Emmes Group Emmes Veridix AI

Hematopoietic Stem Cell Transplantation Studies from a Statistical Perspective

BASS 2024

Laura Morrison

Senior Manager, Biostatistics

General Information

Emmes: Senior Manager, Biostatistics

Vancouver Island Life Sciences (VILS): Vice President

Clinical Trial Experience

01 Over 11 years experience in clinical trial research

02 Lead statistician on cell and gene therapy protocols

03 Oversee protocols spanning therapeutic areas including CGT

04 Substantially contributed to trials of various products that received FDA approval

Agenda

I have kindly received permission to share the information about a series of studies. However, I am not allowed to share the name of the company or product during this talk. I will therefore refer to the company as the sponsor and the product as investigational product.

01 Hematologic Malignancies

02 Investigational Product

03 Overview of Clinical Trials Conducted

04 Phase III Population, Design, and Endpoints

05 Randomization and Re-randomization

06 Statistical Tests

07 Key Results

Agenda

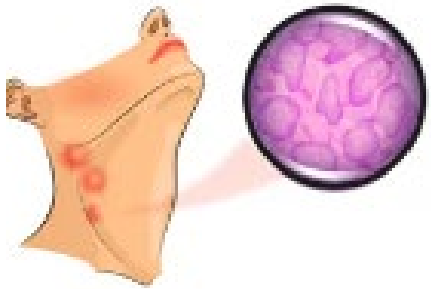
Hematologic Malignancies: What are they?

- Hematologic malignancies are tumors involving the bone marrow, blood, and lymphoid tissue
- They occur when abnormal cells in the blood grow uncontrollably
- They often begin in the bone marrow where stem cells develop into white blood cells, red blood cells, or platelets
- They impact the normal production and function of the blood cells
  - White blood cells help fight infections
  - Red blood cells help carry oxygen throughout the body
  - Platelets help clot blood
- They account for approximately 9% of all cancers and the prevalence is rising

Types of Hematologic Malignancies

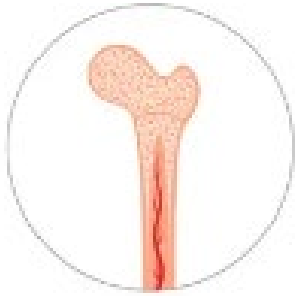
Leukemia

Begins in early blood-forming cells in the bone marrow



Lymphoma

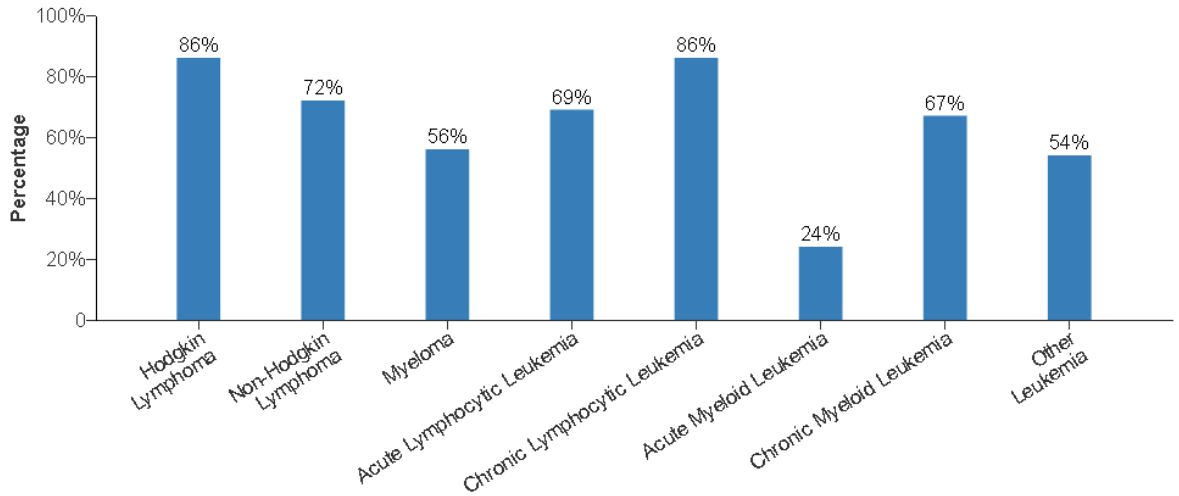
White blood cells that are part of the lymphatic system which helps fight infection



Myeloma

Affects the plasma cells in bone marrow; plasma cells make antibodies to help fight infections

Figure 2. 5-year Relative Survival for Malignant Hematologic Cancers, by Cancer Type[a].



Hematologic Malignancies: Treatments

- Treatment options may include a combination of:
  - Bone marrow transplantation
  - Chemotherapy
  - CAR T cell therapy
  - Immunotherapy
  - Radiation therapy
  - Targeted therapy

Agenda

Cell and Gene Therapy

- Use the term patient rather than subject or participant
- Small sample sizes
- Possibility of production failures of the product
- Endpoints may be complex and require clinical input leading to the need for an Endpoint Review Committee (ERC)

Investigational Product

Why do we need it?

- Often the best results are from Human leukocyte antigens (HLA)-identical sibling transplantations
  - More than 66% of patients lack a matched donor
  - Less than 50% of unrelated donor searches result in identification and availability of a matched donor graft
  - Long searches for suitable donors increases the risk of the patient having progression or relapse

- Patients with no suitable matched related donor or matched unrelated donor in a timely manner have the following options:
  - Mismatched unrelated donor (MMUD)
  - Haploidentical (haplo)-related donor
  - Umbilical Cord Blood (UCB)

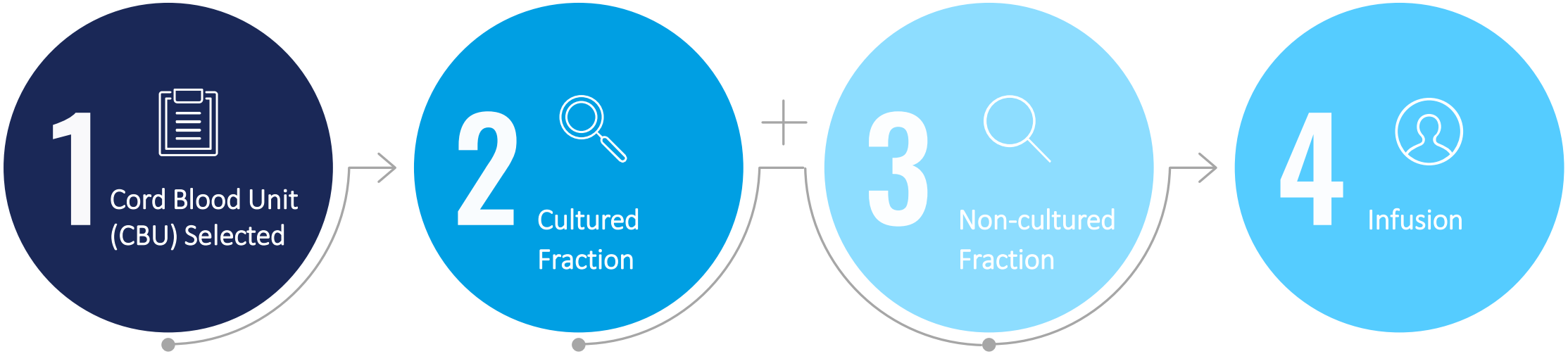
- To eliminate the need for double cord blood transplants

Investigational Product

What is it?

- A stem cell-based product
- Composed of ex-vivo expanded allogeneic cells from one unit of umbilical cord blood
- Utilizes cytokines and nicotinamide (NAM) as an approach to inhibit differentiation and to increase the migration, bone marrow homing and engraftment efficiency of hematopoietic progenitor cells expanded in ex vivo cultures
- Comprised of two components:
  - Ex vivo expanded, cord blood-derived, hematopoietic progenitor cells (cultured fraction)
  - Non-cultured cell fraction of the same CBU consisting of mature myeloid and lymphoid cells

Investigational Product



Cord Blood Unit selected by physician from public cord blood bank

Expanded hematopoietic stem cells using proprietary technology

Differentiated immune cells

Agenda

Clinical Development

Indication	Phase	Treatment	Sample Size
Hematologic Malignancies	Phase I	Investigational Product + UCB	N=11
Sickle Cell	Phase I	Investigational Product + UCB Investigational Product alone	N=13 N=3
Hematologic Malignancies	Phase II	Investigational Product alone	N=36
Hematologic Malignancies	Phase III	Investigational Product alone vs. UCB	N=125
Hematologic Malignancies	Expanded Access	Investigational Product alone	N=33
Hematologic Malignancies	ISS and ISE		

Agenda

Phase III: Study Population

- Study included patients with hematologic malignancies for whom allogeneic stem cell transplants is currently a recommended and potentially lifesaving treatment who do not have an adequate suitably matched and readily available stem cell donor

- 12-65 years of age
- Patients must have one of the following hematological malignancies:
  - ALL
  - AML
  - CML
  - CMMoL or MDS/CMMol overlap
  - MDS
  - Biphenotypic/undifferentiated/Prolymphocytic/Dendritic Cell Leukemias and Natural Killer Cell Malignancies
  - Lymphoma
- Performance score at least 70% by Karnofsky or Lansky
- Sufficient physiologic reserves
  - Cardiac
  - Pulmonary
  - Renal
  - Hepatic

Phase III: Analysis Populations

Intent-to-Treat (ITT)

Patients randomized into the trial

Classified by treatment assigned

Transplanted (TP)

Patients randomized who received a CBT on or before 90 days post randomization. Patients who received a CBT out of specifications are included

Classified by treatment assigned

As Treated (AT)

Patients randomized who received a CBT on or before 90 days post randomization. Patients who received a CBT OOS are not included.

Classified by treatment received

ANC Engrafted (AEP)

Patients who received a CBT on or before 90 days post randomization and achieved neutrophil engraftment by Day 42 with subsequent chimerism

Classified by treatment received

Platelet-Engrafted (PEP)

Patients who received a CBT on or before 90 days post randomization and achieved platelet engraftment

Classified by treatment received

Safety (SP)

Same as As Treated population

Treatments

Conditioning Regimen and Graft versus Host Disease (GVHD) Prophylaxis

All patients were treated with a conditioning regiment prior to transplantation. Sites could select from any of the protocol's three myeloablative conditioning regimens.

All patients were administered GVHD prophylaxis medications pre- and post-transplant.

Unmanipulated Cord Blood

All cord blood units were from public banks that meet regulations.

