Optimal treatment and timing of routine surveillance in children after allogeneic hematopoietic cell transplantaion

Zhezhen Jin

Department of Biostatistics

Mailman School of Public Health

Columbia University

New York, NY 10032

IMA workshop on "Statistical and Computational Challenges in Precision Medicine", University of Minnesota, November 7-9, 2018

1

In biomedical studies:

Study designs: Randomized trials, Case-control, Cohort, Phase I, II, III, IV studies, cross-over study, adaptive design
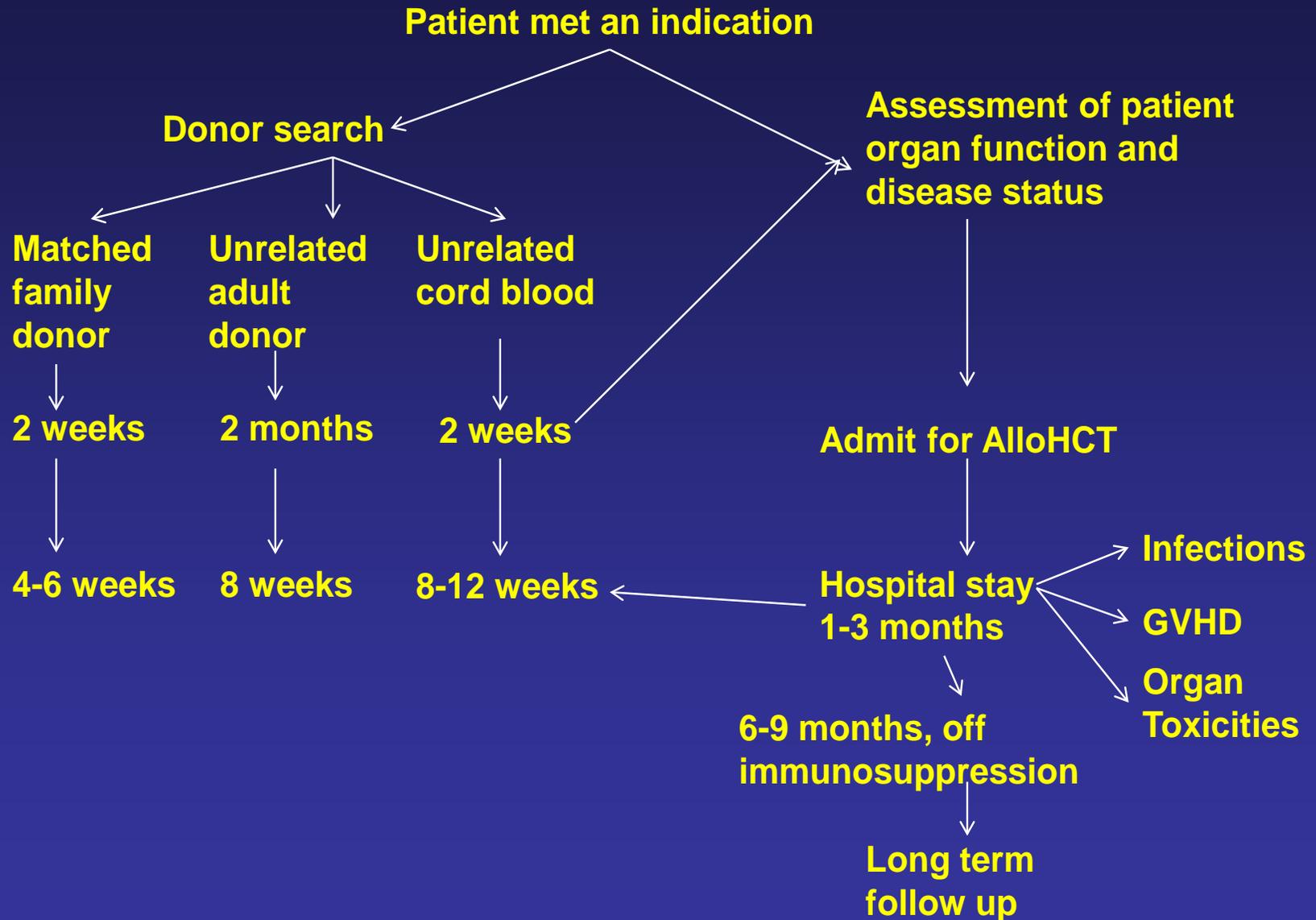
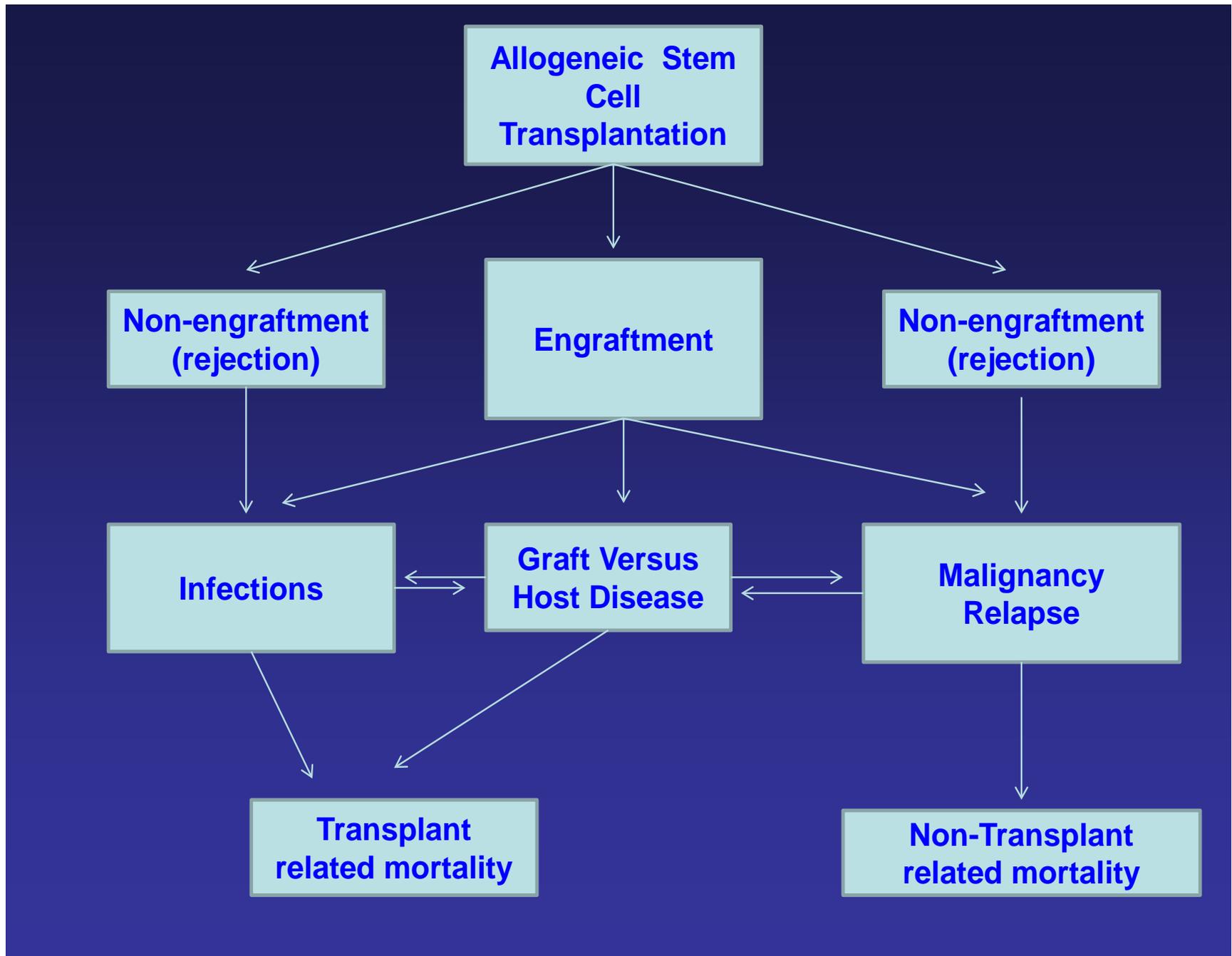Data types: small and large sample sizes, big data (imaging data, -omics data etc.)

