

Multiple Myeloma: Challenges in Identifying regimen modification for a given Line of Chemotherapy

Mrityunjay Kumar, PhUSE EU Connect Birmingham, UK 5th – 8th November 2023

Stream: Coding Tips & Tricks (CT)

Presentation No. - CT03



1. What is Multiple Myeloma ?
2. What treatments are available for Multiple Myeloma (MM) ?
3. How are chemotherapy treatments used in MM ?
4. What are multiple line of treatment regimens ?
5. What are common types of treatment modifications ?
 - challenges in identifying modifications to treatment regimens for a given line of therapy ?
6. What programming steps are involved to derive time interval for a given line of therapy after treatment modification ?

1. What is Multiple Myeloma ?

1. Multiple Myeloma (MM) is a type of blood cancer that develops in plasma cells inside the bone marrow which is soft, spongy tissue at the center of bones.
2. Plasma cells are transformed into cancerous cells that grow out of control and reduce normal blood cells (RBCs, WBCs, and Platelets) that help fight infection.
3. These malignant plasma cells then produce an abnormal antibody called M protein, high levels of which are a characteristic feature of multiple myeloma.
4. In healthy bone marrow, normal plasma cells make antibodies to protect the body from infection.

2. What are treatments available for Multiple Myeloma (MM)

1. Immunomodulatory drugs (IMiDs): Lenalidomide (Revlimid), Pomalidomide (Pomalyst), and Thalidomide (Thalomid) that help immune system fight the cancer.
2. Proteasome Inhibitors: Bortezomib (Velcade), Carfilzomib (Kyprolis) and Ixazomib (Ninlaro) also help trigger death of Myeloma cells by means of defective proteins.
3. Steroids, Antibodies
4. B-cell maturation antigen (BCMA)-targeted therapies
5. Stem Cell Therapy

3. How are chemotherapy treatments used in MM ?

- Chemotherapy treatments are used in combination with cancer drugs.
Example: IMiD drugs in Triplet (XYZ), Quadruplet (ABCD) etc.
- Chemotherapy is most often used in preparation for a Stem Cell Transplant (SCT) i.e before or after SCT.

4. What are multiple line of treatment regimens ?

1. Patients with multiple myeloma are treated for a certain duration which are divided into a number of course.
2. Each course of treatment is known as lines of treatment.
3. Since treatment is given in combination with several anti-cancer drugs its also known as regimen.
4. For example: subject 1001 receives 3 Lines of treatment and 7 types of regimens.
5. Similarly, subject 1002 receives 2 Lines of treatment and 8 types of regimens in current scenario.

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND
1001	23-Feb-17	1	Induction	1	D/C/A	12-Jun-10	08-Dec-11
1001	23-Feb-17	1	SCT	2	A	20-Dec-11	20-Dec-11
1001	23-Feb-17	1	Maintenance	3	Z	23-Apr-12	13-Feb-17
1001	23-Feb-17	2	Induction	4	X/Y/Z	23-Feb-17	12-Sep-17
1001	23-Feb-17	2	Consolidation	5	X	20-Sep-17	11-Nov-17
1001	23-Feb-17	2	Maintenance	6	Y	21-Dec-17	06-Feb-18
1001	23-Feb-17	3	Induction	7	Z	06-Feb-18	26-Jun-18
1002	06-Feb-18	1	Induction	1	A/B	12-Jun-10	08-Dec-11
1002	06-Feb-18	1	Consolidation	2	C	20-Dec-11	20-Dec-11
1002	06-Feb-18	1	Consolidation	3	B/C	23-Apr-12	13-Feb-17
1002	06-Feb-18	1	SCT	4	D/E/F	23-Feb-17	12-Sep-17
1002	06-Feb-18	1	SCT	5	D	20-Sep-17	11-Nov-17
1002	06-Feb-18	1	Maintenance	6	X	21-Dec-17	06-Feb-18
1002	06-Feb-18	2	Induction	7	X/Y	06-Feb-18	26-Jun-18
1002	06-Feb-18	2	Induction	8	X/Y/Z	16-Aug-18	30-Sep-18

Table: Sample Data Structure

USUBJID = Unique subject identifier
 INDEXDT = Index Line Treatment start date
 LINE = Multiple lines of treatment received by a patient
 PHASE = Various phases of treatment
 SEQNUM = Treatment sequence number (also known as regimen number)
 TSTART = Treatment start date,
 TEND = Treatment end date

5. What are common types of treatment modifications ?

Treatment Regimen ($\geq$ TRIPLET) modification can be explained by below 4 subcategories.