Investigational Product

Cyropreserved cell-based product of allogeneic, ex vivo expanded, umbilical cord blood-derived, hematopoietic CD34+ progenitor cells and the non-expanded cell fraction of the same cord blood unit consisting of mature myeloid and lymphoid cells.

Primary Endpoint and Hypothesis

Primary Endpoint

- Primary Endpoint: Time from transplant to neutrophil engraftment

Primary Hypothesis

- Primary hypothesis

  H_0: There is no difference in time to neutrophil engraftment by treatment group

  H_A: There is a difference in time to neutrophil engraftment by treatment group

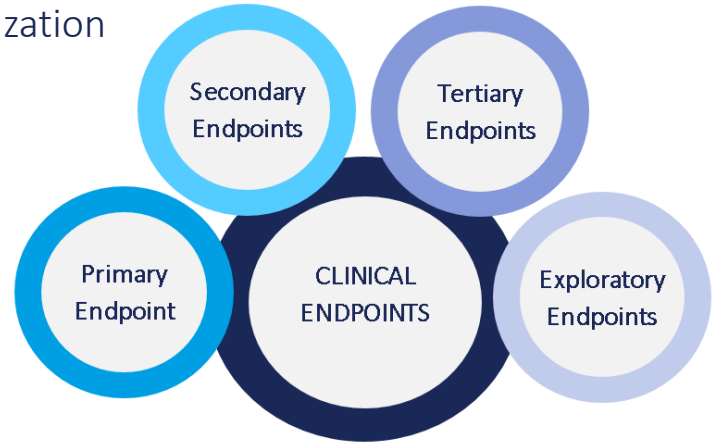
Secondary, Tertiary, and Exploratory Endpoints

Secondary Endpoints

- Incidence of grade 2/3 bacterial or invasive fungal infections by 100 days following transplantation
- Days alive and out of hospital in the first 100 days following transplantation
- Platelet engraftment by 42 days following transplantation

Tertiary Endpoint

- Non-relapse mortality by 210 days following randomization



Exploratory Endpoints

- Neutrophil engraftment by 16 days following transplantation
- Time from transplantation to platelet engraftment
- Duration of primary hospitalization
- Non-relapse mortality by 130 days and 15 months following randomization
- Overall survival by 210 days and 15 months following randomization
- Disease free survival by 15 months following randomization
- Neutrophil engraftment by 42 days following transplantation
- Acute GvHD grade II-IV and III-IV by 100 days following transplantation
- Chronic GvHD (mild/moderate/severe) by 180 days and 1 year following transplantation
- Secondary graft failure by 1 year following transplantation
- Grade 3 viral infections by 180 days and 1 year following transplantation
- Etc.

Sample Size Considerations

- Primary Endpoint: Time from transplant to neutrophil engraftment

- Analysis based on the Mann Whitney test statistic if there are no losses to follow-up before Day 42 following transplantation
- Significance level based on the re-randomization distribution
  - Simulations confirmed the power will be similar to significance based on the usual permutation distribution or it's normal approximation

- Primary sample size calculations were based on Noether's formula

Mann-Whitney U Test

- Alternative name: Wilcoxon-Mann-Whitney test
- Assumptions:
  - Two independent populations
  - Continuous responses that do not have ties
  - If it is additionally assumed that the two populations are identical except for a difference in location, this test can be used as a test of equal means or medians
- Suppose sample size N_1 and N_2 are from two independent continuous random variables Y_1 and Y_2
- The Mann-Whitney U statistic is calculated by counting the number of times an observation randomly selected from the first population is greater than an observation randomly selected form the second population
  - $U = \#\,(Y_{1i} > Y_{2j})$, i= 1, …, N_1; j=1, …, N_2
- $\frac{U}{N_1 N_2}$ is an estimate of the competing probability $P_1 = Pr(Y_1 > Y_2)$

Noether's Formula

- Sample size formula for the Mann-Whitney U two-sided test is provided by Noether (1987) assuming there are no ties in the data
  - $N = \frac{\left(z_{1-\frac{\alpha}{2}} + z_{\beta}\right)^2}{12\frac{N_1}{N_1+N_2}\left(1-\frac{N_1}{N_1+N_2}\right)\left(P_1-\frac{1}{2}\right)^2}$

- P_1 is the probability that a patient who received the investigational product has a shorter neutrophil engraftment time than a patient who received unmanipulated cord blood

Sample Size Assumptions

- The estimate of P ($1 - U/(N_1N_2)$) was based on data from a Phase II study where 36 patients received a single cord transplant of the investigational product and 152 patients in the CIBMTR registry database who received an unmanipulated cord blood transplant and determined to be 0.88.
- No loss to follow-up
- 10% of the patients in the CIBMTR dataset and 0% of the investigational product patients will not achieve neutrophil engraftment
- 10% of patients randomized to the investigational product and 4% of the patients randomized to the unmanipulated cord blood arm will not receive a transplant due to intervening events between the decision to get a transplant and the transplant date itself. The higher percentage in the investigational product arm is due to the time between randomization and transplant being longer.
- 5% of patients assigned to the investigational product arm will receive an unmanipulated cord blood transplant due to investigational product expansion failures
- Under the null, P will be 0.50
- 2-sided test with a 5% significance level

Sample Size Assumptions Continued

- Original estimate of P based on previous data: 0.88
- Given the additional assumptions, the estimate was recomputed
  - P(IP txp)*P(UCB txp)*P(IP>UCB) + P(IP txp)*P(no UCB txp)*P(IP>UCB) + P(no IP txp)*P(UCB txp)*P(IP given UCB) + P(no txp)*P(null)
  - (0.9)*(0.96)*(0.88) + (0.9)*(0.04)*(1.00) + (0.1)*(0.96)*(0.05) + (0.1)*(0.04)*(0.5) = 0.80
- An additional 5% batch failure rate was introduced
  - (0.85)*(0.96)*(0.88) + (0.05)*(0.96)*0.06 + (0.85)*(0.04)*(1.0) + (0.05)*(0.04)*(0.95) + (0.1)*(0.96)*(0.05) + (0.1)*(0.04)*(0.5)
  - This further reduces P to 0.78
- Our best estimate of P is 0.78
- Using a two-sided test with a 5% significance level, a test under the null that P=0.50 and alternative that P=0.78 requires a sample size of 45 for 90% power.

Sample Size Assumptions Continued

- The protocol inflated the sample size to 120 patients for the following reasons:
  - To provide an extensive safety database for the investigational product
  - To determine whether a difference exists between the two groups for the primary endpoint while ensuring the statistical significance is strong and highly convincing
  - To reduce the chance of seeing higher mortality in the investigational product group than the control group even if the investigational product has a beneficial effect on mortality
- With an update to the inclusion criteria allowing patients aged 12-17 years old in the study, the statistical power was recalculated.
  - There was an assumption that patients aged 12-17 would have a median time to engraftment of 16 days rather than the 22 days observed in the CIBMTR database.
  - With 120 patients, power was assessed for 20, 30, and 40 adolescent patients to be 98%, 97%, and 93%, respectively.
  - A decision was made to cap the number of adolescent patients to 30.

Study Design

Open Label

Investigators and patients will be aware of which product was given.

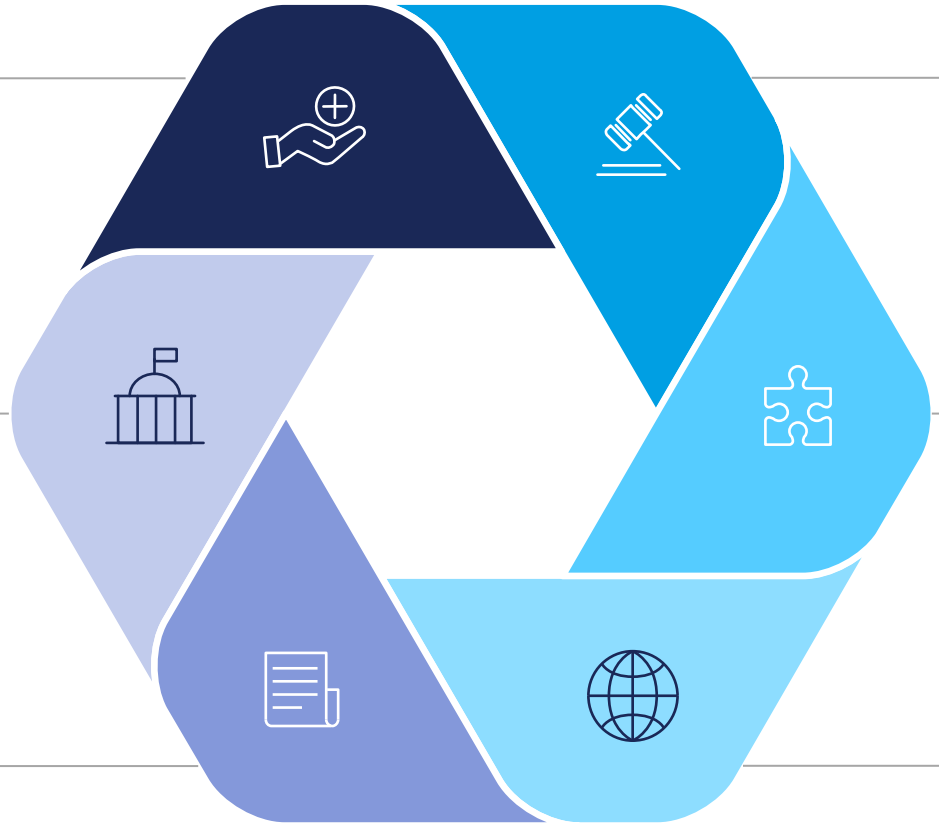
Randomized

Study participants are assigned a treatment arm at random.

Arm 1: Investigational Product
Arm 2: One or Two unmanipulated, unrelated cord blood units

Phase III

Study tests safety and efficacy and compares to standard of care.



Controlled

Compares the investigational product to the current standard of care which in this case is an umbilical cord blood transplant.

