

**LOCF vs. LOV in Endpoint DTYPE Development with CDISC ADaM Standards**

Maggie Ci Jiang, Teva Pharmaceuticals, West Chester, PA

**ABSTRACT**

CDSIC ADaM Implementation Guide v1.0 (IG) has defined the standards on how to use the DTYPE terminology when developing endpoint values in ADaM BDS datasets, and provided users some examples to illustrate how to utilize the standards. However, the definitions and examples from ADaM IG are limited to the applications of the sole definition or an individual case such as DTYPE for AVERAGE, LOV or Endpoint. When it comes to the situation that multi-criteria are defined in SAP, for example, SAP defines that the Last Observed non-missing Value (LOV) is the endpoint value, and also allows Last Observation Carried Forward (LOCF) to be applied to the scheduled visit if the expected value is missing, which terminology should be used for the derived endpoint DTYPE, LOCF or LOV or just Endpoint? This paper will present a review of experiences implementing endpoint DTYPE in ADaM BDS datasets, and discuss the comprehensive utilization of DTYPE terminology step by step with some practical examples.

**INTRODUCTION**

CDISC ADaM IG has defined the DTYPE usage for ENDPOINT variable in BDS datasets. However, when it comes to the real world, the practice seems more complicated than it looks. For example, SAP defines that the Last Observed non-missing Value (LOV) is the endpoint value, and also allows Last Observation Carried Forward (LOCF) to be applied to the scheduled visit if the expected value is missing, which terminology should be used for the derived endpoint DTYPE, LOCF or LOV or just Endpoint? This paper focuses on implementation side of the comprehensive utilization of DTYPE terminology step by step. It assumes basic knowledge of CDISC ADaM data structure, SAS programming and ADaM Implementation Guide v1.0 (IG).

The following three examples Table 1.1, Table 1.2 and Table 1.3 will be used throughout the discussion in this paper, which illustrates a clinical trial that a patient has completed the clinical trial successfully if this patients finished the required four scheduled visits with complete available data; or a patient is considered as an early termination if this patient has not successfully fulfilled the four scheduled visits per the trial design. The three tables represent the different cases of vital signs data collection, usually in an ADaM BDS dataset ADVS.

Table 1.1: Month 3 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL |
|---------|------------------|----------|--------------|---------|--------------|------|
| 0001    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |
| 0001    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |
| 0001    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   |
| 0001    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | .    |

**Table 1.1 Missing one value**

Table 1.2: Month 2 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL |
|---------|------------------|----------|--------------|---------|--------------|------|
| 0002    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |
| 0002    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |
| 0002    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | .    |
| 0002    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | .    |

**Table 1.2 Missing two values**

Table 1.3: There's no missing value, End of Study visit is the last non-missing value

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL |
|---------|------------------|----------|--------------|---------|--------------|------|
| 0003    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |
| 0003    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |
| 0003    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   |
| 0003    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | 90   |

**Table 1.3 No missing values**

## DERIVATION OF LOV FLAG

In clinical trials, there are various types of ENDPOINT values such as SUM, AVERAGE or MAXIMUM etc., among them the most popular ENDPOINT value is often defined as the last observed post baseline value (LOV) or the last non-missing observed post baseline value on Treatment (LVOTFL). Our discussion will skip the details about the types of SUM, AVERAGE or MAXIMUM, please refer to ADaM IG for examples <sup>[1]</sup>, the focus here will be on the application of LOV and LVOTFL.

LVOTFL is Last Value On Treatment Record Flag used to identify last non-missing value LOV on treatment for each group of parameter <sup>[2]</sup>. See the examples below.

