

Secondline treatment for acute graftversushost disease with mesenchymal stromal cells: A decision model

FW Thielen, HM Blommestein, LEM Oosten, FG Calkoen, AC Lankester, JJ Zwaginga K Le Blanc, A Redondo, F SánchezGuijo, M Algeri, F Locatelli, WE Fibbe, CA Uyl-de Groot

European Journal of Haematology. 2018;101(5):676-683

ABSTRACT

Objective: No standard second-line treatment exists for acute graft-versus-host disease steroid-refractory (SR-aGvHD), and long-term outcomes remain poor. Mesenchymal stromal cells (MSCs) have been evaluated as treatment, but no disease model (DM) exists that integrates and extrapolates currently available evidence. The aim of this study was to develop such a DM to describe the natural history of SRaGvHD and to predict long-term outcomes.

Method: The DM was developed in collaboration with experts in haematology-oncology. Subsequently, a model simulation was run. Input parameters for transition and survival estimates were informed by published data of clinical trials on MSC treatment for SR-aGvHD. Parametric distributions were used to estimate long-term survival rates after MSCs.

Results: The newly developed DM is a cohort model that consists of eight health states. For the model simulation, we obtained data on 327 patients from 14 published phase II trials. Due to limited evidence, DM structure was simplified and several assumptions had to be made. Median overall survival was 3.2 years for complete response and 0.5 years for no complete response. **Conclusion:** The DM provides a comprehensive overview on the second-line treatment pathway for aGvHD and enables long-term predictions that can be used to perform a cost-effectiveness analysis comparing any treatment for SR-aGvHD.

INTRODUCTION

Despite decades of research, acute graft-versus-host disease (aGvHD) is still one of the leading causes of death after haematopoietic allogeneic stem cell transplantation (HSCT) for both paediatric and adult patients.^{187,188} Immunosuppressive therapy with systemic corticosteroids is the first-line treatment option. However, latest available studies estimate that about 50%-70% of the patients do not respond to this therapy.¹⁸⁹⁻¹⁹¹ Currently, there is no standard second-line treatment available and outcomes in terms of morbidity and mortality remain poor.¹⁹¹⁻¹⁹⁵

Since 2004, mesenchymal stromal cells (MSCs) are increasingly studied in phase II clinical trials as a therapeutic option, demonstrating positive treatment effects for steroid-refractory (SR)-aGvHD.^{196,197} However, most of these trials are single-arm studies or case studies or include a limited amount of patients (<50 participants), making reliable and meaningful conclusions challenging. Ideally, the effectiveness of MSC should be tested through phase III randomised controlled trials (RCTs), but these trials are difficult to perform due to regulatory and patient population-related issues.¹⁹⁷⁻¹⁹⁹ In 2010, results of one RCT (Prochymal; Osiris Therapeutics, Inc. Columbia, Maryland, United States) were released in the form of an unreviewed abstract.²⁰⁰ Further, a multicentre RCT is currently being conducted by the European Union Horizon 2020-funded research consortium RETHRIM (REgeneration THRough IMmunomodulation)²⁰¹ (i.e. the HOVON 113 MSC trial²⁰²). The latter study aims to determine the efficacy and cost-effectiveness of MSC as part of a second-line treatment for aGvHD.

In the absence of reliable and conclusive RCT data, several options exist to aggregate and synthesise currently available evidence of MSCs as a treatment option for SR-aGvHD. These include metaanalyses, observational databases, the aggregation of expert opinion or decision analytic modelling.²⁰³ Thus far, four reviews that include currently available trials testing MSC as treatment for SR-aGvHD have been published,^{197-199,204} of which two are meta-analyses.^{197,204} However, none of the reviews combine available patient-level data (PLD) of the phase II trials to predict and model long-term outcomes of MSC treatment.