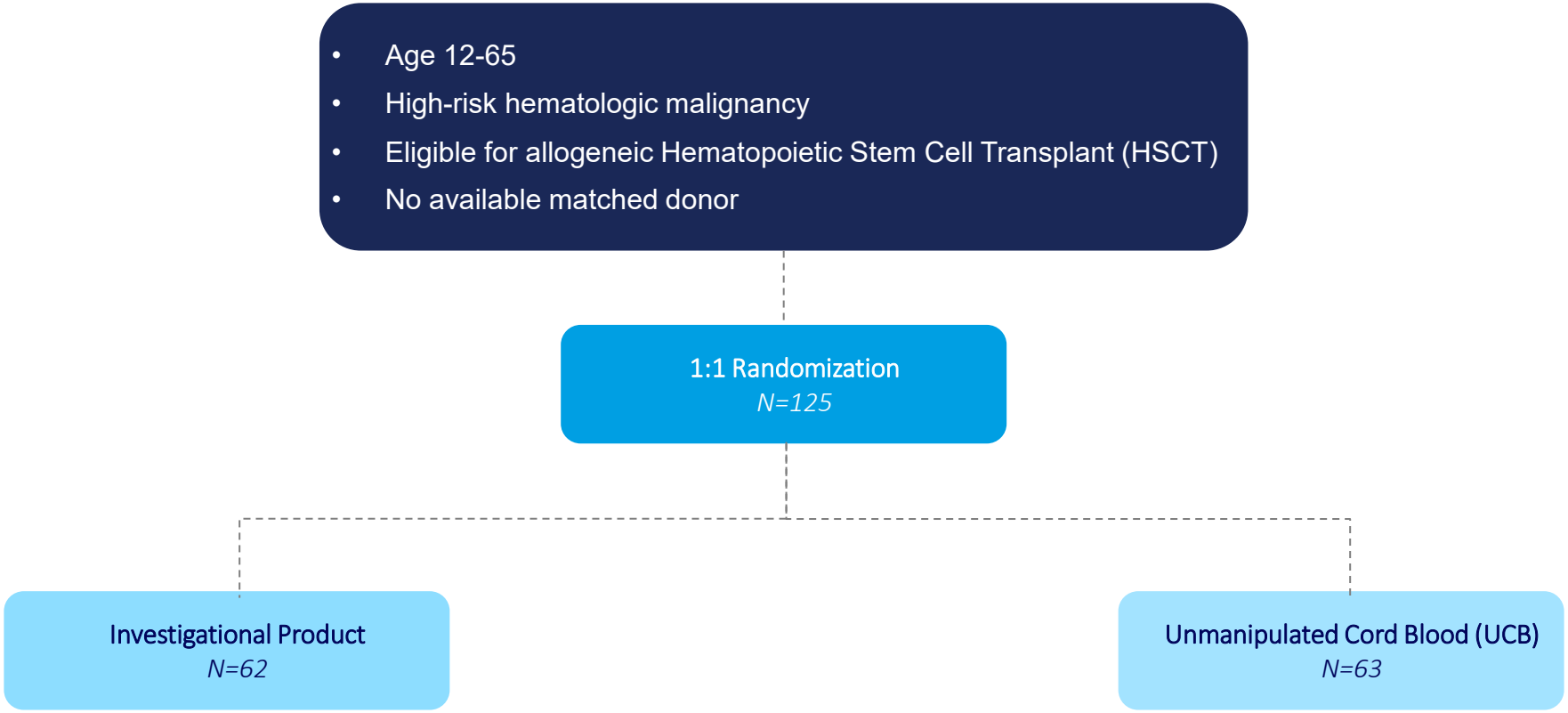
Multicenter

Study utilizes multiple sites.

International

Sites located in North America, South America, Asia, and Europe.

Study Design



Study Design

- Age 12-65
- High-risk hematologic malignancy
- Eligible for allogeneic HSCT
- No available matched donor

1:1 Randomization
N=125

Investigational Product
N=62

Unmanipulated Cord Blood (UCB)
N=63

Primary Endpoint
Day 42 post transplantation

Secondary Endpoint
Day 100 post transplantation

Exploratory Endpoint
Month 15 post randomization

Challenges

- Low accrual at the beginning of the trial
  - The DMC monitored accrual monthly
  - There were discussions as to whether the study was feasible when accrual fell quite far behind
  - The sponsor made the decision to open more sites to help increase accrual
- COVID was at its peak during the trial
  - Missing data was a potential concern if patients couldn't get back to the transplant center
  - Pivoted to allow some information to be collected via tele-medicine visits; others had to remain in person
  - Added questions related to COVID on forms so we could better understand the impact it had on the study

Agenda

Randomization

- An essential component of clinical trials to provide comparison of treatment with precision and validity
- Considered by many researchers to be the optimal approach for assigning participants to a treatment arm in clinical trials
- Process by which participants are enrolled into trial arms only by chance
- If there are factors known to influence participant outcome(s) to treatment, these should be taken into account when randomizing patients

Reasons to Randomize

- Do not want participants in treatment arms to differ in a systematic way
  - If treatment arms are systematically different, trial results will be biased
    - Example: Consider a study looking at a mobility intervention in which a greater proportion of individuals with physically disabilities and older adults are assigned to the treatment arm over the control arm. The imbalance may influence the outcome of the study, and the effects of the treatment may be indistinguishable from the influence of the imbalance of covariates.
- Without randomization, the ability to select a participant's treatment arm may create potential bias and may taint the data
- Potential disadvantage: a consequence of some randomization methods can be imbalance among treatment arms with respect to some prognostic factors, which may invalidate the trial results or require complex secondary analysis to address the source of imbalance

Motivation for Minimization

- Commonly used randomization methods such as Simple Random Sampling, Stratification, and Permuted Blocks may result in treatment imbalance with respect to prognostic factors or with respect to the number of participants in each treatment arm, or may result in empty or sparse strata
- Minimization may be desirable when there are prognostic factors strongly associated with a study endpoint
- Minimization may be advantageous in ensuring only minor differences between treatment arms in the prognostic factors used in the treatment allocation process, particularly in trials with multiple prognostic factors and a small sample size, in which case balance using randomization methods such as stratification randomization is improbable

Minimization Treatment Allocation

- Aims to ensure treatment arms are balanced with respect to predefined participant factors as well as balance the number of participants in each treatment arm
- Construct measures of imbalance for each treatment arm when an eligible participant is ready to be randomized
- The participant is assigned to the treatment which yields the lowest imbalance score. If all imbalance scores are equal, the participant is randomly assigned a treatment.
- Simulation studies show that minimization provides better balanced treatment arms when compared with restricted or unrestricted randomization and that it can incorporate more prognostic factors than stratified randomization methods such as permuted blocks
- Can limit imbalance in randomized clinical trails, including relatively small trials, with more important prognostic covariates
- Important covariates are identified before the start of the trial
- Pure minimization is deterministic meaning one can predict the treatment assignment of the next participant by knowing the fact levels of previously enrolled participants and having the properties of the next participant. To eliminate predictability of treatment allocation, some elements of randomness in the minimization algorithm may be included.

Advantages and Disadvantages of Minimization

Advantages

- May achieve better balance than more conventional randomization schemes
- Treatment arms are balanced with respect to predefined participant factors as well as the number of participants within each treatment arm
- Larger number of stratifying factors may be incorporated

Disadvantages

- Allocating participants to treatment arms via minimization is more complex than many other randomization methods
- Randomization lists cannot be generated in advance
- Treatment assignments may be predicted with certainty in some situations
- There are implications for the analysis method used
- There is a price to pay for imposing extra restrictions on the randomization, such as a lack of control of type I error rate if re-randomization is not used
- Does not perform well with unequal allocation, in particular 2k:1 randomization with an odd number of stratification covariates

Minimization Example

- Two treatment arms (Treatment A, Treatment B)
- Four known key factors that are to be balanced across treatment arms:
  - Factor C (C1, C2, C3, C4)
  - Factor D (D1, D2)
  - Factor E (E1, E2, E3)
  - Factor F (F1, F2, F3, F4, F5, F6)

Participant ID	Factor C	Factor D	Factor E	Factor F
P001	C1	D2	E1	F1
P002	C1	D1	E2	F3
P003	C2	D2	E2	F1
…	…	…	…	…

Minimization Example Continued

- Assign treatment assignments as follows (note: study needs to decide on the amount of randomness to add -- below uses 20%)
  - If numA > numB: assign to treatment A with 20% probability (random number < 0.2)
  - If numA < numB: assign to treatment A with 80% probability (random number < 0.8)
  - If numA = numB: assign to treatment A with 50% probability (random number <0.5)

  where

  - numA = count of those assigned to treatment A in this participant's Factor C
    + count of those assigned to treatment A in this participant's Factor D
    + count of those assigned to treatment A in this participant's Factor E
    + count of those assigned to treatment A in this participant's Factor F
  - numB = count of those assigned to treatment B in this participant's Factor C
    + count of those assigned to treatment B in this participant's Factor D
    + count of those assigned to treatment B in this participant's Factor E
    + count of those assigned to treatment B in this participant's Factor F

Minimization Example Continued

Participant ID	Factor C	Factor D	Factor E	Factor F	numA	numB	Result	Random Number	Treatment Assignment
P001	C1	D2	E1	F1	0	0	numA = numB	0.7371389 (≥0.5)	Treatment B
P002	C1	D1	E2	F3	0	1	numA < numB	0.7670458 (<0.8)	Treatment A
P003	C2	D2	E2	F1	1	2	numA < numB	0.3282159 (<0.8)	Treatment A
…	…	…	…	…	…	…	…	…	…

- Participant P001:
  - Treatment A and B: no participants currently assigned (numA = 0, numB = 0)
- Participant P002:
  - Treatment A: no participants currently assigned (numA = 0)
  - Treatment B: P001 (same Factor C)
    - numB = 1 (Factor C) + 0 (Factor D) + 0 (Factor E) + 0 (Factor F) = 1
- Participant P003:
  - Treatment A: P002 (same Factor E)
    - numA = 0 (Factor C) + 0 (Factor D) + 1 (Factor E) + 0 (Factor F) = 1
  - Treatment B: P001 (same Factor D and F)
    - numB = 0 (Factor C) + 1 (Factor D) + 0 (Factor E) + 1 (Factor F) = 2

Re-randomization

- Preferred analysis method when minimization is used
- The distribution is approximated by re-computing the minimization algorithm using different seeds thousands of times and re-computing the test statistic each time
- The p-value is approximated by the proportion of re-randomized datasets whose test statistic is at least as extreme as the original one
- When unequal allocation is used, it can lead to issues when re-randomized is used with minimization
  - Mean of the re-randomization distribution may be non-zero (note: mean is zero for equal allocation)
  - Re-randomization and t-test results may not be asymptotically equivalent
  - Protects against temporal trends but causes a potential serous loss of power compared to normal theory inference
  - Additional concerns when using an allocation of 2k:1 with an odd number of stratification covariates