Analysis methods

Bone Marrow Transplantation

- Curative for children with malignant and non-malignant disease

- HLA matched sibling donor BMT is associated with best outcomes

Typical Course of a Patient Receiving Allogeneic Stem Cell Transplant

Patient met an indication

Donor search

Assessment of patient organ function and disease status

Matched family donor

Unrelated adult donor

Unrelated cord blood

2 weeks

2 months

2 weeks

Admit for AlloHCT

4-6 weeks

8 weeks

8-12 weeks

Hospital stay 1-3 months

Infections

GVHD

Organ Toxicities

6-9 months, off immunosuppression

Long term follow up

4

Treatment plan and comparison

- **Organ Toxicity:** Veno-occlusive disease; Liver injury; Kidney injury

- **Infection** Viremia; Bacteremia

- **Graft versus host disease:** Acute; Chronic

- **Graft failure**

- **Diagnostic studies:** MRI, CT scans Procedures

- **Care Delivery:** ICU admission

- **Number of inpatient days**

- **Health care cost**

- **Survival:** Overall survival; Relapse

As statisticians, how to adapt to the changing data analysis challenges?

Stick with basic statistical principles and methods (toolbox)

Statistical principles: sufficiency principle, likelihood principle, least squares principle

Statistical methods: descriptive statistics (mean, SD etc), estimation and hypothesis testing (t-test, ANOVA, rank test etc), models (linear and non-linear regression, mixed effects models, additive models, Cox models, logistic models, probit models etc), machine learning (classification, prediction methods),...

parametric methods, nonparametric and semiparametric methods, Bayesian methods, ...

How to apply: art of statistics

Need to good understanding of research questions/problems

Identify pro and cons of various competing statistical methods

Then perform analysis and summary

Most often, complicated statistical methods are difficult to interpret and to explain to non-statisticians

They are often needed for publication

Simple and basic statistical tools are easy to interpret

How to use those simple and basic statistical tools is "**art**"

Every student of the art of data analysis repeatedly needs to build upon his previous statistical knowledge and to reform that foundation through fresh insights and emphases, Mosteller and Tukey (1977, *Data analysis and Regression*).