The aim of this study was to develop a disease model (DM) to describe the natural history of SR-aGvHD progression and its treatment pathways. The DM can be used to predict long-term outcomes and cost-effectiveness of current (eg, MSC treatment) and future treatment options for SR-aGvHD. To test the practicability of the model, we aimed to gather and implement PLD of a second-line treatment option. Ultimately, our model may facilitate clinical decision-making under conditions of uncertainty.²⁰⁵⁻²⁰⁷ When costs are added, this model can be a valuable tool for reimbursement decision-makers.¹⁹

METHODS

Part 1: designing and structuring the model

Model characteristics

The aim of the DM was 2-fold. First, it needed to represent the natural history of SR-aGvHD and its treatment pathways in a simplified manner. Therefore, the DM focusses on the SR-aGvHD only until patients either progress, relapse, develop chronic GvHD or die. Second, it needed to be easily adaptable to a cost-effectiveness model at a later stage. Therefore, the DM was built based on clinical expertise, previously published literature^{189,191,196,208,209} and the R ETHRIM protocol.²⁰¹ According to the ISPOR recommendations for good modelling practice,²¹⁰ we consulted clinical experts to ascertain that the model represents the disease process and addresses the decision problem of determining which one second-line treatment option is (cost-)effective when compared to another therapy option. We employed a convenience sample to include the clinical experts from the R ETHRIM consortium. To be part of R ETHRIM, consortium members needed to prove extensive research and treatment experience in the field of HSCT and work at a HSCT specialised treatment centre. All experts involved in this research thus have various professional backgrounds (e.g., internal medicine, haematology, oncology or transfusion medicine) and originate from five EU member states (Germany, Italy, The Netherlands, Spain and Sweden). A European perspective on the disease and treatment pathway(s) of SR-aGvHD was hence ensured.

Involvement of experts

Experts were consulted via email, during several telephone conferences and consortium meetings. This process was iterative until a final model version was regarded to fully represent the natural disease and treatment pathway. In addition to the several consultations, the experts were asked to give written feedback on an earlier model version by means of a semistructured questionnaire. Choosing the appropriate model type (eg, decision tree, Markov process) was also based on the ISPOR recommendations for good modelling practice.²¹⁰

Part 2: model simulation

Model input parameters

As MSCs are widely studied in numerous phase II clinical trials and case studies since 2004,^{196,199} we selected this treatment option to perform a model simulation. For the model input parameters, we identified relevant studies testing MSCs for the treatment of SR-aGvHD from the recent reviews of Chen et al.,²⁰⁴ Munneke et al.,¹⁹⁸ and Hashmi et al.¹⁹⁷ Additional PLD were obtained for studies whose principal investigator was a member of the R ETHRIM consortium.

Data were extracted using a prespecified extraction form aiming at capturing all available data that describe the disease and treatment pathway (see online Appendix 4A). The clinical effectiveness of MSC treatment was obtained from PLD of the reported first response evaluation or reconstructed from the published Kaplan-Meier curves. Studies not reporting these data were excluded.

To extrapolate the survival data beyond the observed time horizon, we followed the 2013 updated NICE Decision Support Unit recommendations.²¹¹ Accordingly, considered parametric survival models for the DM included exponential, Weibull, Gompertz, loglogistic, lognormal and generalised gamma distributions. Parametric models were evaluated through visual inspection, Akaike's information criterion (AIC) and Bayesian information criterion (BIC) tests as well as clinical validity according to the expert group.

Simulation

We ran a base-case simulation with a hypothetical patient cohort comprising 100 patients with aGvHD grades II-IV. Survival was modelled on a lifetime horizon, whereas it was assumed that patients do not exceed the age of 99 years.²¹² Health outcomes of the simulation were expressed in life-years (LYs). As future health effects are valued lower than immediate effects,¹⁹ we adjusted future health outcomes (LYs) to "present values" according to the Dutch guideline for economic evaluations in health care.^{213,214}

All statistical analyses were conducted in RStudio (version 1.0.143, R version 3.4.1). Comprehensive R Archive Network (CRAN) packages used included survival, flexsurv, survminer and plyr.

RESULTS

Part 1: designing and structuring the model

Model characteristics

We opted for a cohort-based Markov model, which represents the most relevant responses and outcomes to second-line treatment for aGvHD in corresponding health states. A Markov model consists of mutually exclusive health states that are associated with different outcomes and costs and provides an efficient structure to simulate a group of patients over time. Patients can change from one health state to another (ie, transit) at specified time interval (ie, cycles). Outcomes and costs are calculated for the entire time horizon of the model, taking into account the distribution of patients amongst the states at each cycle.^{215,216} Outcomes may entail clinical effectiveness outcomes (eg, response to treatment) as well as health-related quality of life (HRQoL) measures.