Re-randomization Programming

```
***Set maximum size of strata and number of datasets to be created***;
%let stratalim=85; *max size of strata;
%let datanum=8000; *number of datasets;

***Create 8000 treatment tables to be used in the re-randomization
Add 'rr' after variable name to indicate re-randomization variables***;
data trttab;
          length rrstrata $3 trttruerr $30;
          do rriter=1 to &datanum.; *'datanum' datasets to be generated;
                    do rrstrata='001','002','003'; *dataset to contain 3 strata;
                              do trtseqrr=1 to &stratalim.; *each strata to contain 'stratalimit' records;
                                        rv=ranuni(&rrseed.); *set random number;
                                        *assign strata based on rv;
                                        *Strata 1;
                                        if rrstrata='001' then do;
                                                  if rv^=. and rv<=0.10 then trttruerr='Investigational Product';
                                                  else if rv^=. and rv>0.10 then trttruerr='Control';
                                        end;
                                        *Strata 2;
                                        if rrstrata='002' then do;
                                                  if rv^=. and rv<=0.90 then trttruerr='Investigational Product';
                                                  else if rv^=. and rv>0.90 then trttruerr='Control';
                                        end;
                                        *Strata 3;
                                        if rrstrata='003' then do;
                                                  if rv^=. and rv<=0.50 then trttruerr='Investigational Product';
                                                  else if rv^=. and rv>0.50 then trttruerr='Control';
                                        end;
                                        output;
                              end;
                    end;
          end;
run;
```

Re-randomization Programming Continued

```
***Organize treatment table so that there is 1 row per dataset iteration***;
data trttabiter;
          set trttab;

          *organize by dataset iteration;
          by rriter;

          *set array up (3 strata, stratalim variables per strata);
          array trtstrata1 $ trttruestrata001_1-trttruestrata001_&stratalim.;
          array trtstrata2 $ trttruestrata002_1-trttruestrata002_&stratalim.;
          array trtstrata3 $ trttruestrata003_1-trttruestrata003_&stratalim.;
          retain    trttruestrata001_1-trttruestrata001_&stratalim.
                              trttruestrata002_1-trttruestrata002_&stratalim.
                              trttruestrata003_1-trttruestrata003_&stratalim.;

          *assign the array to the trttruerr values based on their strata;
          if rrstrata='001' then trtstrata1[trtseqrr]=trttruerr;
          else if rrstrata='002' then trtstrata2[trtseqrr]=trttruerr;
          else if rrstrata='003' then trtstrata3[trtseqrr]=trttruerr;

          *Array now contains stratalim*datanum rows;
          *Only need the last row of each dataset iteration as it will contain the information
                    for all stratalim variables in each strata;
          if last.rriter then output;

          *drop variables that are no longer needed;
          drop rrstrata trttruerr trtseqrr rv;
run;
```

Re-randomization Programming Continued

```
***Get order in which subjects were enrolled and randomization factor information***; <code>

***Determine number of subjects assigned to each treatment as in the original EDC assignments**; <code>

***Determine number of enrolled subjects***; <code>

***Assign the variables for each iteration all at once and then parse out afterwards***;
data randfact;
          set enrorder;
          do iter=1 to &numsub.;
                    if enterseq<=iter then do;
                              if mdcnumran=. then do;
                                        <set each factor to the original factor entered in EDC>
                              end;
                              else do;
                                        if mdcnumran<iter then do;
                                                  <set each factor to the current factor entered in EDC>
                              end;
                                        else if mdcnumran>=iter then do;
                                                  <set each factor to the original factor entered in EDC>
                                        end;
                              end;
                              output;
                    end;
          end;
run;
```

Re-randomization Programming Continued

```
***Separate out the information by randomization factor;
One factor per dataset
Order dataset by randomization order (enterseq)***;
For each factor:
proc transpose data=randfact out=rand<factor> prefix=<factor>;
          by iter;
          id enterseq;
          var <factor variable>;
run;

***Merge the randomization factor information back together***;
<code>

***create &datnum version of the dataset to be merged with the re-randomization treatment datasets***;
data randfactdup;
          set randfactor;
          do rriter=1 to &datanum.;
                    output;
          end;
run;

***Merge the randomization factor information with treatment table information***;
<code>
```

Re-randomization Programming Continued

```
***Calculate nTreatment1 and nTreatment 2 values at each iteration***;

        *first subject in each dataset will always have counts of 0 since there is nobody to match with; <code>
        *re-set nTreatment1 and nTreatment2 to 0; <code>

        *store variables for randomization factors in array for easier use;
        array <factor> <factor>-<factor>&numsub.;

        *create array to store treatment assignments - retain them for easy use in other rows;
        array trts $ trt1-trt&numsub.;

        do i=1 to &enterseq-1; *go from first record to record immediately prior to desired record;
                nTreatment1= <count of subjects with same treatment factors in treatment 1 group> + nTreatment1;
                nTreatment2= <count of subjects with same treatment factors in treatment 2 group> + nTreatment2;
        end;

        *determine strata;
        *Strata 1;
        if nTreatment1>nTreatment2 then do;
                rrstrata='001';
                nstrata1=nstrata1+1;
        end;
        *Strata 2;
        if nTreatment1<nTreatment2 then do;
                rrstrata='002';
                nstrata2=nstrata2+1;
        end;
        *Strata 3;
        if nTreatment1=nTreatment2 then do;
                rrstrata='003';
                nstrata3=nstrata3+1;
        end;
```

Re-randomization Programming Continued

```
*create arrays to store treatment codes;
array trtstrata1 $ trttruestrata001_1-trttruestrata001_&stratalim.;
array trtstrata2 $ trttruestrata002_1-trttruestrata002_&stratalim.;
array trtstrata3 $ trttruestrata003_1-trttruestrata003_&stratalim.;

*need to assign the same number of treatment1/treatment2 patients as in original EDC assignments;

*if max has not been reached in either trt group;
if nTreatment1assig<&nTreatment1orig. and nTreatment2assig<&nTreatment2orig. then do;
          *set treatment based on sequence number within that strata;
          if rrstrata='001' then trts[enterseq]=trtstrata1[nstrata1];
          if rrstrata='002' then trts[enterseq]=trtstrata2[nstrata2];
          if rrstrata='003' then trts[enterseq]=trtstrata3[nstrata3];
end;
*if max has been reached in a trt group;
else if nTreatment1assig=&nTreatment1orig. then do;
          trts[enterseq]='Treatment2';
          rrstrata='max';
end;
else if nTreatment2assig=&nTreatment2orig. then do;
          trts[enterseq]='Treatment1';
          rrstrata='max';
end;

*assign a variable with the final treatment;
length trtassig $6;
trtassig=trts[enterseq];

*about how many subjects have been assigned to each trt group;
if trts[enterseq]= Treatment1' then nTreatment1assig=nTreatment1assig+1;
if trts[enterseq]='Treatment1' then nTreatment2assig=nTreatment2assig+1;
```

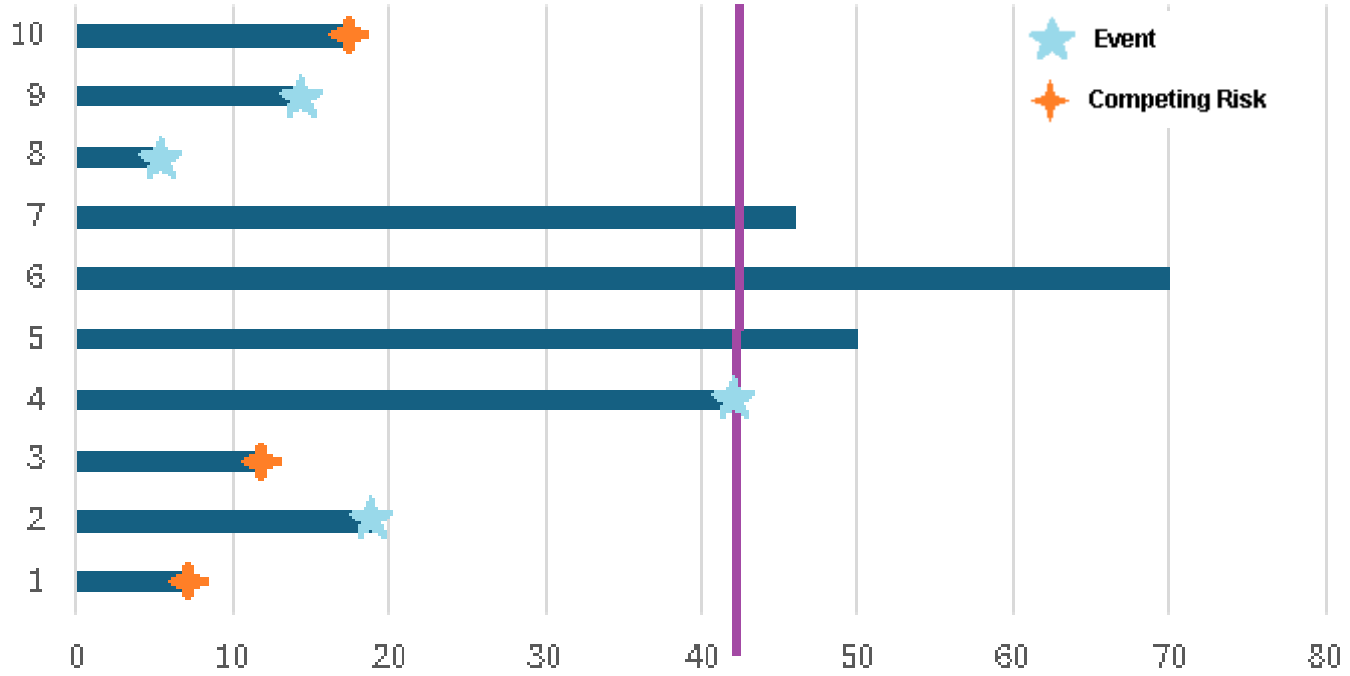
Challenges Encountered With Minimization

- Cannot create a treatment table that shows which treatment patient X will receive
- Had to write code to determine which treatment the patient should be assigned to and have it run automatically each time a patient was to be randomized
- There were instances where the site didn't enter the correct information for the factors used in the minimization algorithm causing a potential for the patient to be randomized to a group other than what they would have if all factor had been entered correctly
- All future patients randomized treatment assignments depend on previously patient treatment assignments. Therefore, when a patient has incorrect factor information entered, it's essential to correct the data in the database as soon as possible to ensure future patients are assigned to the correct treatment group.
- Re-randomization is computationally heavy and took more time to run than other methods and therefore we had to plan accordingly as to when we were going to run the primary analysis on the primary endpoint as our computers would be out of commission for a while

