg*uur/L bleek de optimale fludarabine blootstelling, geasocieerd met 33%

kans op het krijgen van ongunstige uitkomsten. Een te lage blootstelling was geassocieerd met een circa twee keer hogere sterfte, door een vijf keer verhoogde kans op falen van de transplantatie. Een hoge $AUC_{t0-\infty}$ had een vergelijkbaar effect (twee keer verhoogd) op sterfte met te lage $AUC_{t0-\infty}$, maar dit kwam met name door algehele toxiciteit en een meer dan twee keer zo traag immuunherstel.
Uit deze resultaten blijkt dat het belangrijk is zo dicht mogelijk bij de optimale blootstelling van 20 mg*uur/L te blijven. Dit zorgt voor een succesvolle SCT, met minder toxiciteit en een sneller immuunherstel.

Nu duidelijk is dat het huidige doseerregime voor fludarabine een variabele blootstelling geeft en dat deze variabiliteit zorgt voor slechtere uitkomsten aan beide kanten van het spectrum, lijkt de volgende stap het aanpassen van de dosis gericht op de optimale blootstelling. Voor definitieve bewijslast ten gunste van nieuwe doseerregimes, zou echter gerandomizeerd onderzoek nodig zijn. Het is belangrijk een dergelijke studie goed op te zetten, om zo min mogelijk patiënten bloot te stellen aan potentieel minder gunstige uitkomsten van één van beide regimes, maar wel een definitief uitsluitsel over de optimale dosering te geven.
Om te onderzoeken wat de optimale opzet van deze studie zou moeten zijn, is in hoofdstuk 9 een nieuw raamwerk voor simulaties opgezet om een klinische studie virtueel te kunnen doorlopen vanaf dosis tot blootstelling en uiteindelijk SCT uitkomst. De relatie dosis-blootstelling werd gesimuleerd middels het farmacokinetisch model uit hoofdstuk 7 en met behulp van de TTE modellen uit hoofdstuk 8 konden uitkomsten gesimuleerd worden. Met het raamwerk werden vervolgens gerandomiseerde studies gesimuleerd die de huidige dosering van 160 mg/m^2 vergeleken met ofwel doseren gebaseerd op nierfunctie en gewicht danwel op basis van bloedspiegels gemeten op de eerste dag. Het aantal patiënten per arm werd voor beide vergelijkingen gevarieerd en het discriminerend vermogen (het percentage studies dat een definitief uitsluitsel kon geven) van de studie werd uitgerekend voor de eindpunten SCT-falen, ziekte recidief, NRM en algehele overleving. Het ontwerp met het minimaal aantal patiënten per arm om 80% discrimenerend vermogen te bereiken werd als optimaal gedefinieerd en van deze studie werden de resultaten gesimuleerd.
De optimale studie bleek te bestaan uit tweemaal 70 patiënten waarbij gedoseerd werd op basis van spiegelbepaling en waarbij sterfte als gevolg van de behandeling als eindpunt genomen werd. Voor hetzelfde onderscheidend vermogen was n=160 nodig bij dezelfde vergelijking met algehele overleving als eindpunt en n=120 voor doseren op basis van het model met NRM als

eindpunt. De verwachte absolute toename in overleving door de nieuwe doseer methode was 13% door een sterke afname in NRM (5.7% vs 21%).