The DM comprises eight health states: (a) treatment for SRaGvHD II-IV, (b) response to treatment (complete or partial response), (c) sustained response, (d) treatment failure (stable or progressive disease), (e) relapse/persistent aGvHD, (f) third-line aGvHD therapy, (g) relapse or adverse events of haematologic disease requiring reinitiation or intensification of immunosuppression and (h) death (see Figure 1 for the model diagram). These states were primarily based on the HOVON113 MSC treatment protocol and expert opinion (see section “Involvement of experts”). Four response categories are defined: complete response (CR) is defined as the absence of all sign and symptoms of aGvHD; partial response (PR) is the improvement of aGvHD by at least one grade; stable disease (SD) is no change in aGvHD; and progressive disease (PD) is the worsening of aGvHD by at least one grade. The cycle duration was set at 28 days according to the recommendation by Martin and colleagues on the standardised time period to evaluate aGvHD response.²¹⁷ In addition, we applied a half-cycle correction for the calculation of the model outcomes.

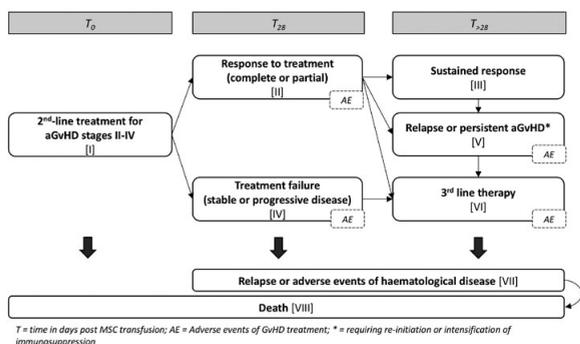


Figure 1 – Design of the decision model for MSC treatment in aGvHD

The model starts with treatment for SR-aGvHD (a) at time 0 (T_0). Within 28 days after treatment (T_{28}), patients either respond [ie, transit to the response to treatment health state (b)] or have a treatment failure (d) or die (h). Responders (b) can have a sustained response (c), relapse or have a persistent aGvHD grade that requires the reinitiation or intensification of immunosuppression (e), or enter third-line therapy (f).

Patients with a treatment failure to the initial MSC treatment (d) will directly receive third-line treatment (f). Adverse events (AEs) of treatment are not defined as a separate health state, but are possible within health states (b), and (d)-(f). At any time and from any health state, patients can transit to relapse or experience of an adverse event of the underlying haematological disease (g), or death (h). Death is an absorbing health state, meaning that once entered, patients remain in this health state.

In the model, all health states are defined as mutually exclusive although in clinical practice patients may sometimes fit the criteria of multiple health states. In these cases, patients are assigned to the health state with the largest impact on outcomes and costs. For example,

patients with treatment failure [health state (d)] may experience a relapse of the disease [health state (g)]. These patients are assigned to the relapse health state (g) although they also fit the criteria for treatment failure. The increasing numbering of the model states indicates the expected increasing impact on both outcomes and costs, according to expert opinion.

Involvement of experts

The team of clinical experts consisted of eleven members from five different countries (Germany, Italy, the Netherlands, Spain and Sweden). A preliminary model based on the HOVON113 MSC treatment protocol was presented to all consortium members at the third RETHRIM consortium meeting. Because the structure of the model was deemed incomplete, alternative treatment pathways and additional health states were discussed in detail. This led to a refinement of the existing model states and to the addition of the following health states: relapse or adverse events of haematological disease, and adverse events. The new version of the model was then presented and discussed at the fourth RETHRIM consortium meeting. To allow for written feedback, experts were asked to fill out a semistructured questionnaire at the fifth consortium meeting. Subsequently, the model was amended and a final version was approved by all experts.

Part 2: Model simulation

Model input parameters

We detected 18 studies that reported on phase II trials and collection of case studies,^{190,218–234} of which 9 reported PLD (see online Appendix 4B).^{218–223,225,228,233} Unpublished^{190,227} and additional data²³⁰ were requested and received for three studies. Two studies could be integrated by reconstructing the published Kaplan-Meier curves.^{232,234} Three studies were excluded because they did not publish survival data,^{224,231} or the proportion of patients in the response categories.⁴¹ Patients reported in the study by Ringdén et al²²⁹ were not included as they had been already included in the data presented by Le Blanc and colleagues.¹⁹⁰ From Lucchini et al,²³³ only four patients with aGvHD stages II-IV were included.