Agenda

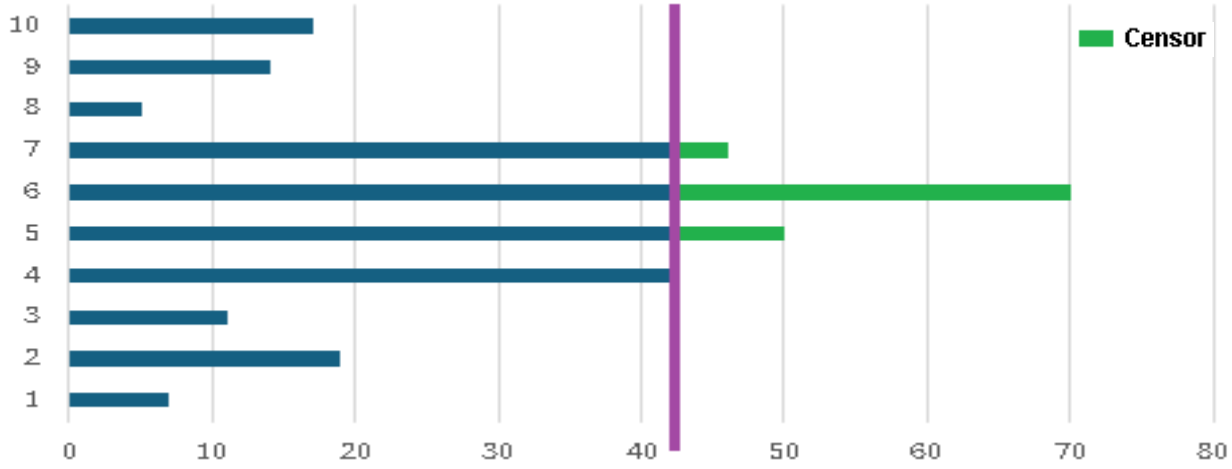
Competing Risks

- In some cases, a patient may fail due to one of K (K ≥ 2) causes, called competing risks
- Typically, in cell and gene therapy trials, we have competing risks
- Example: Neutrophil Engraftment
  - Competing risks: Primary graft failure, secondary graft failure, death, etc.



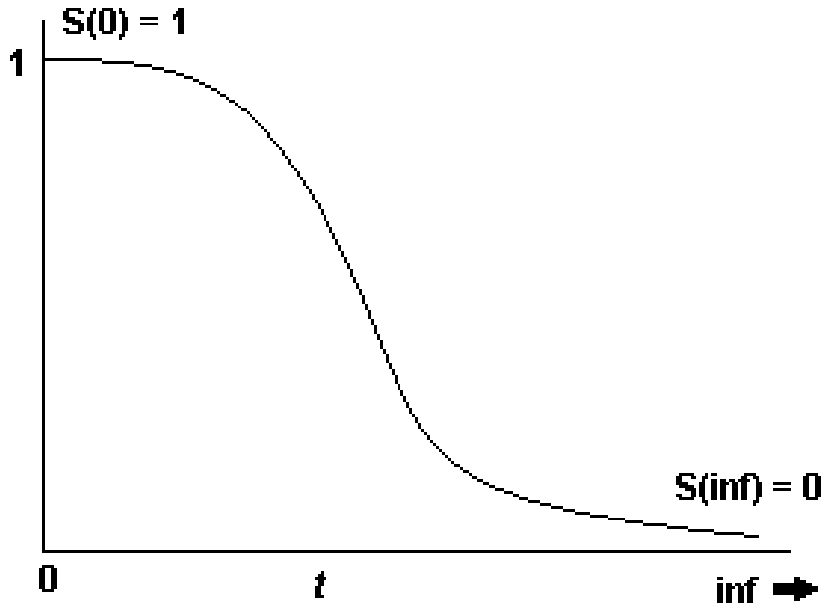
Right Censoring

- Occurs when the event is observed only if it occurs prior to a prespecified time point
- Exact timing of the event is unknown, but it is known that if the event occurred, the event happened after the last time the subject was observed
- Common feature in time-to-event analysis
- Examples of right censoring
  - Study ends before an event occurs
  - Study leaves study before an event occurs (loss to follow-up)



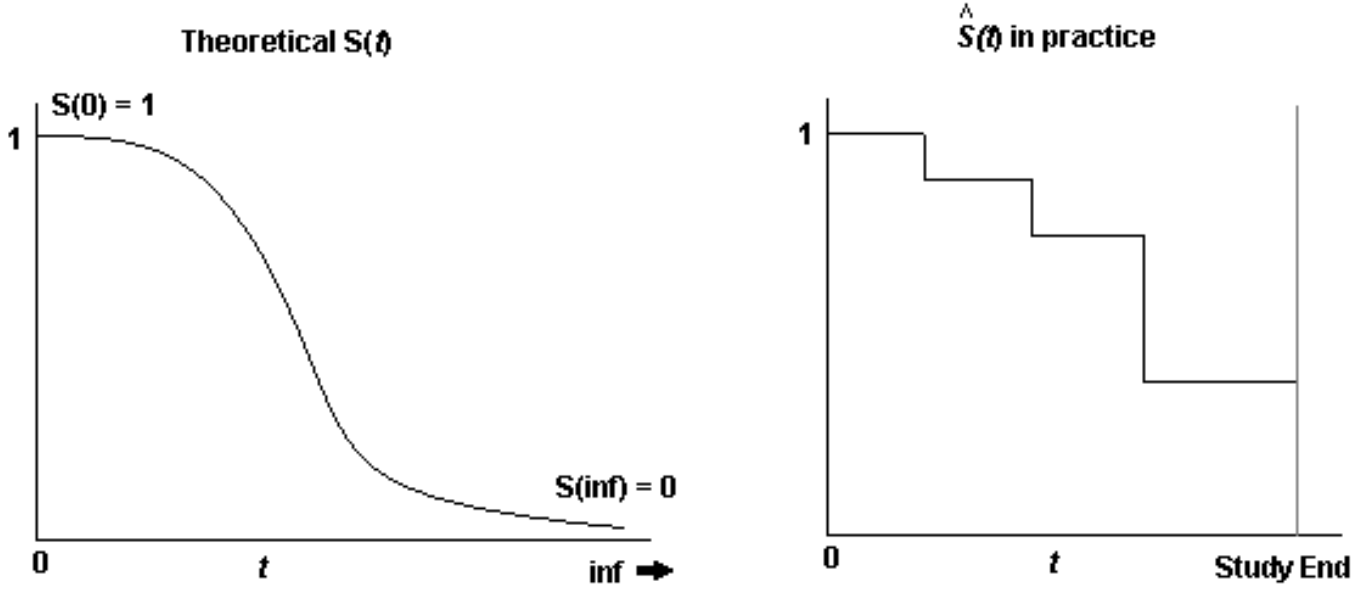
Survival Function

- Probability of not experiencing an event beyond time x
- $S(x) = P(X > x) = \int_x^{\infty} f(t)\, dt$
- Basic properties:
  - Monotone
  - Non-increasing
  - Equal to 1 at time 0 and 0 as time approaches infinity



Kaplan-Meier

- The standard estimator of the survival function is the Kaplan-Meier estimate
- For all values of t in the range where there is data
  - If $t < t_1$ $\quad \hat{s}(t) = 1$
  - If $t_1 \leq t$ $\quad \hat{s}(t) = \Pi_{t_i \leq t} \left[1 - \frac{d_1}{y_i}\right]$
- Step function with jumps at observed event times



Kaplan-Meier in Practice

- Consider a dataset with the follow characteristics
  - Dataset name: adam
  - Treatment codes: trt01pn
  - Event time: aval
  - Indicator variable: cnsr (0 = event, 1 = censor)
  - Interested in estimates at years 1-5

```
ods graphics on;
ods listing close;
proc sort data=adam; by trt01pn; run;
proc lifetest data=adam outsurv=kmest method=km plots=survival(atrisk=0 to 1825 by 365) noprint;
        by trt01pn;
        time aval*cnsr(1);
        title 'KM';
run;
ods listing;
ods graphics off;
```

Hazard Rates and Function

- Hazard rate is defined by

$$h(x) = \lim_{\Delta x \to 0} \frac{P[x \le X < x + \Delta x \mid X \ge x]}{\Delta x}$$

- If X is a continuous random variable, then

$$h(x) = f\,(x)/s\,(x) = -\,d\ln[s(x)]dx$$

- Cumulative hazard function is defined by

$$H(x) = \int_0^x h(u)\,du = -\ln[s(x)]$$

Cause-specific Hazard Rates

- Let i = 1, 2, …., K be the potential unobservable time to occurrence of the i^{th} competing risk
- Observe time at which patient fails from any cause $T = \min(X_1, X_2, \ldots, X_p)$
- Let δ be an indicator of which of the K risks caused the patient to fail

  $\delta = I$ if $T = X_i$

- Cause-specific hazard rate for risk I is defined by

$$h_i(t) = \lim_{\Delta t \to 0} \frac{P[t \le T < t + \Delta t\,,\ \delta = i \mid T \ge t]}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P[t \le Xi < t + \Delta t\,,\ \delta = i \mid Xj \ge t, j = 1,2,\ldots,K]}{\Delta t}$$

- $h_i(t)$ gives the rate at which patients who have yet to experience any of the competing risks are experiencing the i^{th} competing case of failure
- The overall hazard rate of the time to failure, T, is given by the sum of the K cause specific hazard rates

$$h_T(t) = \sum_{t=1}^{k} h_i(t)$$

Cause-specific Hazard Rates Continued

- The cause-specific hazard rate can be derived from the joint survival function of the K competing risks
- Let $S(t_1, t_2, \ldots, t_k) = P(X_1 > t_1, X_2 > t_2, \ldots., X_k > t_k)$
- The cause-specific hazard rate is given by

$$h_i(t) = \frac{-\partial s(t_1,\ldots tk)\partial t_i|_{t_1=\cdots=t_k=t}}{s(t,\ldots,t)}$$