B 3 Conclusies

In dit proefschrift zijn verschillende aspecten van SCT onderzocht, zowel immunologische als farmacologische.
Allereerst is gekeken naar de invloed van stamcelbron op verschillende uitkomsten als immuunherstel, transplantatieziekte en overleving, waarbij navelstrengbloed donor als effectieve stamcelbron naar voren kwam. Verder is de relatie tussen vroeg aspecifiek en laat adaptief immuunherstel gelegd.
Daarna is gezocht naar de optimale vorm van conditionering. Met behulp van reeds bekende literatuur is het busulfan-fludarabine regime vergeleken met hun respectievelijke tegenhangers bestraling en cyclofosfamide, waarbij busulfan-fludarabine positief naar voren kwam als even effectief schema met minder ernstige bijwerkingen. De busulfan dosering kan verder geoptimaliseerd worden door de afname in klaring beter te voorspellen, wanneer de metingen op dag 1 gebruikt worden voor het bepalen van de dosis. Voor fludarabine lijkt het huidige doseerregime niet adequaat en alternatieve doseerregimes zijn voorgesteld, alsmede een studie-opzet om deze regimes te onderzoeken. Een meetmethode is beschreven om dit effectief toe te passen in combinatie met het reeds gebruikte doseren op basis van bloedspiegels bij busulfan.
Samenvattend weten we nu dat de optimale conditionering voor een SCT het busulfan-fludarabine regime is en is veel duidelijk geworden over de juiste dosering van dit regime. Daarbij is uitgezocht hoe de immunologie van donorcellen, aspecifek en adaptief immuunherstel samenhangen. Op deze manier draagt therapie gebaseerd op wiskundige modellen bij aan veilige én effectieve toepassing van SCT.

Appendix C List of co-authors and affiliations

- Bierings, Marc
  Pediatric Blood and Marrow Transplant Program, Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands

- Boelens, Jaap Jan
  Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
  Pediatric Blood and Marrow Transplant Program, Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands
  Stem Cell Transplant and Cellular Therapies; Pediatrics; Memorial Sloan Kettering Cancer Centre, New York City, USA

- Boss, Jill
  Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

- Dahlberg, Ann
  Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, USA

- Delaney, Colleen
  Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, USA

- Dorlo, Thomas P.C.
  Department of Pharmacy & Pharmacology, Antoni van Leeuwenhoek Hospital/Netherlands Cancer Institute, Amsterdam, The Netherlands

- Elfeky, Reem
  Institute of Child Health, Great Ormond Street Hospital for Children, London, United Kingdom

- Egas, Annelies C.
  Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

- Egberts, Toine C.G.
  Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
  Department of Pharmacoepidemiology & Clinical Pharmacology, Faculty of Science, Utrecht University, Utrecht, The Netherlands

◇ Giller, Roger
Blood and Marrow Transplantation, Department of Pediatrics, University of Colorado School of Medicine, Aurora, USA

◇ Huitema, Alwin D.R
Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Department of Pharmacy & Pharmacology, Antoni van Leeuwenhoek Hospital/Netherlands Cancer Institute, Amsterdam, The Netherlands

◇ Keating, Amy K.
Blood and Marrow Transplantation, Department of Pediatrics, University of Colorado School of Medicine, Aurora, USA

◇ van Kesteren, Charlotte
Pediatric Blood and Marrow Transplant Program, Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands
Department of Pharmacy, Albert Schweitzer Hospital, Dordrecht, The Netherlands

◇ de Koning, Coco C.H.
Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Pediatric Blood and Marrow Transplant Program, Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands

◇ Kuball, Jürgen
Department of Hematology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

◇ Kurtzberg, Joanne
Pediatric Blood and Marrow Transplant Division, Duke University Medical Center, Durham, USA

◇ Lalmohamed, Arief
Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Department of Pharmacoepidemiology & Clinical Pharmacology, Faculty of Science, Utrecht University, Utrecht, The Netherlands

◇ Lindemans, Caroline A.
Pediatric Blood and Marrow Transplant Program, Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands

◇ van Maarseveen, Erik M.
Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

◇ Milano, Filippo
Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, USA

◇ Mitchel, Richard
Department of Blood and Marrow Transplant, Royal Manchester Children's Hospital, Manchester, United Kingdom

◇ Nierkens, Stefan
Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Clinical immunologist, Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands

◇ O'Brien, Tracey A.
Blood and Marrow Transplant Program, Kids Cancer Centre, Sydney Children's Hospital, Randwick, Australia; School of Women & Children's Health, University of New South Wales, Sydney, Australia

◇ Page, Kristin M.
Pediatric Blood and Marrow Transplant Division, Duke University Medical Center, Durham, USA

◇ Punt, Arjen M.
Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

◇ Raijmakers, Reinier
Department of Hematology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

◇ van Rhenen, Anna
Department of Hematology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

