The following describes how discontinuation (DISCONT), discontinue reason (ENDTRS\_C) and discontinue time (ENTRT\_PC) are consolidated for both the training and testing datasets. The participants can make adjustment of the dependent variables in the training set if they wish. But the testing set used for scoring will be fixed.

**(1)** **Group raw discontinue reasons intro 5 major categories**

a) As long as adverse events are mentioned in detailed reason of discontinue treatment #1 or #2, or comment of treatment discontinue, the patients are labeled as “discontinue due to adverse event”. This rule assigned “discontinue due to adverse event” status to 116 AZ patients and 1 EFC6546 patient.

b) Continue to group discontinue reasons into 5 major categories.

According to the table on page 2.

**(2) Consolidate discontinue time (ENTRT\_PC)**

Discontinue times for CELGENE and ASCENT2 are already consolidated in the raw data. Discontinue time for most of the AZ dataset is found in treatment #2 end date. For some AZ patients whose “discontinue due to adverse event” status is found in detailed reason for discontinue of treatment #1, treatment #1 end date is used. Maximum of discontinue dates for treatment #1, #2 and #3 is used as a proxy for EFC6546.

**(3) Determine the censored label of patients**

a) Patients in “miscellaneous groups” and patients with missing discontinue time information will be left out of scoring in the testing set. Participants should probably also discard “miscellaneous” patients and patients with missing discontinue time in the training set. But it’s left for them to decide.

b) Patients are labeled as DISCONT=1 if and only if they discontinue due to adverse event or possibly due to adverse event before 3 months. Otherwise, the patients are labeled as DISCONT=0. Some patients discontinued due to certain other reasons that are still possibly related to adverse events in an indirect way. But the patients who discontinued possibly due to adverse event should be considered to have a lower signal-to-noise level relative to patients who discontinued unambiguously due to adverse event.

|  |  |
| --- | --- |
| **Raw discontinue reason** | ENDTRS\_C |
| ADVERSE EVENT | discontinue due to adverse event (AE) |
| ADVERSE EVENT (COMPLETE AE FORM) | discontinue due to adverse event |
| DOCETAXEL TOXICITY (COMPLETE GI, TE, AND/OR AE FORMS AS APPROPRIATE) | discontinue due to adverse event |
| LOST TO FOLLOW UP | miscellaneous groups (misce) |
| SUBJECT LOST TO FOLLOW-UP | miscellaneous groups |
| POOR COMPLIANCE TO PROTOCOL | miscellaneous groups |
| PROTOCOL VIOLATION | miscellaneous groups |
| Empty string (non-EFC6546 patients) | miscellaneous groups |
| LACK OF THERAPEUTIC RESPONSE | death or progression (progression) |
| CONDITION UNDER INVESTIGATION WORSENED | death or progression |
| DEATH | death or progression |
| DEATH (COMPLETE SURVIVAL FORM) | death or progression |
| DISEASE PROGRESSION | death or progression |
| Empty string (one EFC6546 patient) | complete study (complete) |
| CONDITION UNDER INVESTIGATION IMPROVED / SUBJECT RECOVERED | complete study |
| COMPLETED 30 WEEKS OF STUDY TREATMENT PHASE | complete study |
| MAXIMUM CYCLE OF CHEMOTHERAPY REACHED | complete study |
| AT THE REQUEST OF SUBJECT (WILL CONTINUE ON STUDY) | discontinue possibly due to adverse event (possible\_AE) |
| DEVELOPMENT OF STUDY SPECIFIC DISCONTINUATION CRITERIA | discontinue possibly due to adverse event |
| OTHER REASON | discontinue possibly due to adverse event |
| INVESTIGATORS DECISION | discontinue possibly due to adverse event |
| OTHER | discontinue possibly due to adverse event |
| VOLUNTARY DISCONTINUATION BY SUBJECT | discontinue possibly due to adverse event |
| WITHDRAWAL OF CONSENT | discontinue possibly due to adverse event |
| WITHDREW CONSENT | discontinue possibly due to adverse event |