*Impact of varying maintenance strategies post – Autologous stem cell transplantation in standard and high-risk multiple myeloma patients: a single-centre Canadian retrospective study*

Methods: Statistical Analysis

Table of Contents

[Primary Objective/Outcome Measure 3](#_Toc72419419)

[Secondary Objectives 4](#_Toc72419420)

[Multivariate Cox proportional-hazards model 8](#_Toc72419421)

[Competing Risk Analysis 9](#_Toc72419422)

[Theory of Statistical tests/Application in R 10](#_Toc72419423)

# 

# Primary Objective/Outcome Measure

*Primary objective*: The primary objective of this study is to compare the median PFS for patients with TE high risk multiple myeloma receiving maintenance treatment post ASCT with Lenalidomide + PI versus single agent maintenance/no maintenance.

*Primary outcome measure*: The primary outcome is the median PFS in high-risk patients post-ASCT receiving maintenance treatment with Lenalidomide + PI versus other strategies (no maintenance, single agent maintenance).

Notes:

* High risk patients (MM) are defined by presence of one of t (4;14), t (14;16), t (14;20), del(17/17p), (1q) amplification identified by fluorescence in situ hybridization (FISH) and patients will ISS stage 3 disease (5,6)
* Progression-free survival (PFS) is defined as the time from diagnosis until either the first documentation of disease progression or death from any cause. In terms of the excel data, this refers to the time between diagnosis data (column E on “ASCT Study May 14 version”) and relapse 1 (column CI on “ASCT Study May 14 version”).
* Question: If they don’t relapse and hence are censored, should their final follow up date be their end date? For some patients, this end date is either not clear/not given (i.e., column F is a good start, and other columns too, but they tend to be mostly empty).
* Answer: Yes, should be last follow up.
* In terms of post maintenance drugs: Lenalidomide (= R= Revlimide) + PI (protease inhibitor, V=velcade=bortezemib. K= carfilzomib I= Ixazomib)
* Patients who received thalidomide maintenance will be excluded from the study.
* Question: Should results for low-risk patients be investigated too?
* Answer: Yes.

Statistical tests:

* Establish dataset of patients that are all high risk, and use their “status” (i.e., censored or “event of interest”) and the time of this status to establish Kaplan-Meier curves, where the stratification factor is maintenance treatment, producing 3 curves corresponding to: L + PI maintenance treatment, single agent maintenance treatment, and no maintenance treatment. (Categorical representation of 2, 1, 0 respectively)
* Perform log rank test to establish significance between these three curves survival curves. (Question: do one log rank test for all 3 curves, or perform log rank test for the 3 possible comparisons between groups of two curves?)
* Answer: One log rank test.
* For low risk, shouldn’t find that double agents are better than one agent.

# 

# Secondary Objectives

*Secondary objectives*: The secondary objectives include evaluating the outcomes of lenalidomide maintenance with and without protease inhibitor, evaluating the outcomes of different doses of Lenalidomide maintenance and effects of tandem versus single ASCT as detailed below:

*i)* Median OS

Notes:

* OS (overall survival) - time between diagnosis and death from any cause or last follow up (censored).
* In chart data: time of transplant (column: AO corresponds to transplant #1 date), last follow-up/death (column: YY) Question: Should time of transplant correspond to transplant #1, or if tandem, correspond to transplant #2? Also, same issue as in primary objective in determining last follow-up/death.
* Answer: It’s time from diagnosis, not transplant.

Statistical tests:

* Using the OS time data and the status at this time (i.e., either censored or event of interest), can establish a Kaplan Meier curve which determines the Median OS for all patients involved in the study.
* Question: should this be just for high-risk patients?
* Answer: For low risk as well.

*ii)* The ORR and depth of response based on international myeloma working group criteria (IMWG) in the following pre-specified periods: Post induction, 1 to 3 months post ASCT (if tandem, 1 to 3 months post tandem as well), 12 months post ASCT (for those who did not receive maintenance therapy); or 12 months after initiation of maintenance therapy.

