

# Deriving Adverse Event Records Based on Study Treatments

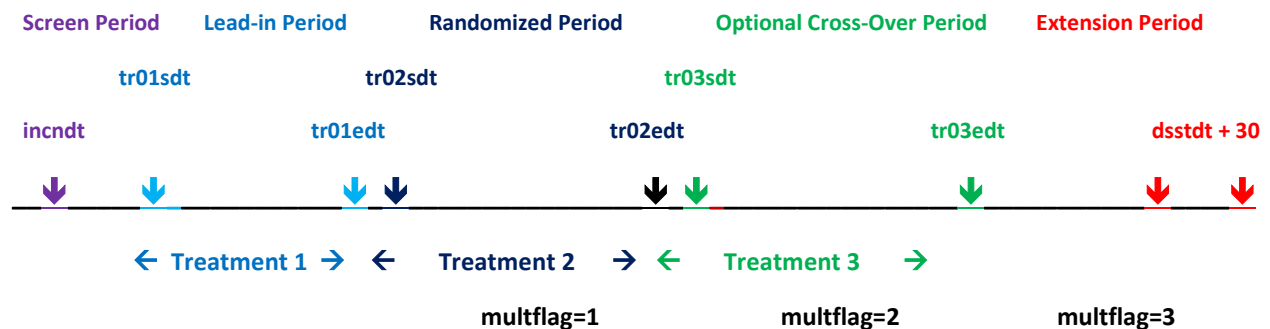
Jonson Jiang, M.S., inVentiv Health, Princeton, NJ, USA

## ABSTRACT

This paper puts forward an approach on how to create duplicate records from one Adverse Event (AE) datum based on study treatments. In order to fulfill this task, one flag was created to check if we need to produce duplicate records depending on an existing AE in different treatment periods. If yes, then create these duplicate records and derive AE dates for these duplicate records based on treatment periods or discontinued dates.

## INTRODUCTION

One clinical trial study was designed with 5 periods as shown below. (1) Informed consent date is start date of "Screen Period". (2) The first treatment is at "Lead-in Period". (3) The second treatment is at "Randomized Period". (4) The third treatment is at "Optional Cross-Over Period", which is for patients in the placebo group. The last period is called "Extension Period".



In the source SDTM AE dataset, there was only one record per patient per AE. So, what we needed to do in target ADaM ADAE dataset was to create duplicate records if one AE crossed-out treatment period. If duplicate records populated, then derive AE start date (ASTDT) as start date of corresponding treatment period, and AE end date (AENDT) as end date of corresponding treatment period. If a patient discontinued, then use `dsstdt+30` as cut-off date.

Faced with an on-going study with many data issues, we had to consider as many situations as possible for treatment data, AE data, and Discontinuation data.

**AE records:** there are 2 cases:

1. With both non-missing AE start date and AE end date (create duplicate records and derive AE start and/or AE end date)
2. Missing AE start date or missing AE end date (don't need for duplicate records and derive AE start and/or AE end date);

**Treatments records:** there could be 8 cases:

1. Without treatment date (in Screen Period);
2. Just start treatment 1 (with tr01sdt and without all other treatment date);
3. Completed treatment 1 (with tr01sdt, tr01edt and without other treatment date);
4. Start treatment 2 (with tr01sdt, tr01edt, tr02sdt and without other treatment date);
5. Completed treatment 2 (without treatment 3 date) - completed all treatments or have not started Cross-Over study for placebo patient only);
6. Start treatment 3 (without treatment 3 end date) ;
7. Completed all three treatments;
8.  $tr02sdt - tr01edt = 1$  (no gaps) or  $tr02sdt - tr01edt > 1$  (there is some gaps).  
Similar as  $tr03sdt - tr02edt$ ;

**Discontinuation records:** there are 2 cases:

1. With discontinuation date (dsstdt);
2. Without discontinuation date (on-going).

Based on the above data situations, two parts in the program were set up to handle this task. One is to check if we need to create duplicate record depending on AE existent in different treatment periods. If yes, create these duplicate records. Another is to derive AE end date or AE start date based on treatment period or discontinued date.

## SAS CODES AND DISCUSSION

In order to achieve this goal, one variable named "multiple flag" (multflag) was created.

```
=====
* if AE start date and AE end date of an AE record spans in multiple treatment periods, then create multiple *
* records based on the treatment start and end dates *;
data step_1;
    set ae2;

    * Case 1 - keep the record for all original AE data (level 0) *;
    multflag=0;
    output;

    * Case 2 - add one duplicate record for AE crossed-out treatment 1 into Randomized period or *
    * Extension period *;
    if (TR02SDT^=. and .<ASTDT<TR02SDT and AENDT>=TR02SDT>.) then do;
        multflag=1;
        output;
    end;
    else if (TR02SDT=. and .<ASTDT<=TR01EDT and AENDT>TR01EDT>.) then do;
        multflag=1;
        output;
    end;

    * Case 3 - add one duplicate record for AE crossed-out treatment 2 into Cross-Over Period or *
    * Extension Period *;
    if (TR03SDT^=. and .<ASTDT<TR03SDT and AENDT>=TR03SDT>.) then do;
        multflag=2;
        output;
    end;
    else if (TR03SDT=. and .<ASTDT<=TR02EDT and AENDT>TR02EDT>.) then do;
        multflag=2;
    end;
```

```

        output;
    end;

    * Case 4 - add one duplicate record for AE crossed-out treatment 3 into Extension Period *;
    if (.<ASTDT<=TR03EDT and AENDT>TR03EDT>.) then do;
        multflag=3;
        output;
    end;
run;
=====