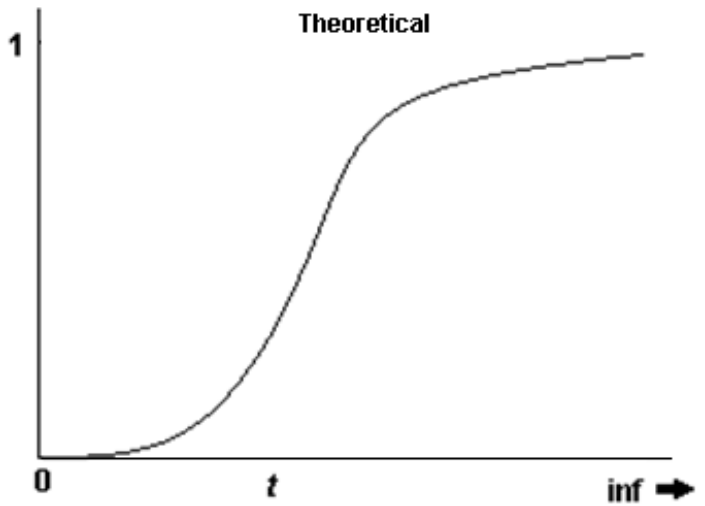
Cumulative Incidence

- Often we are not interested in the hazard rate, but instead in a probability summarizing the likelihood of occurrence of a particular event
- The probability can be expressed by the cause-specific sub-distribution function, known as the cumulative incidence function

$$F_i(t) = P[T \leq t, \delta = i] = \int_0^t h_i(u) \exp\{-H_T(u)\}\, du$$

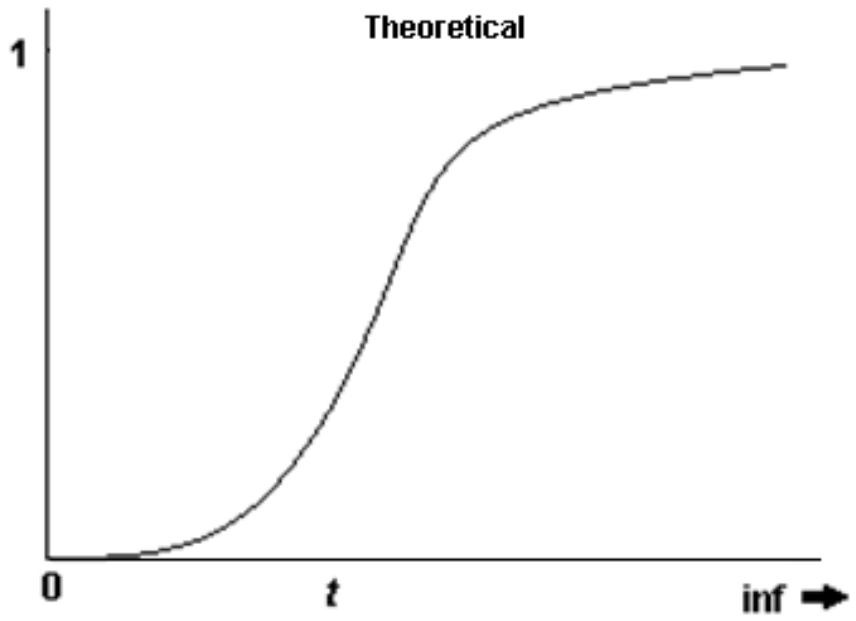
where $H_T(t) = \sum_{j=1}^{k} \int_0^t h_j(u)\, du$ is the cumulative hazard rate of T

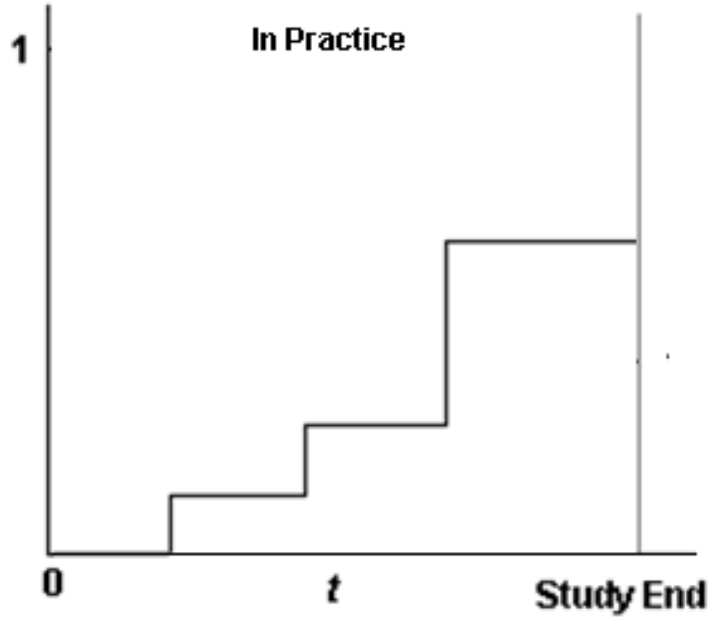
- $h_i(t)$ can be estimated from the observed data, meaning $F_i(t)$ is estimable without making assumptions about the join distribution of the potential failure times
- Properties of $F_i(t)$
  - Non-decreasing
  - $F_i(0) = 0$
  - $F_i(\infty) < 1$



Cumulative Incidence Continued

- Expresses the probability an event has occurred by a given time t
- Starts at 0 and increases stepwise to a maximum of 1
- In the absence of competing risks, this is equivalent to 1 – survival probability





Cumulative Incidence in Practice

- Consider a dataset (adam) with the follow characteristics
  - Treatment codes: trt01pn
  - Event time: aval
  - Indicator variable: cnsr (0 = event, 1 = censor, 2 = competing risk)
  - Interested in N at risk values at years 1-5

```
proc lifetest data=adam outcif=cinic plots=cif noprint;
        by trt01pn2 ;
        time aval*cnsr(1) / eventcode=0;
run;

ods listing close;
ods graphics on; ods output survivalplot=natrisk;
proc lifetest data=adam method=km plots=survival(atrisk=0 to 1825 by 365 atrisklabel
        atrisktickonly) timelist=0 to 1825 by 365;
        by trt01pn;
        time aval*cnsr(1,2) ;
run;
ods listing;
```

Log-Rank Test

- Test the null hypothesis that there is no difference between the populations in the probability of an event at any time point
- Assumes the hazard functions for the two treatment groups are proportional
- Most likely to detect a difference between groups when the risk of an event is consistently greater for one group than the other; unlikely to detect a difference when survival curves cross
- For each of the K failure times across two randomized groups at times t_1, t_2, …, t_K, a 2x2 table is constructed. For failure time t_k, k = 1,2,3…,K, the table is:

	Treatment 1	Treatment 2
# events	d_{Pk}	d_{Ek}
# non events	n_{Pk}-d_{Pk}	e_{Ek}-d_{Ek}

- The log-rank statistic is:

$$Z_L = \frac{0-E}{\sqrt{V_L}} = \frac{\sum_{k=1}^{K}\left(\frac{n_{Pk}\,dEk \;\; n_{Ek}\,dPk}{n_{Pk} \mp n_{Ek}}\right)}{\sum_{k=1}^{K}\left(\frac{(dPk+dEk)(n_{Pk}+n_{EK}-dPK-dEK)n_{Pk}n_{Ek}}{(n_{Pk}+nEk-1)(n_{Pk}+n_{Ek})^2}\right)}$$

Log Rank Test in Practice

- Consider a dataset (ADTTE) with the follow characteristics
  - Treatment codes: trtgrpn
  - Event time: AVAL
  - Indicator variable: CNSR (0 = event, 1 = censor)

```
proc lifetest data=ADTTE;
strata trtgrpn / test=logrank;
time AVAL*CNSR(1);
ods output HomTests=out_lr(keep = ChiSq) HomStats=out2(where=(trtgrpn='1'))
LogRankHomCov=out3(where=(trtgrpn='1') keep=trtgrpn _1 rename=(_1=V));
run;

data out_lr;
set out_lr;
call symput('ChiStat', ChiSq);
run;
%put ChiStat=&ChiStat;
```

Gray's Test

- Comparison of the cumulative incidence probabilities of two or more groups of patients
- It is an adaptation of the log-rank test developed for competing risk data
- Consider a dataset (adam) with the follow characteristics
  - Treatment codes: trtgroup
  - Event time: aval
  - Indicator variable: cnsr (0 = event, 1 = censor, 2 = competing risk)

```
ods select none;
ods output GrayTest=GrayTest;
proc lifetest data=adam  plots=cif;
        strata trtgroup;
        time aval*cnsr(1) / eventcode=0;
run;
ods output close;
```

Agenda

01 Hematologic Malignancies

02 Investigational Product

03 Overview of Clinical Trials Conducted

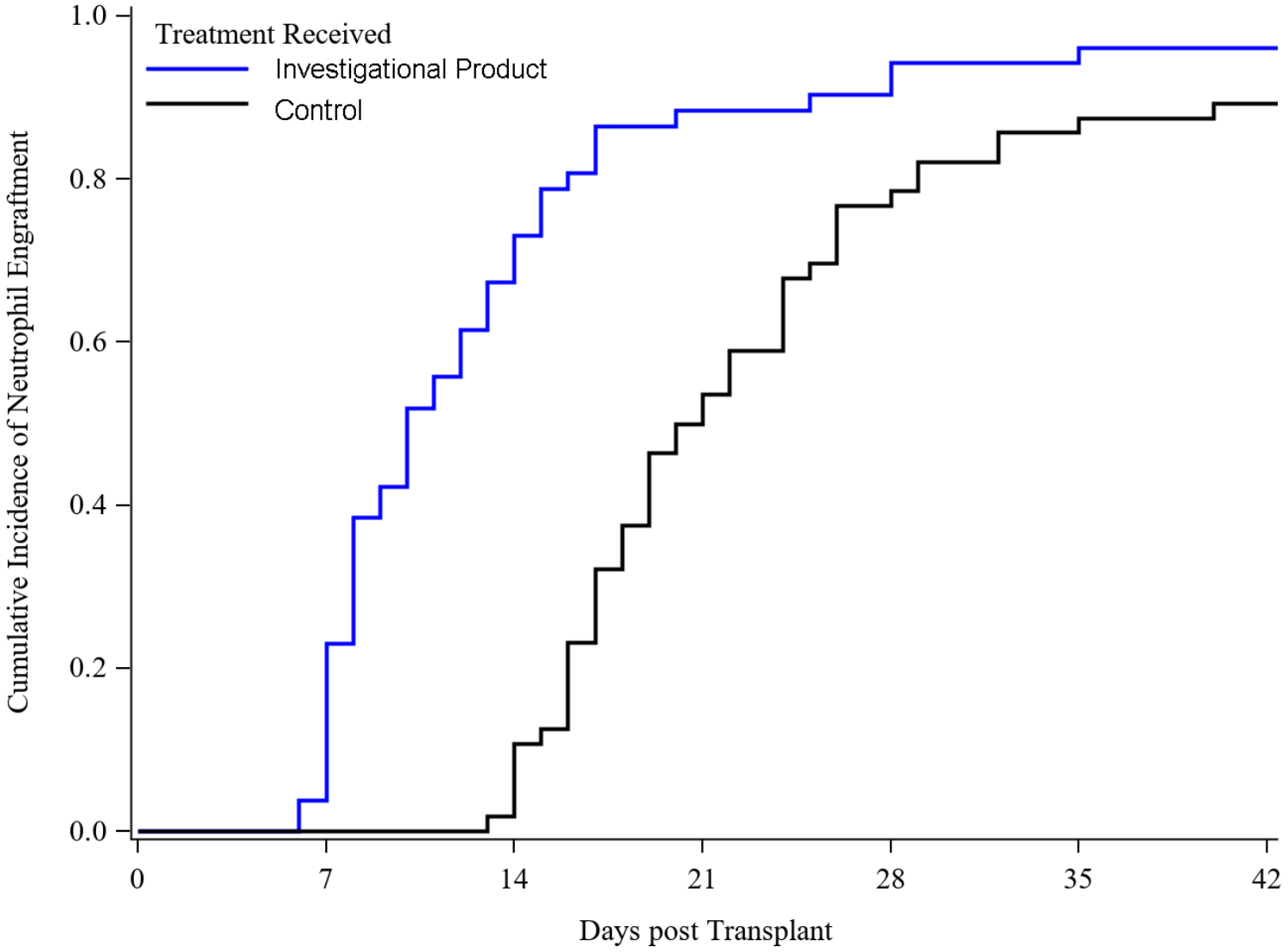
04 Phase III Population, Design, and Endpoints