Notes:

* ORR is the overall response rate, which corresponds to the data like VGPR, CR etc. It is indicated in columns: AG (post induction), AW (1–3-month post ASCT 1), BI (1–3-month post ASCT 2), XX (12-month post ASCT, for those who didn’t receive maintenance therapy), YY (12 months after initiation of maintenance therapy). Question: Were the right columns chosen, and what are the missing columns?
* Answer: Not necessarily AW, the correct data is being collected. Not necessarily BI, and XX and YY all still being collected.
* Question: What does depth of response correspond to? Or is it the same as ORR? And what columns?
* Answer: Same as ORR

Statistical tests:

* For each pre-specified period where data is collected, can determine the amount of each type of ORR response, i.e., a bar graph (also amount of depth of response if this is different from ORR)
* Question: should this be just for high-risk patients
* Answer: And for low risk as well.
* Can do these graphs on excel.

*iii)* Subgroup analysis to assess ORR, median OS and median PFS in high-risk patients according to their maintenance strategy.

Notes:

* High risk patients only
* Median PFS vs. maintenance strategy already completed in primary objective
* Question: Maintenance strategy post-transplant or after one of the relapses?
* Answer: Post-transplant

Statistical Tests:

* For OS. Vs. maintenance strategy, follow primary objective procedure identically, except the time-to-event of interest is OS, instead of PFS.
* For ORR vs. maintenance strategy: Only use ORR data corresponding to 12 months post ASCT (for those who did not receive maintenance therapy); and 12 months after initiation of maintenance therapy (for those who did receive maintenance therapy). Will have 3 groups corresponding to the 3 maintenance strategies (0, 1, or 2 drugs), and compare their relative rates of ORR. I.e., should produce a contingency table with ORR (VGPR, PR etc.) versus maintenance strategy (0, 1, 2). Can use a chi-square test to explore if there is a significant relationship between the type of maintenance strategy, and the ORR.

*iv)* To assess the median lenalidomide maintenance dose

Tau squared test for heterogeneity

Notes:

* Question: High risk patients only?
* Answer: and Low risk patients
* Question: LI Maintenance dose post-transplant or after one of the relapses?
* Answer: Post-transplant

Statistical Tests:

* For only the patients who underwent maintenance treatment with Lenalidomide (exclude people that had two agent or zero agent treatment), obtain their dosages and determine the median. (Answer is likely 10 mg. That’s the go to dose, while 5mg is the go-to reduced dose, which doesn’t happen more often)
* Question: Is my above analysis, okay?
* Answer: Yes. But maybe also check for 2 agent , since they will still probably get 10mg.
* Question: Don’t think a survival table is needed here?
* Answer: No.

*v)* To assess the rate of lenalidomide dose reduction and discontinuation.

Notes:

* Question: High risk patients only?
* Answer: And low risk as well.
* Info on dose reduction in columns BV (indicates if dose was reduced) and CD (indicates what it was reduced too, or if was discontinued)
* Looks like reduction happened more often, and that discontinuation occurred only few times, and occurred because of extenuating circumstances.
* Question: Maintenance strategy post-transplant or after one of the relapses?
* Answer: Post-transplant

Statistical Tests:

* This refers to the proportion (percentage) of people who received lenalidomide maintenance treatment and had their doses either reduced or discontinued.
* I.e. (Li dose reduction = # of individuals on LI treatment that had their dose reduced / total # of people on LI; Li dose discontinuation = # of individuals on LI treatment that had their dose discontinued / total # of people on LI).

*vi)* To assess Impact of lenalidomide maintenance treatment at full doses versus reduced doses in OS.