Exploratory data analysis and confirmative data analysis.

- On the treatment on acute graft-versus-host disease

- On bone marrow harvest from pediatric sibling donors

- On the utility and optimal timing of routine bone marrow and cerebrospinal fluid surveillance

Biology of Blood and
Marrow Transplantation

ASBMT
American Society for Blood
and Marrow Transplantation

journal homepage: www.bbmt.org

ELSEVIER

Efficacy of Tacrolimus/Mycophenolate Mofetil as Acute Graft-Versus-Host Disease Prophylaxis and the Impact of Subtherapeutic Tacrolimus Levels in Children after Matched Sibling Donor Allogeneic Hematopoietic Cell Transplantation

CrossMark

Katharine Offer [1], Michelle Kolb [2], Zhezhen Jin [3], Monica Bhatia [1], Andrew L. Kung [1], Diane George [1], James H. Garvin [1], Chalitha Robinson [1], Jean Sosna [1], Esra Karamehmet [1], Prakash Satwani [1,*]

On the treatment on acute graft-versus-host disease

- Donor immune cells reacts to patients body, this phenomenon is called graft versus host disease.

- Incidence of 20-60%.

- Can damage skin, intestines, liver, mucosa and gastrointestinal tract.

- Treatment of GVHD with steroids can lead to severe infections.

- Prevention of GVHD might be best strategy.

Tacrolimus and mycophenolate: GVHD prevention for Matched Sibling donor Transplants

- Tacrolimus (Tacro) and mycophenolate (MMF) are utilized at Columbia for GVHD prevention.

- Best therapeutic range for tacrolimus in children not well described.

- High levels can lead to kidney injury and low levels can lead to GVHD.

- Our practice was to keep levels between 15-20 ng/mL which based on adult data.

- To examine the currently accepted target range of 15-20 ng/mL for Tacrolimus, by measuring the association between whole blood levels <10 ng/mL during the first four weeks post-AlloHCT, and the cumulative incidence of grade II-IV aGVHD in children.

- To identify lower value for whole blood Tacrolimus range at various time points after AlloHCT that may be more appropriate targets for prophylactic dosing.

- We hypothesized that subtherapeutic Tacrolimus levels were associated with a higher incidence of aGVHD.

- Retrospective analysis evaluating the efficacy of prophylactic Tacro/MMF and the competing risk factors that may impact the occurrence of grade II-IV aGVHD in pediatric patients undergoing matched sibling donor AlloHCT.

Data collection

Medical record: age, gender, ethnicity, nonmalignant/malignant diseases, HLA matching, donor source, conditioning regimen, and CMV status.

Treatments: r-ATG or alemtuzumab and concurrent medications

All tacrolimus and MMF levels through day +100 were collected, as were levels of creatinine and bilirubin on days 0, 7, 14, 21, and 28.

Outcome: Acute GVHD, Renal injury, Liver toxicity, survival

The number of days during which each patient had subtherapeutic tacrolimus levels (<10 ng/mL) were counted. MMF levels during weeks 1 to 4 post-transplant were collected.

Sixty patients from March 2003 to September 2012.

Repeated measures

Time to event (acute GVHD/ renal injury/ liver toxicity)

Statistical analysis: competing risk analysis, ROC analysis, logistic regression analysis

For a binary outcome D, a cut point on continuous biomarker X helps classify subject into one of the two classes ($D = 0,\ 1$).
— Let $X^{(D)}$ be the measure from group D with distribution functions $F_D(x)$.
— Suppose that group ($D = 1$) tends to have smaller X.

Sensitivity - probability of correctly classifying subjects in $D = 1$ group with cut point x, $\quad Se(x) = P(X^{(D=1)} \leq x) = F_1(x)$.

Specificity - probability of correctly classifying subjects in $D = 0$ group with cut point x, $\quad Sp(x) = P(X^{(D=0)} > x) = 1 - F_0(x)$.

Receiver Operational Characteristic (ROC) curve
Plot of $Se(x)$ vs. $1 - Sp(x)$ for all possible cut points x.

ROC curve related criteria for cut point selection

Let W be the set of possible cut points.

Youden index
$J(x) = Se(x) + Sp(x) - 1 = S_1(x) - S_0(x)$ takes value in $[-1, 1]$.

Selected cut point $x_J = \max_{x \in W} J(x)$.