05 Randomization and Re-randomization

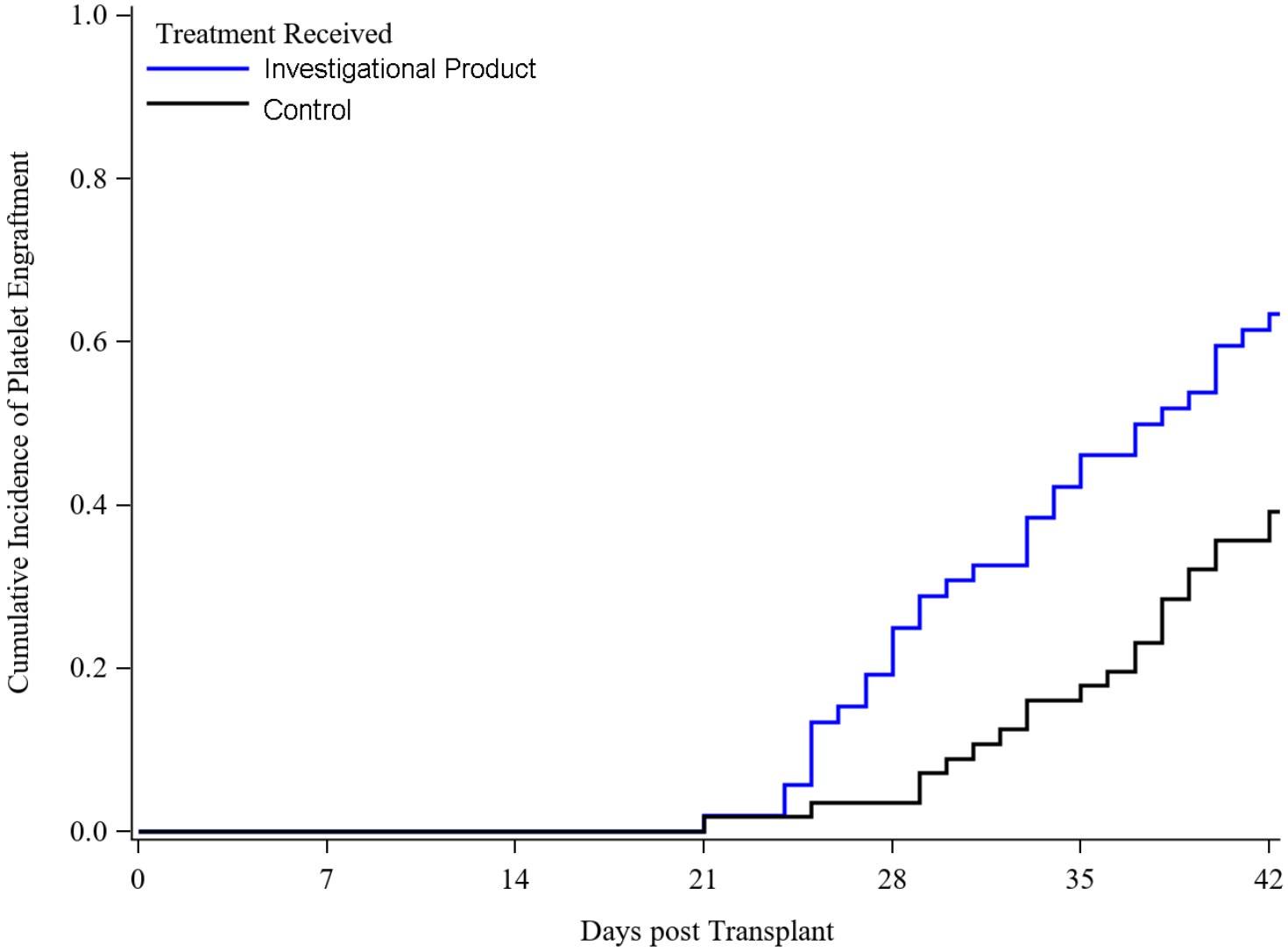
06 Statistical Tests

07 Key Results

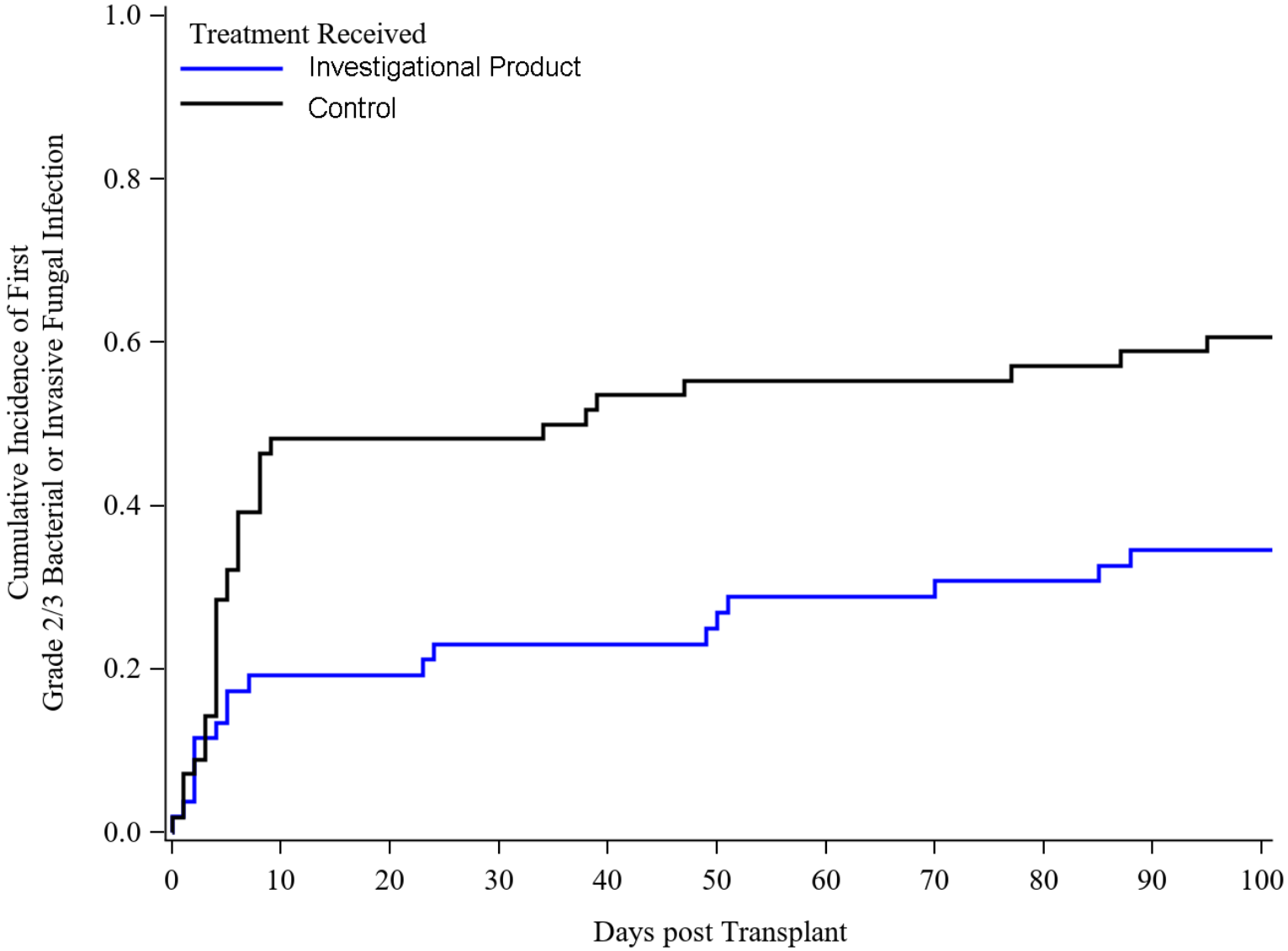
Cumulative Incidence of Neutrophil Engraftment