```

**There are four cases:**

**First:** Keep all original AE data to set multflag as 0.

**Second:** Check if AE crossed-out from treatment 1 into 'Randomized' Period or Extension Period. If yes, add one duplicate record and then set multflag=1 for this record. That means if AE start date is before treatment 2 and AE end date is in or after treatment 2 then set multflag as 1. There are two situations here.

[1] If treatment 2 start date non-missing, then use it to identify if we need create duplicated record and multflag;

[2] If treatment 2 start date is missing, then use treatment 1 end date to identify duplicated record and multflag.

**Third:** Check if AE crossed-out from treatment 2 into 'cross-over' Period or Extension Period. If yes, add one duplicate record and then set multflag=2 for this record. That means if AE start date is before treatment 3 and AE end date is in or after treatment 3 then set multflag as 2. There are two situations here.

[1] If treatment 3 start date non-missing, then use it to identify if we need create duplicated record and multflag;

[2] If treatment 3 start date is missing, then use treatment 2 end date to identify duplicated record and multflag.

**Fourth:** Check if AE crossed-out treatment 3 into Extension Period. For this situation, we only need treatment 3 end date to identify duplicated records and multflag.

For example, if a patient completed 3 treatments and one AE start date was before treatment 2 and AE end date was after treatment 3 end date, keep original record and create another three duplicated records. For those records, Multflag will be 0, 1, 2 and 3. If this patient's AE start date was in treatment 2 and AE end date was after treatment 3 end date, keep original record and create another two duplicated records. For those records, Multflag will be 0, 2 and 3...

After creating multiple records and their multiple flags, we need to think about how to impute AE start date and AE end date for those records. This leads to the second part of the program.

```

=====

data step_2;
    set step_1;

    if multflag=0 then do; * only derive AE end date based on AE start date in which period *;

```

```

* 1. AE start in or before treatment 1 period and AE end after treatment 1*;
  if TR02SDT^=. and .<ASTDT<TR02SDT and AENDT>=TR02SDT>. then AENDT=TR02SDT-1;
  else if TR02SDT=. and .<ASTDT<=TR01EDT and AENDT>=TR01EDT>. then AENDT=TR01EDT;

* 2. AE start in treatment 2 period and AE end after treatment 2*;
  if TR03SDT^=. and .<TR02SDT<=ASTDT<TR03SDT and AENDT>=TR03SDT>. then AENDT=TR03SDT-1;
  else if TR03SDT=. and .<ASTDT<=TR02EDT and AENDT>=TR02EDT>. then AENDT=TR02EDT;

* 3. AE start in treatment 3 period and AE end after treatment 3 in Extension*;
  if .<TR03SDT<ASTDT<=TR03EDT and AENDT>=TR03EDT>. then AENDT=TR03EDT;

* 4. AE start in Extension Period before discontinuation and AE end after discontinuation*;
  if dsstdt>. and ASTDT<=dsstdt and AENDT>(dsstdt+30) then do;

    * [1] completed treatment 3 *;
    if ASTDT>TR03EDT>. then AENDT=(dsstdt+30);

    * [2] completed treatment 2 without treatment 3 *;
    Else if ASTDT>TR02EDT>. and TR03SDT=. and TR03EDT=. then AENDT=(dsstdt+30);

    * [3] completed treatment 1 without treatment 2 and treatemnt 3 *;
    Else if ASTDT>TR01EDT>. and TR02SDT=. and TR02EDT=. and TR03SDT=. and TR03EDT=.
    then AENDT=(dsstdt+30);

  end;
end;
* AE in Randomized (treatment 2) Period to impute both AE start date and end date*;
else if multflag=1 then do;
  if TR01EDT>. then ASTDT=TR01EDT+1;
  if TR02SDT>. then ASTDT=TR02SDT;

  * with treatment 'Rand' end date and 'CO' (treatment 3) start date*;
  if AENDT>=TR02EDT>. then AENDT=TR02EDT;
  if AENDT>=TR03SDT>. then AENDT=TR03SDT-1;

  * without 'Rand' end date and 'CO' (treatment 3) start date *;
  if (TR02EDT=. and TR03SDT=.) and dsstdt>. and AENDT>(dsstdt+30)
  then AENDT=(dsstdt+30);
end;
* AE in Cross-Over (treatment 3) Period to impute both AE start date and end date *;
else if multflag=2 then do;
  if TR02EDT>. then ASTDT=TR02EDT+1;
  if TR03SDT>. then ASTDT=TR03SDT;

  * with treatment 'CO' end date*;
  if AENDT>=TR03EDT>. then AENDT=TR03EDT;
  * without 'CO' end date *;
  if TR03EDT=. and dsstdt>. and AENDT>(dsstdt+30) then AENDT=(dsstdt+30);
end;
* AE after treatment 3 in Extension Period to impute both AE start date and end date *;
else if multflag=3 then do;
  ASTDT=TR03EDT+1;
  if dsstdt>. and AENDT>(dsstdt+30) then AENDT=(dsstdt+30);
end;
run;