Closest-to-(0,1) criterion
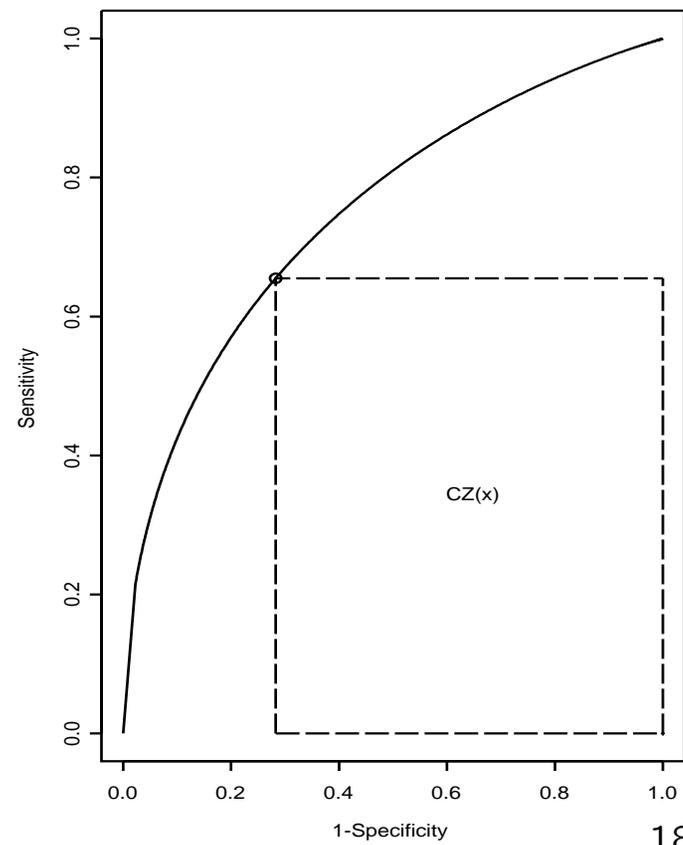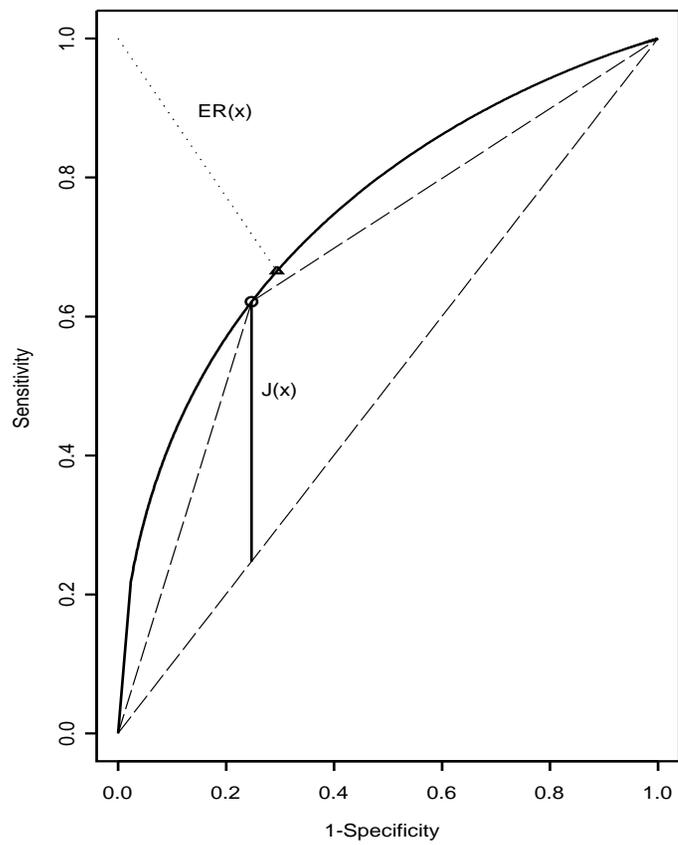$ER(x) = \sqrt{(1 - Sp(x))^2 + (1 - Se(x))^2}$ takes value in $[0, 1]$.

Selected cut point $x_E = \min_{x \in W} ER(x)$.

Concordance criterion
$C(x) = P(X^{(1)} \leq x < X^{(0)}) = Se(x)Sp(x)$ takes value in $[0, 1]$.

Select cut point $x_C = \max_{x \in W} C(x)$.

Figure 1. ROC curve related criteria

The three criteria may not yield the same optimal cut-point.

The cut-point optimizing either Youden index or Closest-to-(0,1) criterion may not optimize the concordance probability for binary classification.

The common cut-point optimizing the three criteria exists in some spacial cases. When $X^{(D)} \sim N(\mu_D, \ \sigma^2)$, the three criteria will have the same optimum, but variation in $\widehat{X}_C$ is much smaller than that of $\widehat{X}_J$, while slightly larger than that of $\widehat{X}_E$.

In practice, it is important to check all three criteria.

Time dependent ROC curve

Disease status at t: $D(t) = 0$ for no disease at t and $D(t) = 1$ for disease.