1. Discontinuation of part of 2L (XYZ $\rightarrow$ XY or YZ)
2. Addition to part of 2L (XY $\rightarrow$ XYZ or YZ $\rightarrow$ XYZ)
3. Discontinuation of complete 2L
4. Changed to different regimens from $\geq$ TRIPLET regimens (XYZ $\rightarrow$ ABC)

Challenges in identifying treatment modification and calculating duration in days

- A. Subject disposition form captures subject level discontinuation data.
- B. Since, each subject is taking multiple regimens within each line of treatment, it becomes difficult to correctly identify which treatment group was modified ?
- C. Find exact date of treatment modified and calculate time duration from 2nd Line start date which is also known as Index date.

6. Programming steps involved

We are going to see what are various programming steps involved to derive time interval followed by treatment modification.

1. Discontinuation of part of 2L (XYZ → XY or YZ)

- We want to identify which subject has discontinued part of second line treatment regimen.
- We know that here each subject is taking triplet (XYZ) regimen, and the challenge is to correctly identify the subject who is now taking (XY or YZ) after discontinuation.
- Then calculate time interval between Index date of 2L and modified regimen.

Table 1: Input data - RDROP

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND
2001	12-Jun-10	2	Induction	2	D/C/A	12-Jun-10	08-Dec-11
2002	17-Jun-11	2	SCT	3	A	17-Oct-11	20-Dec-11
2003	02-Jun-13	2	Maintenance	4	Z	23-Apr-12	13-May-13
2003	02-Jun-13	2	Induction	5	X	13-May-13	28-May-13
2003	02-Jun-13	2	Consolidation	6	X/Y/Z	02-Jun-13	25-Apr-16
2003	02-Jun-13	2	Consolidation	7	X/Z	26-Apr-16	15-Jun-16
2004	08-Jul-17	2	Induction	3	Z	12-Jul-17	12-Jul-18
2004	08-Jul-17	2	Consolidation	4	G	04-Oct-19	15-Dec-19
2004	08-Jul-17	3	Maintenance	5	O	15-Jan-20	21-Dec-20

1. Discontinuation of part of 2L (XYZ $\rightarrow$ XY or YZ) contd..

Step # 1. List subject who is taking multiple triplet (XYZ) regimen

```
data rdrop_subj;
  set rdrop;
  by usubjid;
  if first.usubjid then count = 1;
  else count + 1;
  if count > 1 ;
run;

proc sort data=rdrop_subj out=rdrop_subj (keep=usubjid) nodupkey;
  by usubjid;
run;
```

Step # 2. Merge subject list back with 2L triplet (XYZ) regimen

```
data rdrop2L;
  merge rdrop
        rdrop_subj (in=a);
  by usubjid;
  if a;
  drug_count = count(drug, '/') + 1;
  d_flag = lag(drug_count);
  if first.usubjid then drug_count = .;
run;
```

1. Discontinuation of part of 2L (XYZ → XY or YZ) contd..

Step # 3. Create drug discontinuation flag and calculate duration from 2L Index date

```
data rdrop_dur;
  set rdrop2L;
  by usubjid;
  if .z< drug_count < d_flag then dropfl = 'Y';
  if dropfl = 'Y' and tstart > indexdt then drop_dur = (tstart - indexdt +1);
run;
```

Table 1.1: Output data RDROP_DUR

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND	DROP_DUR
2003	02-Jun-13	2	Consolidation	7	X/Z	26-Apr-16	15-Jun-16	1060

2. Addition to part of 2L (XY → XYZ or YZ → XYZ)

Refer to the example below in Table 2 regimen data (RADD), where new treatment is added to existing regimen (treatment sequence). We want to identify for which subject there has been addition of new drug as part of second line regimen.