◇ Stefanski, Heather
Blood and Marrow Transplantation, Department of Pediatrics, University of Minnesota, Minneapolis, USA

◇ Verneris, Michael R.
Blood and Marrow Transplantation, Department of Pediatrics, University of Colorado School of Medicine, Aurora, USA

◇ Veys, Paul
Institute of Child Health, Great Ormond Street Hospital for Children, London, United Kingdom

◇ Wagner John E.
Blood and Marrow Transplantation, Department of Pediatrics, University of Minnesota, Minneapolis, USA

◇ de Witte, Moniek A
Department of Hematology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Laboratory of Translational Immunology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

◇ Wynn, Robert F.
Department of Blood and Marrow Transplant, Royal Manchester Children's Hospital, Manchester, United Kingdom

Appendix D List of publications related to this thesis

Punt, A. M., Langenhorst, J. B., Egas, A. C., Boelens, J. J., van Kesteren, C. & van Maarseveen, E. M. Simultaneous quantification of busulfan, clofarabine and F-ara-A using isotope labelled standards and standard addition in plasma by LC-MS/MS for exposure monitoring in hematopoietic cell transplantation conditioning. *J Chromatogr B Analyt Technol Biomed Life Sci* **1055-1056,** 81–85 (2017).

De Koning, C., Langenhorst, J., van Kesteren, C., Lindemans, C. A., Huitema, A. D., Nierkens, S. & Boelens, J. J. Innate Immune Recovery Predicts CD4+ T Cell Reconstitution after Hematopoietic Cell Transplantation. *Biology of Blood and Marrow Transplantation* **25,** 819–826 (2019).

Langenhorst, J. B., Dorlo, T. P. C., van Maarseveen, E. M., Nierkens, S., Kuball, J., Boelens, J. J., van Kesteren, C. & Huitema, A. D. R. Population Pharmacokinetics of Fludarabine in Children and Adults during Conditioning Prior to Allogeneic Hematopoietic Cell Transplantation. *Clinical pharmacokinetics.* Epub ahead of print (Oct. 2018).

Langenhorst, J. B., Keating, A. K., Wagner, J. E., Page, K. M., Veys, P., Wynn, R. F., Stefanski, H., Elfeky, R., Giller, R., Mitchel, R., Milano, F., O'Brien, T. A., Dahlberg, A., Delaney, C., Kurtzberg, J., Verneris, M. R. & Boelens, J. J. The Influence of Stem Cell Source on Transplant Outcomes for Pediatric Patients with Acute Myeloid Leukemia. *Blood advances.* Accepted (Feb. 2019).

Langenhorst, J. B., van Kesteren, C., Huitema, A. D. R. & Boelens, J. J. *Busulfan and fludarabine in conditioning:a potent duo adjustable to precision?* Submitted. Jan. 2019.

Langenhorst, J. B., Boss, J., van Kesteren, C., Lalmohammed, A., Kuball, J., Egberts, T. C. G., Boelens, J. J., Huitema, A. D. R. & van Maarseveen, E. M. *A semi-mechanistic model to describe the reduction in busulfan clearance by depletion of glutathione* In preparation. Jan. 2019.

Langenhorst, J. B., van Kesteren, C., van Maarseveen, E. M., Dorlo, T. P. C., Nierkens, S., Lindemans, C. A., de Witte, M. A., van Rhenen, A., Raij-

makers, R., Bierings, M., Kuball, J., Huitema, A. D. R. & Boelens, J. J. *Fludarabine exposure in the conditioning prior to allogeneic hematopoietic cell transplantation predicts outcomes* Submitted. Dec. 2018.

Langenhorst, J. B., Dorlo, T. P. C., van Kesteren, C., van Maarseveen, E. M., Nierkens, S., de Witte, M., Boelens, J. J. & Huitema, A. D. R. *A novel simulation framework for rational design of trials evaluating optimal fludarabine dosing in allogeneic hematopoietic cell transplantation* In preparation. Jan. 2019.