In total, we extracted data from 327 patients from 14 studies to estimate the proportion of individuals in the different health states and the proportion of patients changing between these states (ie, transition probabilities).³¹ Age and sex were only reported for 177 and 152 patients, respectively, which made the originally planned subgroup analysis for age and sex impossible. All patients had aGvHD grades II–IV prior to MSC treatment; for 204 cases, the exact grade was known (18.6% aGvHD II, 45.6% aGvHD III and 35.8% aGvHD IV). Response categories in the underlying studies were defined heterogeneously. Whereas one study did not provide any definition for the response categories employed,³⁷ only complete response (CR =the complete resolution of all aGvHD symptoms) was unanimously defined

in all other studies. Therefore, all response categories other than CR were grouped as “no complete response” (nCR).

Due to the lack of sufficiently detailed observational data, we were forced to use a simplified version of our model simulation. In this version, we could only model the first response to treatment (at day +28) and long-term survival of patients after the first response evaluation to MSCs. The health states sustained response,¹⁹⁰ relapse or persistent aGvHD,¹⁹¹ third-line therapy, chronic GvHD,¹⁹³ relapse or adverse events of haematological disease,¹⁹⁴ and adverse events in general could not be modelled.

Transition probabilities for the first two model cycles

Reported mean and median time of first evaluation was 26.6 and 28 days, respectively (range = 2-58). To integrate all available observations into the model, we assumed that all responses were evaluated within the first 28 days after MSC treatment. Based on this, we calculated a transition matrix showing the probability of response to MSC after day 28 (see Table 1).

Table 1 shows that 43.4% patients with aGvHD II-IV had a CR at first response evaluation whereas 43.7% had nCR, and 12.8% died within the first 28 days.

Table 1 - Transition rates from pre- to post-MSC treatment

		To		
		CR	nCR	Death
From	aGvHD II-IV	43.4%	43.7%	12.8%

Long-term survival estimation

To extrapolate survival estimates beyond study observations, survival data, available for 235 patients (115 CR, 120 nCR), were used. For the excluded cases, the last time of follow-up or time to death was not reported. Patients who died before day 28 were already included in the 28-day transition rates to death. Median survival times were reached at approximately five years (1819 days) for complete responders and at approximately four months (115 days) for nCR.

Kaplan-Meier curves and fitted parametric models for survival are depicted in Figure 2 for both response categories. AIC and BIC values are presented in online Appendix 4C. From a statistical point of view, the fit of the parametric models to the empirical data was similar. Nevertheless, the extrapolation after observation was different. For instance, the extrapolation according to the generalised gamma function predicted that 40% of the CR patients survived more than 25 years after MSC treatment. This was not deemed plausible from a clinical perspective by the experts. Based on clinical expertise, the lognormal distribution would present the best balance between the statistical tests of fit and visual inspection for both CR and nCR.

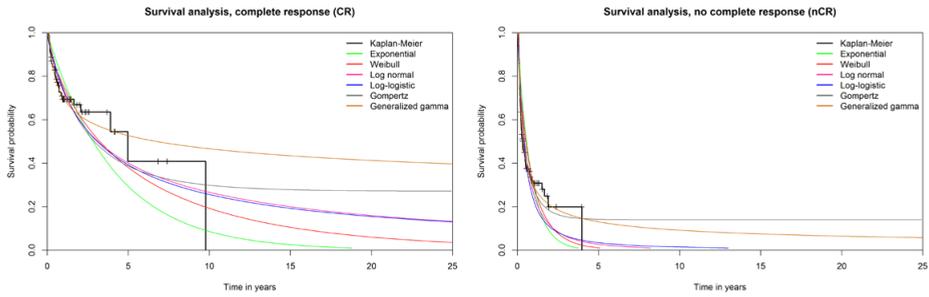


Figure 2 - Empirical Kaplan-Meier curves with parametric model estimations

Results of the model simulation

At the 28-day treatment evaluation, of the 100 simulated cases with aGvHD II-IV, approximately 43 and 44 observations had a CR and nCR, respectively. The median survival time modelled for CR was 3.2 and 0.5 years for nCR. Overall median survival for all patients irrespective of their response category was 9.6 months. Average per-person life-years were