Time dependent sensitivity: $Se(x,t) = P(X \leq x | D(t) = 1)$,

Time dependent specificity: $Sp(x,t) = P(X > x | D(t) = 0)$.
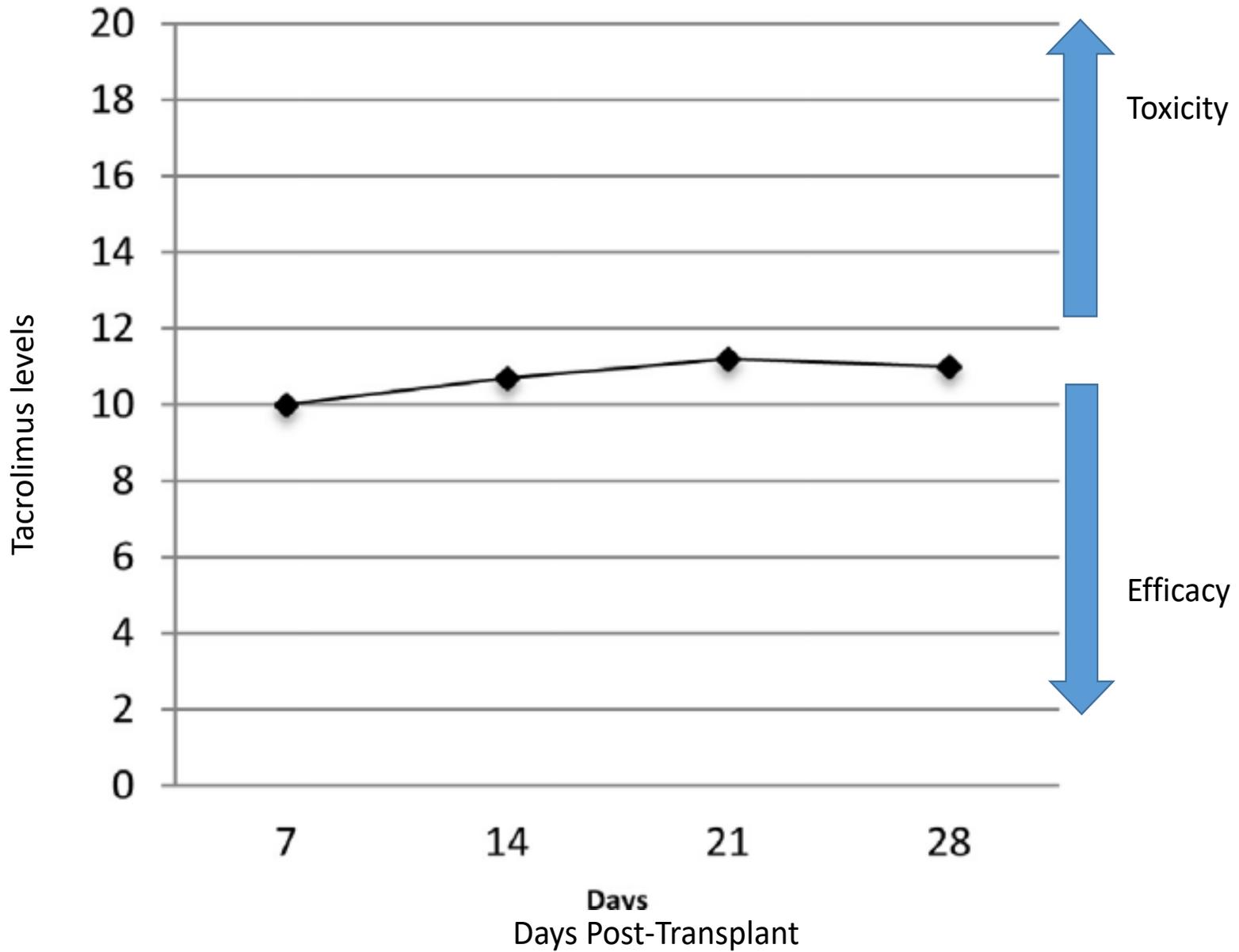
Then, the three time-dependent

$$J(x,t) = Se(x,t) + Sp(x,t) - 1$$

$$ER(x,t) = \sqrt{(1 - Sp(x,t))^2 + (1 - Se(x,t))^2}$$

$$Cz(x,t) = P(X^{(D(t)=0)} \leq x < X^{(D(t)=1)}) = Se(x,t)Sp(x,t).$$

cut-point: maximize $\int_{t_1}^{t_2} W_1(t) J(x,t) dt$, $\int_{t_1}^{t_2} W_3(t) Cz(x,t) dt$ and minimize $\int_{t_1}^{t_2} W_2(t) ER(x,t) dt$

Conclusion

- Optimal level for tacrolimus during first 4 weeks is between 10-12.

- Sub-therapeutic levels during 3rd week was associated with higher incidence of aGVHD.

- Practice was changed at Columbia to keep levels between 10-15 from 15-20 since 2015.

Pediatric

Bone Marrow Harvest in Pediatric Sibling Donors: Role of Granulocyte Colony-Stimulating Factor Priming and CD34+ Cell Dose

Aimee Furey [1], Sonal Rastogi [2], Remi Prince [3], Zhezhen Jin [2], Elana Smilow [1], Courtney Briamonte [1], Justine M. Kahn [3], Yvette Tanhehco [4], Nita Patel [4], Diane George [3], James Garvin [3], Monica Bhatia [3], Prakash Satwani [3,*]

[1] Department of Nursing, New York-Presbyterian Morgan Stanley Children's Hospital, New York, New York
[2] Mailman School of Public Health, Columbia University, New York, New York
[3] Department of Pediatrics, Columbia University, New York, New York
[4] Department of Pathology, Columbia University, New York, New York

On bone marrow harvest from pediatric sibling donors

- To ensure optimal clinical outcomes with adequate protection for donors, the National Marrow Donor Program (NMDP) developed guidelines stating that up to 20 mL/kg of bone marrow can safely be harvested from donors.