Here, the challenge is to correctly identify the subject who is taking (Z) and after addition and treatment regimen becomes XY → XYZ and thereafter calculate time interval between Index date of 2L and modified regimen. (Note: consider 1st occurrence in case of multiple cases of addition of new drug).

Table 2. Input data – RADD

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND
3001	06-Jun-14	2	Induction	2	D/C/A	06-Jun-14	19-Jul-14
3002	22-Jun-15	2	Maintenance	3	A	20-Dec-11	20-Dec-11
4003	04-Apr-16	2	Induction	2	X/Y/Z	04-Apr-16	14-Jun-16
4003	04-Apr-16	2	SCT	3	X/Y	02-Aug-16	02-Aug-16
4003	04-Apr-16	2	Consolidation	4	X/Y/Z	17-Nov-16	22-Apr-17
4003	04-Apr-16	2	Maintenance	5	X	22-Apr-17	14-Mar-18
5004	13-Jul-17	2	Induction	2	A/B/C/D/E	13-Jul-17	19-Jul-17
5004	13-Jul-17	2	SCT	3	Z	29-Jul-17	30-Jul-17
5004	13-Jul-17	3	Maintenance	4	O	14-Aug-17	11-Sep-17

2. Addition to part of 2L (XY → XYZ or YZ → XYZ)

Refer to previous Step # 2 for source data RDROP2L and Create drug addition flag to calculate duration from 2L Index date.

```
data radd_dur;
  set rdrop2L;
  by usubjid;
  if .z< drug_count > d_flag then addfl = 'Y';
  if addfl = 'Y' and tstart > indexdt then add_dur = (tstart - indexdt +1);
run;
```

Table 2.1: Output data RADD_DUR

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND	ADD_DUR
4003	04-Apr-16	2	Consolidation	4	X/Y/Z	17-Nov-16	22-Apr-17	228

3. Discontinuation of complete 2L

Treatment discontinuation data gets captured in subject's disposition form where we can see treatment discontinuation dates with reasons as why the subject discontinued the treatment and there could be various reasons such as Completed study, Lost-to follow up, Death etc.

However, in cases where multiple lines of treatment regimens are given to treat patients such as Multiple myeloma (MM), it becomes tricky for the programmer to understand logic of correctly identifying discontinuation date because of below two reasons.

1. Not all subjects may have treatment discontinuation dates available in disposition data since treatment may be ongoing.
When treatment discontinuation date is present but it's after 2L treatment end date, take minimum of DISCDT and TEND as last date and calculate time interval from INDEXDT i.e 2nd Line start date.
2. Many subjects may have started the next line of treatment regimens i.e., after 2nd Line, 3rd Line treatment got started.
When treatment discontinuation date is absent, and subjects started 3rd line of treatments (3L) followed by 2nd line of treatment (2L). In this situation consider taking minimum of 3L treatment start or 2L treatment end as last date and calculate time interval from INDEXDT i.e., 2nd Line start date.

3. Discontinuation of complete 2L

Table 3. Input data – DISC

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND	DISCDT
6001	30-Oct-15	1	Induction	1	A/B/C/D/E	13-Jul-13	19-Jul-14	13-Aug-16
6001	30-Oct-15	1data	Induction	2	A/Y/Z	29-Jul-14	30-Jul-15	13-Aug-16
6001	30-Oct-15	1	Maintenance	3	Y	14-Aug-15	11-Sep-15	13-Aug-16
6001	30-Oct-15	2	SCT	4	X/Y/Z	30-Oct-15	15-Aug-16	13-Aug-16
6002	12-Apr-17	1	Induction	1	X/Y/Z	14-Jan-16	11-Apr-17	08-Feb-18
6002	12-Apr-17	2	Maintenance	2	X/Y/Z	12-Apr-17	04-Jan-18	08-Feb-18
7001	06-Jul-16	1	Induction	1	D/E/G	22-Jun-14	22-Sep-14	
7001	06-Jul-16	1	Maintenance	2	E/F	10-Mar-15	10-Jun-15	
7001	06-Jul-16	1	Maintenance	3	D/E	15-Jun-15	05-Jul-16	
7001	06-Jul-16	2	SCT	4	X/Y/Z	06-Jul-16	27-Jul-17	
7001	06-Jul-16	3	Induction	5	X/Z	17-Aug-17	29-Jan-18	
7001	06-Jul-16	3	Maintenance	6	X/Z	12-Feb-18	28-Jan-19	
7001	06-Jul-16	4	Induction	7	X/Y/Z	25-Jul-19	07-Feb-21	