```

=====

**There are two Scenarios:**

**First:** Consider **multflag=0**. Because it comes from all of level 0 AE data, it could be cut off by different treatment date. So we only need to derive AE end date based on AE start date in which treatment period.

1. If AE starts in or before treatment 1 period and AE ends after treatment 1, then we need to derive AE end date as start date of treatment 2 minus 1 if treatment 2 start date is non-missing; or derive AE end date as end date of treatment 1 if treatment 2 start date is missing.
2. If AE start in treatment 2 period and AE end after treatment 2, then we need to derive AE end date as start date of treatment 3 minus 1 if treatment 3 start date is non-missing; or derive AE end date as end date of treatment 2 if treatment 3 start date is missing.
3. For AE start in treatment 3 period and AE end after treatment 3, we only consider treatment 3 end date as imputed AE end date.
4. If AE starts in Extension Period before discontinuation and AE ends after discontinuation, we need to set discontinuation plus 30 as the limit for imputed AE end date. In order to make sure AE is in Extension Period, the condition of "AE start date > last non-missing treatment end date" was assumed.

There are three cases for AE starts in Extension before discontinue and AE ends after discontinue.

- [1] Completed treatment 3 - check AE start date is greater than tr03edt.
- [2] Completed treatment 2 without treatment 3 - check AE start date is greater than tr02edt.
- [3] Completed treatment 1 without treatment 2 and 3 - check AE start date is greater than tr01edt.

**Second:** Consider multflag greater than 0. In this situation, all of records are duplicated records. Both AE start date and AE end date would be imputed based on multflag equals 1 or 2 or 3.

**Multflag=1** means this duplicate AE is in treatment 2 (Randomized Period). The treatment 2 data will be used to impute its AE start date and AE end date.

1. Imputed AE start date: If treatment 2 start date is missing, the treatment 1 end date plus 1 is used as imputed AE start date; If treatment 2 start date is not missing, then treatment 2 start date is used as imputed AE start date for this duplicate record. So the second condition here has priority and will overwrite the first one.
2. Imputed AE end date: If treatment 3 start date is missing, then treatment 2 end date is used as imputed AE end date; If treatment 3 start date is not missing, then treatment 3 start date minus one is used as imputed AE end date for this duplicate record. Also the second condition here has priority and will overwrite the first one; If without both treatment 2 end date and treatment 3 start date records, then discontinuation date plus 30 (cutoff date) will be used to check or set as imputed AE end date. For no discontinuation, we just keep as is AE end date.

**Multflag=2** means this duplicate AE is in treatment 3 (Cross-Over Period). The treatment 3 data will be used to impute its AE start date and AE end date.

1. Imputed AE start date: If treatment 3 start date is missing, the treatment 2 end date plus 1 is used as imputed AE start date; If treatment 3 start date is not missing, then treatment 3 start date is used as imputed AE start date for this duplicate record. So the second condition here has priority and will overwrite the first one.
2. Imputed AE end date: For this situation, treatment 3 end date is used as imputed AE end; If without treatment 3 end date, then discontinuation date plus 30 (cutoff date) will be used to check or set as imputed AE end date. For no discontinuation, we just keep as is AE end date.

For **multflag=3**, the record is in Extension Period. Treatment 3 end date should not be missing. So treatment 3 end date plus one is used as imputed AE start date and also check if we need to derive AE end date based on discontinuation date plus 30.

#### **SUMMARY AND CONCLUSION**

Only one flag is used to create duplicate records. The two simple steps in the program cover many data situations for treatment data, AE data and Discontinuation data. We can easily check and debug using this approach. If there are duplicate records problems in final dataset, just go to 'Step one' to check the value of multiple flag; If there is a problem for derived AE date, just go to 'Step two' to debug. Because we divided one AE into multiple duplicate records, two ways can be used to check if those multiple records are correct. One way is to check if multiple records are divided based on treatments; and the second way is to see if those multiple records can be restored back without gaps.

#### **ACKNOWLEDGMENTS**

I am grateful to my Manager (Maman Bio) for his support, guidance, and review of my paper.

#### **CONTACT INFORMATION**

Author Name: Jonson Jiang  
Company: inVentiv Health  
Address: 910 183<sup>rd</sup> PL SE, Bothell, WA 98012  
Email: [jonson.jiang@inventivhealth.com](mailto:jonson.jiang@inventivhealth.com)