- These guidelines, developed originally for unrelated adult donors, are followed for children as well

- Children, on average, have more robust bone marrow than adults do. In light of this, harvesting fewer cells from young donors could theoretically yield an adequate stem cells.

- Smaller harvest would be of benefit to donors specifically because it could decrease time spent under general anesthesia, reduce overall blood loss, decrease post-procedure pain and expedite the recovery period.

Data collection

The Columbia University Medical Center

All 5-6/6 HLA-matched sibling bone marrow harvests done between January 2005 and July 2017 were included in the study, and for clinical outcomes, all 6/6 HLA-matched sibling bone marrow transplantations performed between January 2005 and July 2016 were included. Sibling donors who were 5/6 HLA-matched were excluded.

Bone Marrow Harvest Procedure/Bone Marrow Processing (CD34$^+$, CD3$^+$)

There were 92 consecutive bone marrow harvests (82 from 6/6 HLA matched and 10 from 5/6 HLA-matched sibling donors) and 69 recipients.

Medical record: age, gender, ethnicity, nonmalignant/malignant diseases, HLA matching, donor source, conditioning regimen, CMV status, infections, length of hospital stay.

Treatments: r-ATG or alemtuzumab and concurrent medications

Outcome: A composite index known as the GRFS (GVHD-free/relapse-free survival) which has recently been used to assess the true success of allo-HCT instead of the patient's transplant-related outcome. The GRFS is determined by both early and late outcomes and includes events such as grade III-IV acute GVHD, chronic GVHD requiring systemic therapy, relapse, or death in the first year after allo-HCT.

Statistical methods Goal of this study: to find a CD34$^+$ cell dose associated with GRFS, and to conceptualize a personalized prescription for bone marrow harvest among pediatric donors.
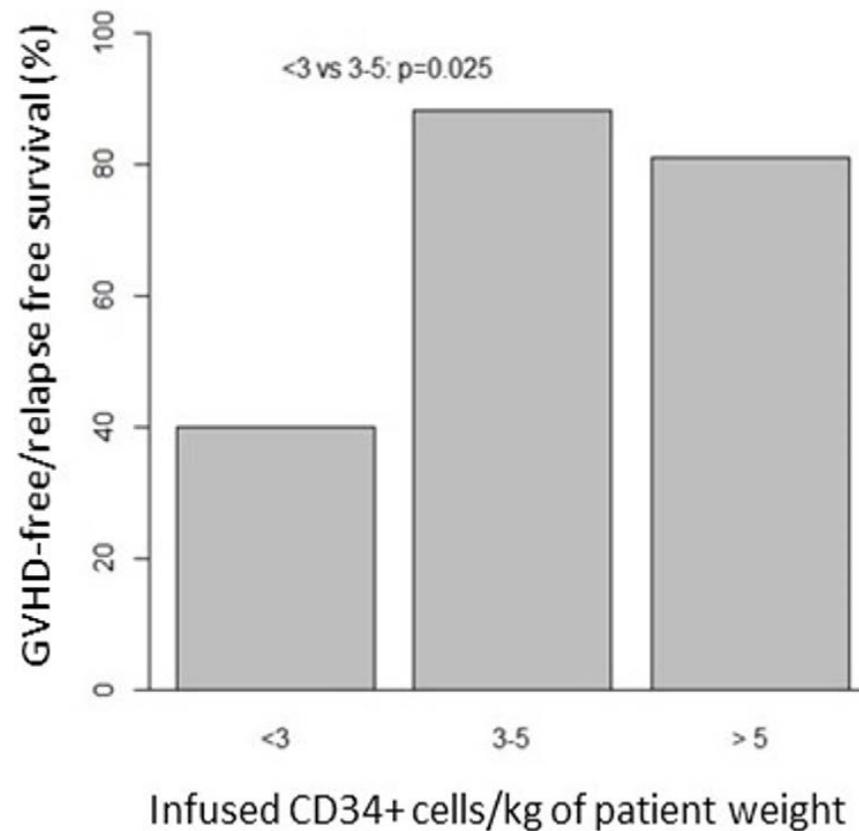
Wilcoxon rank-sum test

Chi-square test or Fisher's exact test

Pearson's correlation coefficient

Logistic regression analysis

Receiver operating characteristic curve analysis for the optimal cutoff of CD34$^+$ cells

Infused CD34: What is the best cell dose after matched sibling transplant

Bone Marrow Harvest
Custom Prescription Based on Donor Age

Donor age (years)	Median CD34+ count/ microliter	Recipient weight	Donor weight	Calculation for harvesting **5 x 10^6/kg** of CD34+ cells based on recipient weight (ml/kg of donor weight)	NMDP guideline bone marrow harvest up to 20ml/kg of donor weight	Our algorithm (bone marrow harvest/kg of donor weight)
<6	700	40kg	25kg	(patient weight x5)/ (donor weight x 0.7)	500ml	285ml (11.4ml/kg)
6-12	360	40kg	40kg	(patient weight x5)/ (donor weight x 0.36)	800ml	555ml (13.9 ml/kg)
>12	300	40kg	50kg	(patient weight x5)/ (donor weight x 0.3)	1000ml	833 ml (16.6 ml/kg)