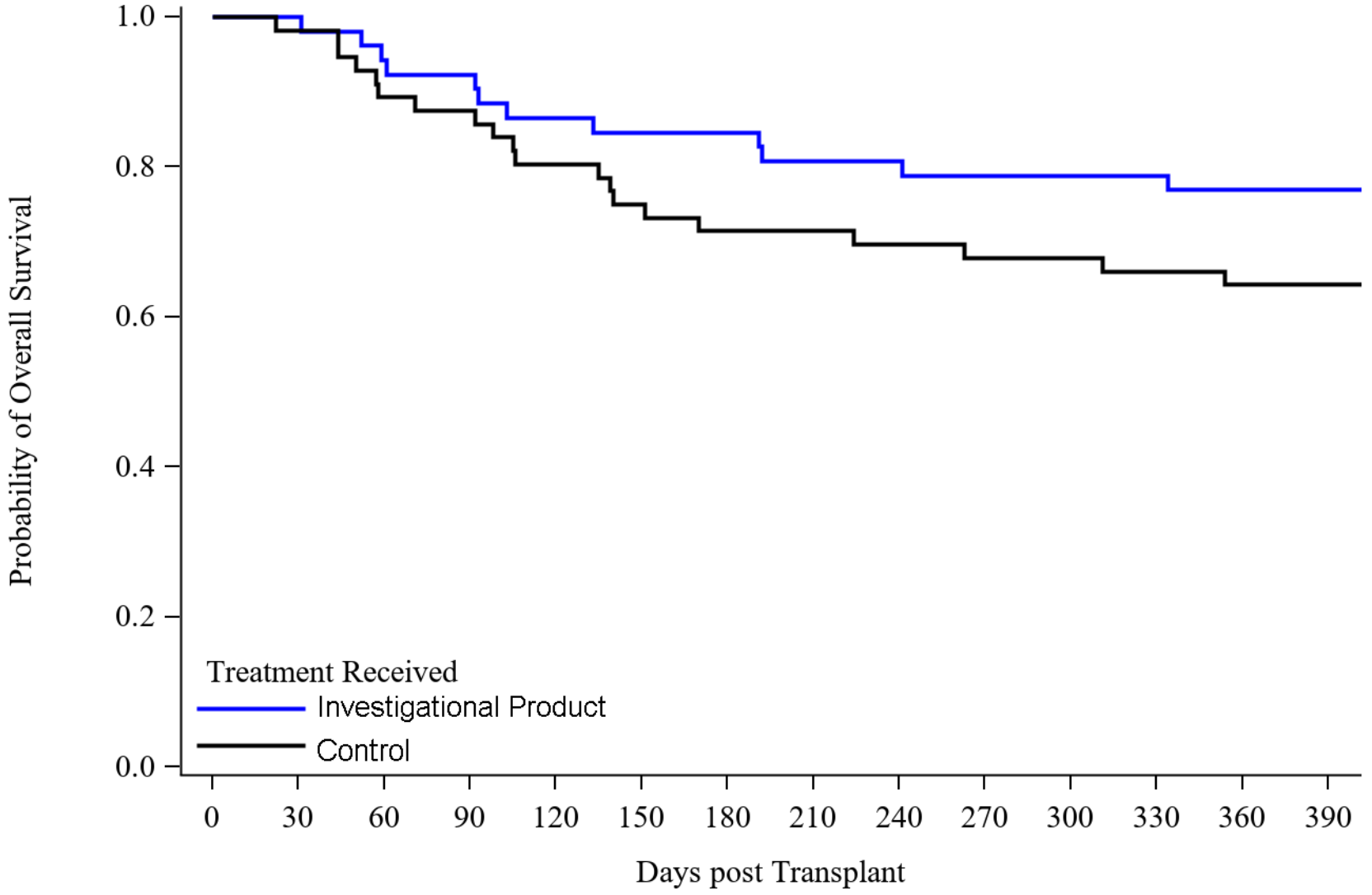
Cumulative Incidence of Platelet Engraftment



Cumulative Incidence of Infections



Overall Survival



Status of Neutrophil Engraftment

		Randomized Treatment Group			
Primary Endpoint Outcome[1]	**Assigned Analysis Day**	**Investigational Product**		**Control**	
		N	**%**	**N**	**%**
Intent-to-Treat population		62	100.0%	63	100.0%
Met definition for neutrophil engraftment on or before Day 42	Time to first day of 3 different days of consecutive measurements of ANC>=0.5 x 10^9/L	55	88.7%	53	84.1%
No neutrophil recovery by (and including) Day 42	43	1	1.6%	4	6.3%
ANC recovery by Day 42 without subsequent donor chimerism <=10% of host cells with applicable chimerism data provided in the Day 100 window or up to the date of relapse, secondary graft failure, or death	43	1	1.6%	0	0.0%
Failure to receive a transplant within 90 days following randomization	43	3	4.8%	4	6.3%
Relapse[2] on or prior to Day 42 without prior neutrophil recovery	43	0	0.0%	0	0.0%
Second transplant on or prior to Day 42 without prior neutrophil recovery	43	2	3.2%	1	1.6%
Death on or prior to Day 42 without prior neutrophil recovery	43	0	0.0%	1	1.6%
Loss to follow up on or prior to Day 42 without prior neutrophil recovery or ANC recovery by Day 42 without subsequent qualifying donor chimerism <=10% of host cells and with applicable chimerism data not available in the Day 100 window	N/A	0	0.0%	0	0.0%

[1]For those with multiple competing risk events, classification of status will be according to the first competing risk event.
[2]Not including relapses that occur before transplant if the transplant is carried out within 90 days of randomization

Test of Time to Neutrophil Engraftment by Day 42 (ITT Population)

				Probability (time to engraftment is shorter in Investigational Product)					
Randomized Treatment Group	N	Median Time to Engraftment[1]	Number Engrafted	P Statistic	Re-randomization 95% Lower CL	Re-randomization 95% Upper CL	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL	Re-randomization *P*-Value
Investigational Product	62	12.0	55	0.77	0.66	0.85	0.68	0.86	<.001
Control	63	22.0	53						

		Cumulative Incidence Analysis							
Randomized Treatment Group	N	Time for Cumulative Incidence of 50% Engraftment (Median Time)	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL	Difference in Median Time Investigational Product vs Control	Rerandomization 95% Lower CL	Rerandomization 95% Upper CL	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL
Investigational Product	62	12	10.0	14.0	-10	-12.9	-6.4	-14.0	-6.0
Control	63	22	19.0	25.0					

Logrank Test Statistic	Logrank Test Statistic Re-randomization *P*-Value	Hazard Ratio	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL
14.50	<.001	2.25	1.42	4.25

Test of Time to Neutrophil Engraftment by Day 42 (AT Population)

		Cumulative Incidence Analysis					
Treatment Received	N	Time for Cumulative Incidence of 50% Engraftment (Median Time)	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL	Difference in Median Time (Investigational Product vs. Control)	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL
Investigational Product	52	10	8.0	13.0	-10	-14.0	-7.0
Control	56	20	18.0	24.0			

				Probability (time to engraftment is shorter in Investigational Product)			
Treatment Received	N	Number Engrafted	Median Time to Engraftment[1]	P Statistic	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL	Wilcoxon Rank Sum Test *P*-Value
Investigational Product	52	50	10.0	0.86	0.78	0.94	<.001
Control	56	50	20.5				

Cumulative Incidence of Platelet Engraftment by Day 100 (AT Population)

			Median Time to Engraftment			Gray's Test of Platelet Engraftment by Day 100
Treatment Received	Number of Participants	Proportion Achieving Platelet Engraftment by Day 100	Cumulative Incidence Median Day	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL	*P*-Value
Investigational Product	52	0.83	37	33.0	42.0	0.023
Control	56	0.73	50	42.0	58.0	

Overall Survival by Day 365 (AT Population)

				Difference in Kaplan-Meier Probability			
Treatment Received	Number of Participants	Number of Deaths by Day 365 post Transplant	Kaplan-Meier Probability of Overall Survival by Day 365 post Transplant	Difference in Probability (Investigational Product vs Control)	Bootstrap 95% Lower CL	Bootstrap 95% Upper CL	*P*-Value
InvestigationalProduct	52	12	0.77	0.13	-0.04	0.29	0.145
Control	56	20	0.64				

Additional Study Information

- Sensitivity analysis was done
  - Worst Rank Assigned
  - Tipping Point

- Study received Orphan designation
- Study received Breakthrough designation
- Study received Priority Review by the FDA following the BLA submission

What does this all mean?

- Patients with no matched related or matched unrelated donor no longer need to rely on mismatched unrelated donor, haplo-related donor, or unmanipulated cord blood transplantations
- No longer a need for double cord blood transplants
- Patients who received the investigational product had a significant reduction in time to neutrophil engraftment
- Patients who received the investigational product recovered platelets faster
- There were fewer infections in patients who received the investigational product
- Although overall survival didn't show statistical significance between the groups, the curves did how potential for better overall survival. We are currently running a long-term follow-up study which includes looking at overall survival
- Faster neutrophil and platelet recovery times in those that received the investigational product, and decreased infection rates is great news for patients and their loved ones. This may allow these patients to have longer times with their families with a higher quality of life. This gives hope to those patients who may not have had other options previously.

REFERENCES & IMAGE SOURCES

1. Noether, G. E. (1987). Sample Size Determination for Some Common Nonparametric Tests. Journal of the American Statistical Association 82(398): 645-647.
2. Klein, J.P., Moeschberger, M.L. Survival Analysis: Techniques for Censored and Truncated Data.
3. https://www.pennmedicine.org/cancer/navigating-cancer-care/programs-and-centers/hematological-malignancies-program
4. https://www.sciencedirect.com/science/article/abs/pii/B9780702047626000990?via%3Dihub
5. https://www.pennmedicine.org/cancer/navigating-cancer-care/programs-and-centers/hematological-malignancies-program
6. https://www.cdc.gov/cancer/uscs/about/data-briefs/no30-hematologic-incidence-surv-prev.htm#:~:text=Hematologic%20cancers%20begin%20in%20the,immune%20system%20that%20fights%20infection.
7. https://www.pennmedicine.org/cancer/navigating-cancer-care/programs-and-centers/hematological-malignancies-program
8. client.wvd.microsoft.com/arm/webclient/index.html
9. https://www.accc-cancer.org/home/learn/cancer-types/hematologic-malignancies
10. https://www.ons.org/sites/default/files/publication_pdfs/HM%20Ch%20%201.pdf
11. https://datatab.net/tutorial/log-rank-test
12. https://pmc.ncbi.nlm.nih.gov/articles/PMC403858/
13. https://www.mcw.edu/-/media/MCW/Departments/Biostatistics/TR070.pdf
14. https://www.linkedin.com/pulse/balancing-gender-equality-equation-meghna-singh
15. https://agri-spareparts.com/blog-post/automatic-milking-systems-advantages-and-disadvantages/
16. https://humboldt-wi.github.io/blog/research/information_systems_1920/group2_survivalanalysis/
17. https://ielts-up.com/writing/advantages-disadvantages-essay.html
18. https://www.freepik.com/free-photos-vectors/challenge

Thank you