Notes:

* Question: High risk patients only?
* Answer: And low risk patients
* Question: Maintenance strategy post-transplant or after one of the relapses?
* Answer: post-transplant
* OS (overall survival)- time between diagnosis and death from any cause or last follow up.
* Columns BV (indicates if dose was reduced) and CD (indicates what it was reduced too, or if was discontinued)
* Basically, normal dose = 10mg and reduced dose = 5mg.
* Question: Should only look at individuals that were given exclusively Li? (i.e., excluding people who had Li + PI or had no maintenance treatments). Otherwise, effects of PI’s or other agents may affect OS extraneously
* Answer: Can do both.
* Something to note: after determining the correlation between LI maintenance dose and OS, it may be difficult to establish or suggest causation of the results because of a certain factor:

1. It’s possible the reason that doses get reduced in the first place is because the patient is responding well to treatment; hence it isn’t the reduced dose of LI that maybe changes OS, but rather the circumstances that lead to LI being reduced in the first place that changes OS.
2. Overall, determining whether the reduced dose or the circumstance leading to the reduced dose, correlates with OS, is not trivial.

* It’s only ever reduced if the patient has side effect, if patient is responding, they keep going.

Statistical Tests:

* Establish dataset with individuals that were given exclusively LI (i.e., excluding people who had Li + PI or had no maintenance treatments). These individuals can be separated into two categories: full dose or reduced dose, since only the doses of 10mg and 5mg were used, respectively. This will be the stratification used in the Kepler Meier curve.
* The time-to-event metric is OS,
* The log rank test will be used to establish significance between the 5mg vs. 10mg dose survival curves.

*vii)* To assessImpact of maintenance approach in PFS2

Notes:

* Question: High risk patients only?
* Answer: and low risk patients
* PFS2 - time from diagnosis to second relapse (= initiation of 3rd line of therapy)
* Time of diagnosis is column E, Initiation of 3rd line therapy (column DS). Question: Were the right columns chosen?
* Answer: Yes
* “Impact of Maintenance” – Lenalidomide + PI versus single agent maintenance versus no maintenance.
* Question: Maintenance strategy post-transplant or after one of the relapses?
* Answer: Post-transplant

Statistical Tests:

* For PFS2. Vs. maintenance strategy, follow primary objective procedure identically, except the time-to-event of interest is PFS2, instead of PFS.

*viii)* To assess impact of single versus tandem ASCT on ORR, PFS and OS in patients with high-risk multiple myeloma.

Notes:

* High risk patients only.
* Single (1) vs. tandem (2) ASCT

Statistical Tests:

* Impact of Single (1) vs. tandem (2) ASCT on ORR. The ORR data is the 12 months post ASCT (for those who did not receive maintenance therapy), i.e., excluding those who received maintenance treatment, in order for maintenance treatment to not be an exogenous factor. The ORR data could also be the 1–3-month post ASCT (if tandem, 1 to 3 months post tandem as well). If those who received maintenance treatment aren’t excluded, the ORR data could be 12 months post ASCT (for those who did not receive maintenance therapy); or 12 months after initiation of maintenance therapy.
* Idea is to establish a contingency table and check the chi-square statistic to establish significance.
* Question: Is excluding those who underwent maintenance therapy the right idea here? Same question holds for the two below analysis. Thinking maybe not, but it may be the best idea to apply multivariate analysis here, so the effect of maintenance therapy gets averaged out.
* Answer: 12 months past only single ASCT or tandem ASCT.
* Impact of Single (1) vs. tandem (2) ASCT on PFS. Kepler Meier curve analysis, log rank to establish significance.
* Impact of Single (1) vs. tandem (2) ASCT on OS. Kepler Meier curve analysis, log rank to establish significance.