Table 2 - Estimated overall survival probability in per cent

Response category	Years after the first response evaluation							
	1	2	5	10	20	50	80	99
CR	73.1	59.6	40.1	26.7	16.0	6.8	4.1	0
nCR	25.7	11.4	2.6	0.6	0.1	0	0	0
All	43.0	30.9	18.6	11.9	7.0	3.0	1.8	0

Modelling results plotted on empirical Kaplan-Meier (KM) curves

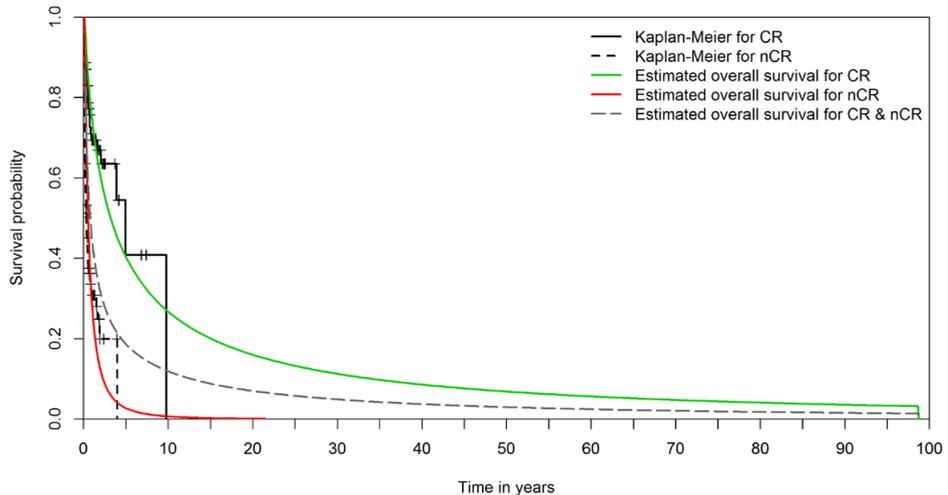


Figure 3 - Modelling results; KM: Kaplan-Meier; Markov trace = simulated, estimated survival of patients over time

estimated at 0.32 years (ie, 3.8 months). In Figure 3, modelling results are plotted on the empirical Kaplan- Meier curves for CR and nCR. Table 2 depicts the estimated overall survival probabilities for different years after the first response evaluation.

DISCUSSION

In this study, we present a DM for the second-line treatment of aGvHD. In line with current recommendations for good modelling practice, we involved an international group of clinical experts to develop this DM and used evidence available from the literature.^{189,191,208,210} Consequently, we ensured that the model sufficiently represents current clinical practice.

The DM can serve manifold purposes. Updated with clinical evidence, it can be used by clinicians and researchers to estimate longterm health outcomes of different treatment alternatives. When costs are added to the clinical evidence, for example, from RETHRIM, economic evaluations can be performed to inform reimbursement decision-makers on the implementation of treatment alternatives.

To run a first model simulation, we searched for available evidence on both costs and clinical evidence on treatment alternatives for aGvHD. We found that MSCs for the treatment of SR-aGvHD are widely studied. Therefore, we were able to integrate more than a decade of empirical data into our DM. Nevertheless, mainly due to the absence of randomised phase III studies, the number of patients included and the restricted follow-up periods, we faced several challenges in integrating the collected information. Consequently, we had made a number of assumptions.

First, MSC products and their administration varied between the studies. With the exception of the study by Fang et al²¹⁸ where MSCs were derived from human adipose tissue, all other studies used bone marrow-derived products. In addition, the number of MSC infusions (between 1 and 21 infusions) as well as the dosage of infused MSCs (between 0.6 and 20 x10⁶ cells per kg body weight) varied across the studies. In this regard, we assumed that type and administration of MSC products had no effect on transition probabilities or mortality rates. Whereas this made the integration of the study results possible, the data did not allow for further stratification to test potential MSC derivation-related or dose-related effects on the response rates.