Timing and Utility of Relapse Surveillance after Allogeneic Hematopoietic Cell Transplantation in Children with Leukemia

Anya Levinson [1], Staci Arnold [2], Zhezhen Jin [3], Monica Bhatia [1], Diane George [1], James H. Garvin [1], Prakash Satwani [1,*]

[1] Department of Pediatrics, Columbia University, New York, New York
[2] Department of Pediatrics, Emory University, Atlanta, Georgia
[3] Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York

No published prospective or retrospective studies investigated the optimal frequency or timing of routine surveillance with either bone marrow (BM) or cerebrospinal fluid (CSF) analyses in children after alloHCT.

Patients may be undergoing unnecessary, harmful procedures or, conversely, experiencing delayed diagnosis of relapse

The frequency or timing of routine surveillance would be physician-dependent and varies across centers.

A study was designed both to characterize the existing practice nationwide and to gain some preliminary insight into the utility of post-alloHCT surveillance based on data from our center.

- **National survey**: a 14-question e-mail survey via Surveymonkey to 152 pediatric BM transplant physicians nationwide, identified from the Children's Oncology Group members' roster.

- **Single-Center Review (Columbia)**: retrospective review on children who underwent alloHCT for leukemia (ALL, AML) at New York-Presbyterian Morgan Stanley Children's Hospital between 2000 and 2012 (n=108). Baseline data collected included disease, sex, age, and time to relapse. Surveillance data were collected, at approximately days 100, 180, 270, and 365, as per our institutional protocol.

Survey Questions:

1. How many pediatric allogeneic hematopoietic cell transplants (AlloHCT) are performed at your center/year?

2. Do you perform routine bone marrow analyses after AlloHCT in all patients with acute lymphoblastic leukemia (ALL) who are not enrolled on a research study?

3. If you perform routine post-AlloHCT bone marrow analyses in patients with ALL who are not enrolled in a research study, what is the preferred schedule?

4. Do you perform routine bone marrow analyses after AlloHCT in all patients with acute myeloid leukemia (AML) who are not enrolled on a research study?

5. If you perform routine post-AlloHCT bone marrow analyses in patients with AML who are not enrolled in a research study, what is the preferred schedule?

6. After the first year post-AlloHCT, do you perform yearly routine bone marrow analyses?

7. If yes, for how many years?

8. Do you perform routine lumbar punctures after AlloHCT in all patients with ALL?

9. If you perform routine lumbar punctures in patients with ALL, what is the preferred schedule?

10. Do you routinely administer intrathecal chemotherapy after AlloHCT in patients with ALL?

11. Do you perform routine lumbar punctures after AlloHCT in patients with AML?

12. If you perform routine lumbar punctures in patients with AML, what is the preferred schedule?

13. Do you routinely administer intrathecal chemotherapy after AlloHCT in patients with AML?

14. Would you be interested in collaborating with our group on future studies regarding routine 1-year procedures following AlloHCT?

Survey Results
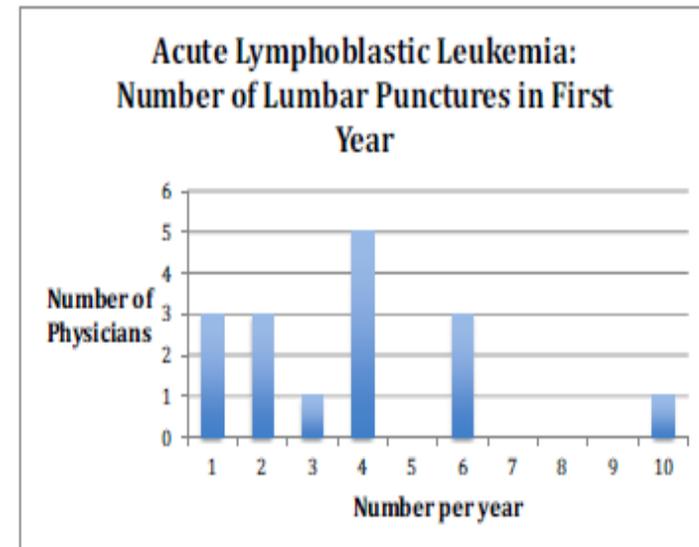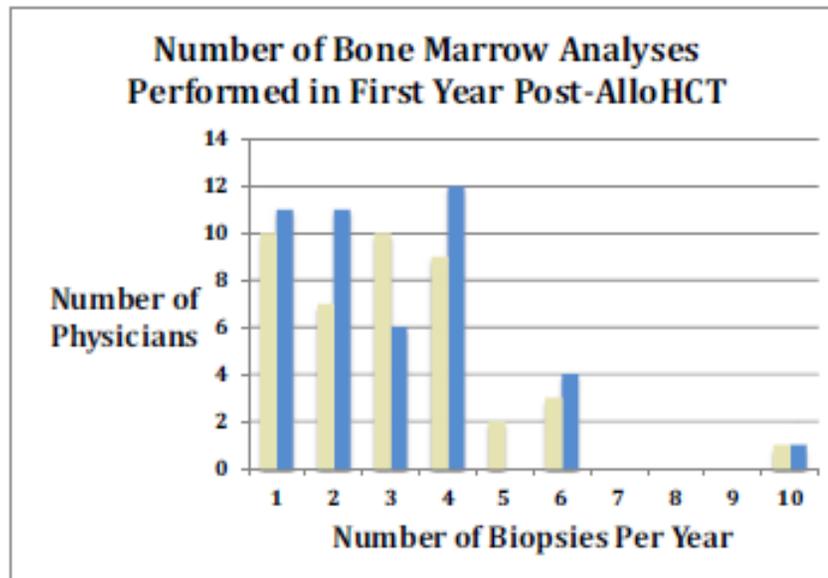
(A) Histogram of the number of physicians who perform between 1 and 10 BM analyses in the first year post-alloHCT (blue bars indicate ALL; yellow bars indicate AML).