Table 3.1: Output data DISC_DUR

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND	DISCDT
6001	30-Oct-15	2	SCT	4	X/Y/Z	30-Oct-15	15-Aug-16	13-Aug-16
6002	12-Apr-17	2	Maintenance	2	X/Y/Z	12-Apr-17	04-Jan-18	08-Feb-18
7001	06-Jul-16	2	SCT	4	X/Y/Z	06-Jul-16	27-Jul-17	

```
proc sort data = disc out = sorted_disc;
  by usubjid descending tstart;
run;
```

```
data disct;
  set sorted_disc;
  next_tstart = lag(tstart);
  lag_usubjid = lag(usubjid);
  format next_tstart date9.;
  if line = 2;
run;
```

```
data disc_dur;
  set disct;
  if subjid = lag_usubjid then do;
    disc_dur = min(next_tstart, tend, discdt)
               - indexdt + 1;
  end;
run;
```

4. Changed to different regimens from >= TRIPLET regimens (XYZ → ABC)

In the next example of regimen modification, our interest is to identify when a subject’s triplet regimen is changed within the same line of chemotherapy. The duration of treatment discontinuation will be calculated by considering minimum of discontinuation or death date and next regimen change start date. In this case, we would be interested in knowing the 1st date of regimen change which can be seen in blue color.

Table 4: REGIMEN

USUBJID	INDEXDT	LINE	PHASE	SEQNUM	DRUG	TSTART	TEND
2004	02-Jun-13	1	Induction	1	D/C/A	12-Jun-10	08-Dec-11
2004	02-Jun-13	1	SCT	2	A	20-Dec-11	20-Dec-11
2004	02-Jun-13	1	Maintenance	3	Z	23-Apr-12	13-May-13
2004	02-Jun-13	2	Induction	4	X	13-May-13	28-May-13
2004	02-Jun-13	2	Consolidation	5	X/Y/Z	02-Jun-13	25-Apr-16
2004	02-Jun-13	2	Maintenance	6	A/B/C	26-Apr-16	15-Jun-17
2004	02-Jun-13	2	Induction	7	Z	12-Jul-17	12-Jul-17
2004	02-Jun-13	2	Induction	8	G	04-Oct-19	15-Dec-19
2004	02-Jun-13	3	Maintenance	9	O	15-Jan-20	21-Dec-20

```
data next_rgm;
  set regimen;
  if tstart > indexdt then reg_change = 'y';
  if reg_change = 'y' and line = 2;
run;

proc sort data = next_rgm;
  by usubjid;
run;

data nregmn;
  set next_rgm;
  by usubjid;
  reg_dur = min(tstart, discdt, dthdt) - indexdt + 1;
  if first.usubjid;
run;
```

Conclusion

- There has been an increase in the number of treatments as well as new types of treatment regimens are being used to treat patients diagnosed with multiple myeloma.
- Patients who are severely affected are advised to undergo Stem Cell Transplant (SCT) procedure.
- To see the significant impact of treatments across various lines over the given time interval has always been an interesting parameter for the statisticians to observe the effectiveness of treatment drug.
- Statistical programmers are requested to summarize treatment duration in the form of summary statistics when treatment regimens are modified.
- It may be helpful to understand data structure before applying any derivation logic to correctly derive time interval.
- In this paper, four types of treatment modifications criteria have been explained for second line (2L) when patients are taking multiple lines of treatment with the help of dummy data.
- However, programmers are advised to have in-depth understanding of data before reaching out to any derivation logic since there may be other scenarios as well.

Questions ???



Authors:

Main Author: Mrityunjay Kumar
Co-author: Shailesh Gupte