Table 2.1: Month 3 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL |
|---------|------------------|----------|--------------|---------|--------------|------|--------|
| 0001    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |
| 0001    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |        |
| 0001    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   | Y      |
| 0001    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | .    |        |

**Table 2.1 Missing one value**

Table 2.2: Month 2 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL |
|---------|------------------|----------|--------------|---------|--------------|------|--------|
| 0002    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |
| 0002    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   | Y      |
| 0002    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | .    |        |
| 0002    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | .    |        |

**Table 2.2 Missing two values**

Table 2.3: There's no missing value, End of Study visit is the last non-missing post-baseline value

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL |
|---------|------------------|----------|--------------|---------|--------------|------|--------|
| 0003    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |
| 0003    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |        |
| 0003    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   |        |
| 0003    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | 90   | Y      |

**Table 2.3 No missing values**

In Table 2.1, Table 2.2 and Table 2.3 LVOTFL has been assigned to last non-missing visit for each PARAM Pulse Rate as long as the subject has a post-baseline value available. It's straight and easy, there's no derivation method involved so AVISIT keeps the same collected VISIT value as is for the row identified by LVOTFL=Y.

## DERIVATION OF LOCF DTYPE

In some clinical trials, when a scheduled visit is missing the assessment value, the last observed value can be carried over from the previous scheduled visit for the missing visit, which is called LOCF. When it's defined in the SAP that LOCF is required for analysis, the derivation type in ADaM dataset needs accommodate such required derivation procedure. For example, VISIT 3 can carry over the assessment value from VISIT 2 if this VISIT 3 is missing the assessment value and the VISIT 2 assessment value is available [Table 3.1], more, VISIT 4 can continue to carry over the value from VISIT 3 if VISIT 4 is also missing assessment value [Table 3.2].

Below shows how to identify the LOCF as the derived type DTYPE per ADaM IG.

Table 3.1: Month 3 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT   | AVISITN | AVISIT       | AVAL | DTYPE |
|---------|------------------|----------|---------|---------|--------------|------|-------|
| 0001    | Pulse Rate (bpm) | 1        | Month 1 | 1       | Month 1      | 60   |       |
| 0001    | Pulse Rate (bpm) | 2        | Month 2 | 2       | Month 2      | 70   |       |
| 0001    | Pulse Rate (bpm) | 3        | Month 3 | 3       | Month 3      | 80   |       |
| 0001    | Pulse Rate (bpm) | 3        | Month 3 | 4       | End of Study | 80   | LOCF  |

**Table 3.1 Missing one value**

Table 3.2: Month 2 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT   | AVISITN | AVISIT       | AVAL | DTYPE |
|---------|------------------|----------|---------|---------|--------------|------|-------|
| 0002    | Pulse Rate (bpm) | 1        | Month 1 | 1       | Month 1      | 60   |       |
| 0002    | Pulse Rate (bpm) | 2        | Month 2 | 2       | Month 2      | 70   |       |
| 0002    | Pulse Rate (bpm) | 2        | Month 2 | 3       | Month 3      | 70   | LOCF  |
| 0002    | Pulse Rate (bpm) | 2        | Month 2 | 4       | End of Study | 70   | LOCF  |

**Table 3.2 Missing two values**

Table 3.3: There's no missing value, End of Study visit is the last non-missing post-baseline value

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | DTYPE |
|---------|------------------|----------|--------------|---------|--------------|------|-------|
| 0003    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |       |
| 0003    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |       |
| 0003    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   |       |
| 0003    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | 90   |       |

**Table 3.3 No missing values**

For the above three tables, Table 3.1, Table 3.2 and Table 3.3, keep in mind that only the first two subjects "0001" and "0002" require LOCF applied to the missing values of the scheduled visits so the analysis can have the values available for all visits; the third subject does not need LOCF at all because this subject has completed the trial successfully with all four assessment values available for all four scheduled visits. After LOCF has been applied to the missing scheduled visits, the change made to the missing VISIT and VISITNUM is very critical per ADaM IG, this change informs the reviewers that the ENDPOINT value has been carried over from which previous VISIT. In Table 3.1, LOCF for End of Study visit has carried VISIT 3 over; In Table 3.2, LOCF for End of Study visit has carried VISIT 2 over. In Table 3.3, there's no LOCF involved.