# Multivariate Cox proportional-hazards model

Stratification Factors:

* Question: High risk patients only?
* Answer: and for low risk as well.
* Question: What should the stratification factors be that are implemented into the model?
* Answer: Maintenance, ORR (patients that had VGPR, vs.), age, LDH (normal vs. abnormal, 225 is the cut-off of abnormality), and ISS as markers (ISS 1 + 2 vs. 3), LI dosage, single vs. tandem, creatinine, gender, HB levels, type of monoclonal chain (IGG, IGA, IGM, IGD)( kappa, lambda), (Maybe look to group together light chains vs heavy chains diseases?), age
* Ideas listed below:

1. Maintenance treatment: LI + PI maintenance treatment (2), single agent maintenance treatment (1), and no maintenance treatment (0).
2. ORR? (If yes, when? Post-ACST?) (don’t think this one makes sense)
3. LI dosage, i.e., either no dose, reduced dose or full dose (0mg (0), 5mg (1) or 10mg (2) respectively). If the 0, 1, 2 classifications are assigned, I’ll be treating LI dosage more categorically, rather than if the exact values (0, 5, 10) are used, where they’ll be treated more quantitatively.
4. Single (1) vs. tandem (2) ASCT
5. Age, gender, blood levels, drug type etc.

Statistical analysis:

The time-to-event will be PFS and OS (each analyzed separately). I.e., both PFS and OS will have their own set of hazard ratios with respect to the stratification factors.

Once have stratification factors set and multivariate analysis complete, can create survival curves that are dependent on a given stratification factor. Basically, whenever a Kepler Meier analysis was performed, could also create a multivariate survival curve for that given stratification factor.

Question: For a Multivariate Cox regression analysis, can a stratification factor have some empty entries? (I’m assuming the answer is no)

Answer: Should hopefully still work, but must make sure that not too many values are missing.

# Competing Risk Analysis

* Competing risks is when subjects have multiple possible events in a time-to-event setting.
* Competing risk analysis is a way to isolate different causes of the “event of interest”, without just attaching one cause to the event of interest and labelling everything else as a censored event. For example, if our event of interest was relapse post chemotherapy, but some people were also dying from the cancer at the same time, this establishes a competing risk event; therefore, relapse and death need to be considered separately. However, for our study, since our “event of interest” includes first documentation of disease progression or death from any cause, this functionally encompasses all possibilities and so a competing risk analysis may not directly apply here. Question: is, how should it be applied in our study, if at all? Unless the idea is that while our original study may assume one overarching event of interest, our competing risk analysis is performed to break down the event of interest into its component parts and analyze them individually.
* Answer: Don’t have to do this. Consdier later maybe.

# Theory of Statistical tests/Application in R

* 80% power and 2-sided alpha of 0.05, we require a total of 170 patients
* Chi-square (i.e., contingency tables) vs. fisher exact (2-by-2 contingency tables): When we try to compare proportions of a categorical outcome according to different independent groups, we can consider several statistical tests such as chi-squared test, Fisher's exact test, or z-test. The chi-squared test and Fisher's exact test can assess for independence between two variables when the comparing groups are independent and not correlated. The chi-squared test applies an approximation assuming the sample is large, while the Fisher's exact test runs an exact procedure especially for small-sized samples.
* Students t test – t test in chapter 10.8 of stats textbook
* Kaplan-Meier Curves, log rank test (2 sided stratified), hazard ratios (briefly mentioned): <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3932959/> This article gives a brief theoretical explanation.
* Using R for Kaplan-Meier curves and log rank test: <http://www.sthda.com/english/wiki/survival-analysis-basics>
* Multivariate Cox proportional hazards //// Multivariate cox proportional hazards model adjusted for stratification factors to estimate adjusted hazard ratios and 95% CI’s 🡪
* Using R for Cox Proportional-Hazards Model: <http://www.sthda.com/english/wiki/cox-proportional-hazards-model>
* Competing risk analysis to assess the effect of censoring events on progression free survival<https://www.nature.com/articles/1705727>. <https://www.emilyzabor.com/tutorials/survival_analysis_in_r_tutorial.html#Part_3:_Competing_Risks> (2nd link is better)