(B) Histogram of the number of physicians who perform between 1 and 10 lumbar punctures in the first year post-alloHCT in ALL patients.

(C) At our center, we performed routine Bone Marrow Aspirate at biopsy and spinal tap on: Day +100, Day +180, Day +270 and Day +365 post AlloHCT.

Cumulative probability of relapse as a function of time for children with leukemia at Columbia
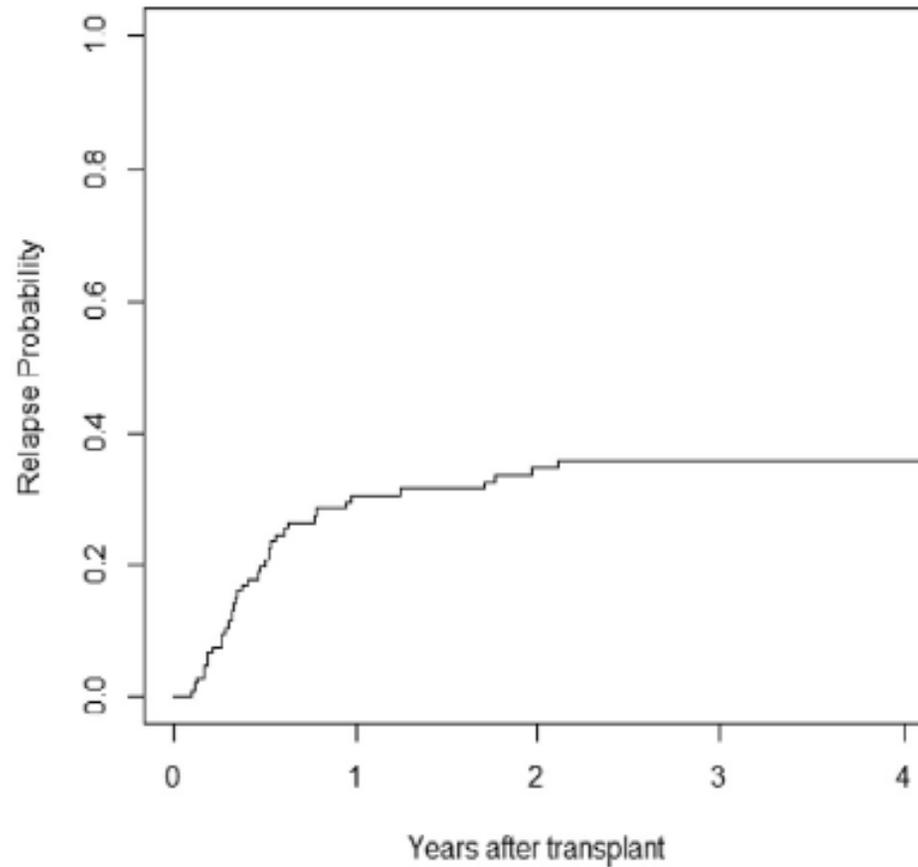
Results

- At Columbia, routine surveillance at days 100, 180, and 270 detected 36 relapses, representing 33% of relapses in the first year.

- These relapses were all identified before the patients presented clinically with disease.

- This finding may be important for the non-trivial percentage of practitioners who do no routine surveillance in the first year (10% and 15%, respectively, for ALL and AML patients).

- However, median survival was not meaningfully increased for those in whom relapse was detected at the subclinical stage.

- BM and CSF surveillance could be ceased by 1 year after alloHCT in children with leukemia. However, more study is needed.

Summary and Discussion

Three specific examples which yield precision and personalized treatments are presented

Real data analysis sometimes is challenging

Need to understand scientific questions and exploratory data analysis is necessary

Integrated data analysis is often needed.

Existing statistical can be used to address scientific questions

New development is often based on existing methods

Acknowledgement

Dr. Prakash Satawani

Dr. Monica Bhatia

Dr. Diane George

Dr. James H. Garvin

Dr. Staci Arnold