Appendix E List of Communications related to this thesis

Langenhorst, J., van Kesteren, C., van Maarseveen, E., Kuball, J., de Witte, M. & Boelens, J.-J. *High Exposure to Fludarabine in Conditioning Prior to Allogeneic Hematopoietic Cell Transplantation Predicts for Impaired CD4 Reconstitution and Lower Survival Chances* Oral: ASBMT Tandem meetings 2017.

Langenhorst, J. B., van Kesteren, C., Dorlo, T. P., van Maarseveen, E. M., Nierkens, S., Kuball, J., Witte, M. A. D., Boelens, J. J. & Huitema, A. D. R. *High exposure to fludarabine in conditioning prior to allogeneic hematopoietic cell transplantation predicts impaired CD4 reconstitution and lower probability of survival* Oral: Population Approach Group Europe meeting 2017.

Langenhorst, J. B., Dorlo, T. P. C., van Kesteren, C., van Maarseveen, E. M., Nierkens, S., Witte, M. A. D., Boelens, J. J. & Huitema, A. D. R. *Cause-specific hazard models with markovian elements to quantify the fludarabine exposure-response relationship: from learning to confirming in allogeneic hematopoietic cell transplantation* Oral (Lewis Sheiner Student Session): Population Approach Group Europe meeting 2018.

Appendix F Other publications

De Koning, C., Gabelich, J. A., Langenhorst, J., Admiraal, R., Kuball, J., Boelens, J. J. & Nierkens, S. Filgrastim enhances T-cell clearance by antithymocyte globulin exposure after unrelated cord blood transplantation. *Blood Adv* **2,** 565–574 (2018).

Appendix G Curriculum vitae

Jurgen Langenhorst was born on the 30th of April 1988 in Hoogeveen. He graduated from the Gymnasium Camphusianum in 2005, after which he started his bachelor and subsequently his master of Pharmacy at Utrecht University. During this period he investigated the potential role of erythropoietine in the treatment of myocardial infarction and heart failure, performed research on Gallium68 and its role in nuclear medicine, and on poly(n-vinylcaprolactam) aiming to be used as a thermosensitive polymer for controlled chemotherapy delivery.

After obtaining his pharmacist degree early 2015, he started working as a pharmacist in the manufacturing department at the pharmacy of the University Medical Center Utrecht. In December 2015, he started his PhD, to which this thesis is related, in the same center at the group of Boelens-Nierkens under supervision of Prof. Dr. Huitema, Dr. Boelens, and Dr. van Kesteren. For work related to this thesis, he received the Lewis Sheiner award at the Population Approach Group Europe (PAGE) meeting in 2018.

Currently he works as a pharmacometric consultant at Pharmetheus AB, Uppsala. He lives together with his partner Kayleigh Polman in Utrecht. They have shared a passion for rowing during their studies and are currently equally passionate about triathlon.

Appendix H Dankwoord

Het wordt wel gezegd dat promoveren een eenzame bezigheid is, dus het lijkt erop dat ik een zeer gunstige omgeving gehad heb. Zelden heb ik het gevoel gehad dat enkel ikzelf belang had bij het tot stand komen van dit werk, waarvoor ik een aantal mensen in het bijzonder wil bedanken.

Allereerst mijn promotor, Alwin. Er moest een kleine horde genomen worden, namelijk je aanstelling als professor, alvorens je officieel promotor werd, maar dit uitstel heeft je verantwoordelijkheidsgevoel niet minder groot gemaakt. Structuur aanbrengen in mijn denken, werk en planning was je grootste uitdaging, maar ik denk dat de afronding van dit proefschrift bewijs is dat je hier behoorlijk in geslaagd bent. Ook kijk ik met veel plezier terug naar de keren dat je nieuwe ideëen uittekende op een kladblaadje, die je zo uit je mouw leek te schudden. Als toonbeeld van je betrokkenheid, ben je tot en met de laatste stukken altijd kritisch gebleven en waarschijnlijk had je het liefst in dit stuk ook nog verbeteringen aangebracht.