Second, there were not enough data available to estimate transitions between response categories after a first response evaluation. We had assumed survival can be predicted based on the initial response category at 28 days post-MSD transfusion. However, in clinical practice response categories may change in any direction after the first evaluation. Of the included studies, only Prasad et al³⁶ assessed response to MSCs in paediatric patients more than once. In this study, five of twelve patients further improved after day 32 to a complete remission. Future studies measuring response on several time points after MSC treatment

could inform our model on subsequent changes in response categories. Whereas this enables a better estimate for subgroups, the effect on the average survival of the entire population is most likely negligible.

Third, it needs to be noted that the simulation did not consider the underlying haematological disease, nor the patients' age or sex. This choice was made as not all studies did report on these variables and any further stratification would have resulted in a reduction in the population on which the different estimates are based. However, we acknowledge that the underlying haematological disease and age and sex can be important determinants for long-term survival. Detailed reporting on patient characteristics and their diagnosis may help to enable further analyses for these subgroups.

Our modelling results, however, did show the expected longer survival after MSC transfusion for patients that achieved CR at first response evaluation, when compared to patients with nCR. These estimates are in line with previous findings highlighting the importance of complete response for long-term survival.^{235–240}

To our knowledge, there are no long-term survival results published for SR-aGvHD patients treated with MSCs other than the two reports^{241,242} that are based on studies^{190,229} included in our study. Therefore, we attempted to compare our results to survival estimates of studies that tested other second-line treatment options for aGvHD. To find suitable studies, we consulted the most recent NHS England clinical commissioning policy on the treatments for GvHD following HSCT.¹⁹⁵ This guideline is based on an extensive and updated review of the literature and concludes that there is sufficient evidence only for extracorporeal photopheresis (ECP) to be routinely commissioned for the treatment of aGvHD. Therefore, we focussed on a comparison with ECP. For other treatment options such as infliximab, etanercept, inolimomab, alemtuzumab, pentostatin or MSCs, the reviewers found that there was no sufficient evidence available to propose the routine commissioning for aGvHD.

Regarding survival of aGvHD patients treated with ECP, the largest series was published by Greinix et al.²³⁸ in a phase II prospective study.¹⁹⁵ Every week, 59 patients with steroid-refractory or steroid-dependent GvHD received two consecutive ECP treatments. CR and nCR was defined as in our study. The reported median survival was below 6 months after HSCT for patients with nCR, confirming our median survival estimations for nCR to MSC of approximately 6 months.²³⁸ In the study of Greinix et al,²³⁸ median survival for complete responders was never reached during the reported follow-up period of 9 years after HSCT. For this study, this implies that the probability of surviving nine years after HSCT would be approximately 59% for patients with a CR. With an estimate of approximately 27% at ten years post-MSCs, our estimates for patients with a CR are significantly lower. This may be explained by differences in patient population before treatment (ie, Greinix et al.²³⁸ included a higher number of patients with aGvHD grade II (61%)). In addition, the study by Greinix and colleagues was based on a limited amount of complete responders (n = 41).

Although we were able to demonstrate longer survival for patients with CR when compared to nCR, the relatively high mortality rate of SR-aGvHD, even with MSCs, can still be regarded as unacceptably high. And although there were numerous studies reporting on MSCs as a treatment alternative for SR-aGvHD, none of them included patient-reported outcomes (PROs) such as HRQoL. Updating our results with HOVON113 findings, including HRQoL measures, might improve the current survival estimates and show favourable treatment outcomes in terms of quality of life for MSCs when compared to placebo. However, until these study results are presented, this remains subject to speculation and the search for alternative (pre)treatments helping patients to achieve a complete response will have to continue.

CONCLUSION

The designed DM provides a comprehensive overview on the second-line treatment pathways for aGvHD in general. The model simulation with data from previously published studies on MSCs as a second-line treatment option for aGvHD presented outcomes matching the literature as well as clinical expectations. This demonstrates the practicability and usefulness of the model. However, to date, only insufficiently detailed data are available to fully model all health states and to perform a cost-effectiveness analysis. The yet restricted model simulation would therefore benefit from additional data, preferably from a phase III RCT. The integration of effectiveness results together with health-related quality of life measures (e.g., from the EQ-5D questionnaires) and different cost components derived from the RETHRIM trial could overcome this limitation and enable a full cost-utility analysis.