Mijn promotieteam zou nooit compleet geweest zijn zonder Jaap Jan en Charlotte. Jaap Jan, jij bent vanaf het begin af sturend en enthousiast geweest wat betreft de resultaten. Soms mailde ik in mijn enthousiasme wat preliminaire 'curves' naar je, waarop je vaak (bijna standaard) de reactie 'Cool!' gaf. Ik wist echter pas zeker dat ik potentieel goud had als je tijdens een bespreking de opmerking 'volgende week klaar?' maakte. Ook je klinische manier van denken en je vooruitstrevendheid is iets wat me nog jaren zal helpen.

De methode is iets waar Jaap Jan zich in mijn geval weinig mee bemoeide. Gelukkig was dit goed mogelijk met Charlotte als co-promotor. In het begin heb je me geholpen aan mijn eerste werkende NONMEM run, waar je me later behoedde voor onzorgvuldigheid door kritisch te blijven als ik wat te snel wilde gaan. Ook kon je mijn ongestructureerde gedachtesprongen heel scherp opschrijven met weinig woorden. Vaak bij jouw revisies heb ik gedacht, 'ja precies, dat bedoelde ik!'. Als kamergenoot van Alwin tijdens je promotie en door jarenlange samenwerking met Jaap Jan en zelfs oud-collega van mijn moeder, was je de perfecte lijm van dit team.

Hoewel inhoudelijk de BoNi groep vaak een ver van mijn bed show was, heb ik me er vanaf het begin thuis gevoeld. De onderzoeksrichtingen lopen misschien uiteen, maar de passie ervoor is iets wat we allemaal delen en

kunnen waarderen in elkaar. Dit fanatisme heeft ons helaas niet geholpen in de LTI pubquiz, al is mijn helft wel zeer passievol laatste geworden.

De drijfveer achter de hoeveelheid energie in de groep is zeker niet alleen Jaap Jan, maar ook de andere helft van BoNi: Stefan. Ik kan het natuurlijk hebben over mijn sporadische biologie vragen aan je, waarbij ik hartstochtelijk probeerde eventueel labwerk te ontlopen of uit besteden. De beste herinnering aan je is eerder je inspirerende manier van presenteren en natuurlijk de kudo's op Strava.

Als je het lab vermijdt, breng je nogal veel tijd door in je kamer. Gelukkig is dit geen vervelende plek geweest om te zijn met zulke medebewoners. Anke, Celina, Coco, Niek en Simone, ik heb het erg gewaardeerd dat jullie wilden luisteren naar mijn gekke oraties, waarbij ik vaak later bedacht dat mijn theorieën misschien iets minder steekhoudend waren. Hopelijk hebben jullie hierbij nog een klein percentage van mij geleerd van wat ik van jullie heb meegekregen. Ik weet zeker dat mijn huidige kamergenoot jullie niet kan overtreffen.

Ook wil ik natuurlijk mijn andere collega's van het LTI bedanken voor jullie tolerantie jegens mij en in het bijzonder de mannen van de lunch. Marco's favoriete film ben ik vreemd genoeg alweer vergeten, maar ik ga hem zeker kijken. Door Ruud zal ik voortaan heel goed uitkijken voor vissers, als ik nog eens in een roeiboot zit. Koen, ik zal je niet uit het wiel fietsen, mocht ik je een keer achter me zien schuilen met harde wind tegen.

Een substantieel deel van dit proefschrift gaat niet over kinderen, waar Jaap Jan de expert is, maar over volwassenen. Hierbij heb ik altijd veel hulp gehad met de specifieke klinische vragen die hierbij komen kijken van Jürgen en Moniek. Ook kan ik dankzij jullie iets fysieks achterlaten via de klinische studie voor fludarabine, die jullie met veel moeite hebben opgezet.

Buiten het LTI, ben ik via mijn promotor en vorige werkgever nog met twee andere afdelingen verbonden geweest, de apotheken van het UMC en het NKI. Toine, in afwachting van de aanstelling van Alwin ben ik prima door je opgevangen. Ook in de loop van het traject heeft het me veel geholpen om bij jou en de rest van de apotheek onderzoeksgroep betrokken te blijven.

Binnen de UMC apotheek is er nog een belangrijke drijvende kracht achter

mijn promotie geweest: Erik. Natuurlijk is je bijdrage als verantwoordelijke voor het meten van alle fludarabine monsters al groot, maar dit was niet het enige. Dit hele traject begon namelijk met een opmerking van jou in de wandelgangen van de apotheek: ïk weet een leuke promotieplek voor je". Daarna en nog steeds is je aanstekelijke enthousiasme een belangrijke drijfveer geweest.

Vanaf het moment dat ik dat modelleren een beetje dacht te begrijpen, mocht ik elke dinsdag de trein pakken richting Amsterdam Zuid om aan te sluiten bij het NKI-NONMEM-overleg. Natuurlijk hebben jullie me daar inhoudelijk veel bijgebracht, maar het is vooral prettig geweest met mede vakidioten te kunnen praten over moeilijke (en minder moeilijke) modellen. Een deel van jullie zal ik zeker nog vaker terugzien in onze kleine nerdgemeenschap. In het bijzonder wil ik Thomas bedanken voor al zijn hulp bij de meer technische kanten van de modellen, maar zeker ook voor je tolerantie bij je zorgvuldige revisies van mijn stuken.

Zonder mijn ouders was ik er natuurlijk niet geweest, maar was ik zeker niet hier gekomen. Door jullie beroep ging de medische terminologie al vroeg over tafel. Toch besloot ik op het laatste moment jullie niet achterna te gaan en farmacie te gaan doen, maar er is geen moment van weerstand of twijfel van jullie kant geweest. Voor mijn moeder, dit is de eerste opleiding, waar je me niet meer inhoudelijk met mijn huiswerk hebt geholpen, maar de steun is er zeker niet minder om geweest. Nog altijd ben je een voorbeeld door je energie, die je ondanks je weinige slapen, overal in steekt, met als kenmerkende leus: 'Je bent pas ziek als je collega's (of patiënten red.) meer last van je hebben, dan je toevoegt'. Voor mijn vader, van jou heb ik geleerd dat je hetgeen moet doen dat je leuk vindt, dan komt de grenzeloze inzet vanzelf. Dit leidt uiteindelijk tot een positieve vicieuze cirkel, zoals jouw topsportarts carrière is geweest. Ook een vleug zelfvertrouwen, die soms best het realistische mag overstijgen, kan meer goed dan kwaad doen.

Een kleine ode aan mijn laatste oma is ook wel gepast. U heeft sinds de datum bekend is, elke keer als ik u zag, (waarschijnlijk nog vaker) gezegd 'als ik het maar haal'. Hoewel ik er het volste vertrouwen in heb dat het gaat lukken, bent u op deze manier in ieder geval vereeuwigd in dit boek. Heeft u weer wat leuks te vertellen aan uw vriendinnen.

Ter afsluiting natuurlijk Kayleigh, de promotor van het project 'mijn leven'. Als ik jou nooit had ontmoet, was ik waarschijnlijk nog steeds mijn

master scriptie aan het schrijven, of beter gezegd, hem niet aan het inleveren. Voor dit werk had ik een harde deadline om me hieraan te herinneren, dus kon je me op andere vlakken bijstaan. De wiskunde is misschien niet jouw favoriet, maar toen ik begon over een 'Lewis Sheiner award', zei je zonder twijfelen 'aha, dat is die van Beal'. Dit illustreert hoe betrokken je bent geweest en in hoeverre je luistert naar mijn niet altijd even interessante verhalen. Ik ben bang dat Sheiner, Beal en hun nazaten plus verhalen nog wat langer langs zullen gaan komen. Mijn excuses alvast hiervoor en bedankt voor alles, zonder jou was dit nooit gelukt.

Velen anderen hebben in een vorm bijgedragen aan het tot stand komen van dit proefschrift, ook al zijn zij hier niet genoemd. Voor jullie en iedereen, sluit ik af met een sommetje. Geen zorgen ik ben snel klaar, $jullie + 1 = \infty$.

Appendix I List of abbreviations

ACN	acetonitrile.
aGvHD	acute graft-versus-host disease.
ALL	acute lymphocytic leukemia.
AML	acute myeloid leukemia.
ANOVA	one-way analysis of variance.
AUC	area under the plasma-concentration-time curve.
$AUC_{t_0-\infty}$	area under the plasma-concentration-time curve from the first dose until infinity.
$AUC_{t_0-t_x}$	area under the plasma-concentration-time curve from the first dose until time of transplantation.
$AUC_{t_x-\infty}$	area under the plasma-concentration-time curve from time of transplantation until infinity.
BM	bone marrow.
BMI	body-mass index.
BMT	bone marrow transplantation.
BSA	body-surface-area.
Bu	busulfan.
BuCy	the busulfan-cyclophosphamide regimen.
BuFlu	the busulfan-fludarabine regimen.
BW	actual body weight.
CB	cord blood.
CBT	cord blood transplantation.
CC	calibration curves.
cGvHD	chronic graft-versus-host disease.
CI	confidence interval.
Clo	clofarabine.
CML	chronic myeloid leukemia.
CNS	history of central nervous system disease.
CR	complete remission.
CsA	cyclosporine A.
CSH	cause-specific hazard.
CTS	clinical trial simulation.

CV	coefficient of variation.
CWRES	conditional weighted residuals.
Cy	cyclophosphamide.
dCB	double cord blood.
EFS	event-free survival.
eGFR	estimated glomerular filtration rate.
EMA	European Medicine Agency.
FDA	Food and Drug Administration.
FFM	fat-free mass.
Flu	fludarabine.
FluP	fludarabine as the prodrug F-ara-AMP.
FluTP	fludarabine as the intracellular active moiety F-ara-ATP.
G-CSF	granulocyte colony-stimulating factor.
GRFS	GvHD-free/relapse-free survival.
GSH	glutathione.
GST	glutathione-S-transferase.
GvHD	graft-versus-host disease.
GVL	graft-versus-leukemia effect.
HCT	allogeneic hematopoietic cell transplantation.
HLA	human leukocyte antigen.
HR	hazard-ratio.
IIV	inter-individual variability.
IMPMAP	Monte Carlo importance sampling estimation maximization assisted by mode a posteriori estimation.
IOV	inter-occasion variability.
IR	CD4+ T-cell reconstitution.
IV	intravenous.
LCMS	liquid chromatography-tandem mass spectrometry.
LFS	leukemia free survival.
LLOQ	lower limit of quantification.

MDS	myelodysplastic syndrome.
MOF	multi-organ-failure.
MRD	minimal residual disease.
MSD	matched sibling donor.
MUD	matched unrelated donor.
NK	natural killer.
NPDE	normalized prediction distribution error.
NRM	non-relapse mortality.
OFV	objective function value.
ΔOFV	objective function value change.
OS	overall survival.
PAGE	Population Approach Group Europe.
PD	pharmacodynamic.
PK	pharmacokinetic.
PKPD	pharmacokinetic-pharmacodynamic.
Q	inter-compartmental clearance between V_1 and V_2.
Q_2	intercompartmental clearance between V_1 and V_2.
Q_3	intercompartmental clearance between V_1 and V_3.
QC	quality controls.
rATG	rabbit anti-thymocyte globulin.
RCT	randomized controlled trial.
RSE	relative standard error.
SAEM	stochastic approximation and estimation maximization.
SILS	stable isotopically labelled standards.
SIR	sampling importance resampling.
SOS	sinusoidal obstruction syndrome.
SRM	selected reaction monitoring.
TBI	total body irradiation.

TDM	therapeutic drug monitoring.
TTE	time-to-event.
V_1	volume of distribution of the central compartment.
V_2	volume of distribution of the peripheral compartment.
V_3	volume of distribution of the second peripheral compartment.
VPC	visual predictive check.

Appendix J Lijst met afkortingen

$AUC_{t0-\infty}$	oppervlakte onder de concentratie-tijd-curve met extrapolatie naar oneindig.
GF	geschatte glomerulaire filtratie.
NRM	overlijden niet geassocicieerd met ziekte recidief.
SCT	allogene stamceltransplantatie.
TTE	tijd-tot-event.