## ENDPOINT DTYPE

Per CDISC ADaM IG, a row identifier, analysis timepoint is used to classify values within an analysis parameter into temporal or conceptual groups used for analysis. These groups may be observed, planned or derived<sup>[3]</sup>. The previous Tables 2.1, Table 2.2 and Table 2.3 are good examples of observed groups identified by LVOTFL; Tables 3.1, Table 3.2 and Table 3.3 are good examples of planned groups identified by LOCF. Now let's move on to derived groups.

When an ENDPOINT value is derived for a BDS dataset, a new row should be added with a corresponding description in AVISIT, and the DTYPE column should contain a description on that row such as "AVERAGE" or "ENDPOINT" to indicate both that the row has been derived, and also the derivation method<sup>[4]</sup>.

We've seen how DTYPE is practiced with LOV or LOCF applied separately to ADVS dataset. Let's discuss how to best present the endpoint DTYPE when both LOV and LOCF are required by analysis for a BDS dataset.

## ENDPOINT DTYPE WITHOUT LOCF INVOLVED

When a clinical trial analysis defines ENDPOINT is the last non-missing value, no LOCF has been applied to the missing scheduled visits. It's pretty straight forward, see examples below Table 4.1, Table 4.2 and Table 4.3. What we need to do is simply to add a new row by identifying AVISIT=Endpoint and AVISITN=99, AVAL equals the value that has been identified by LVOTFL, at the same time, assign derivation type DTYPE=LOV<sup>[5]</sup>.

Table 4.1: Month 3 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL | DTYPE |
|---------|------------------|----------|--------------|---------|--------------|------|--------|-------|
| 0001    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |       |
| 0001    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |        |       |
| 0001    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   | Y      |       |
| 0001    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | .    |        |       |
| 0001    | Pulse Rate (bpm) | 3        | Month 3      | 99      | Endpoint     | 80   |        | LOV   |

Table 4.1 Missing one value

Table 4.2: Month 2 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL | DTYPE |
|---------|------------------|----------|--------------|---------|--------------|------|--------|-------|
| 0002    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |       |
| 0002    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   | Y      |       |
| 0002    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | .    |        |       |
| 0002    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | .    |        |       |
| 0002    | Pulse Rate (bpm) | 2        | Month 2      | 99      | Endpoint     | 70   |        | LOV   |

Table 4.2 Missing two values

Table 4.3: There's no missing value, End of Study visit is the last non-missing post-baseline value

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL | DTYPE |
|---------|------------------|----------|--------------|---------|--------------|------|--------|-------|
| 0003    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |       |
| 0003    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |        |       |
| 0003    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   |        |       |
| 0003    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | 90   | Y      |       |
| 0003    | Pulse Rate (bpm) | 4        | End of Study | 99      | Endpoint     | 90   |        | LOV   |

Table 4.3 No missing values

The key note in above examples Table 4.1, Table 4.2 and Table 4.3 is that when AVISIT is assigned the value “Endpoint”, the VISIT of the LOV is also assigned to the VISIT for this derived endpoint row. This derivation clearly passes the message to the reviewers that the endpoint value has been defined to use the last non-missing observed value (LOV). You may have also noticed that the analysis visit AVISIT remains the same as the VISIT for the row identified by LVOTFL=“Y”, means that this value flagged by LVOTFL is also the last available value to be used for the analysis.

## ENDPOINT DTYPE WITH LOCF INVOLVED

It’s been somewhat challenging when a clinical trial analysis defines ENDPOINT is using the last non-missing value, and LOCF is also allowed to be applied to the missing assessment values for the scheduled visits. When such criteria occur, it has been observed that LOCF has been used for DTYPE very often, or LOV has been used sometimes. The important concept here is how not to get confused when the two conditions need be considered at the same time for the same set of data.

Let’s go through the three examples one more time. We do the three steps just like when we do one by one in previous sessions. First step is to derive the LVOTFL to identify the last observed value LOV, the second is to apply LOCF per SAP requirement, then to complete the derivation for DTYPE of the ENDPOINT.

Table 5.1: Month 3 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT   | AVISITN | AVISIT       | AVAL | LVOTFL | DTYPE    |
|---------|------------------|----------|---------|---------|--------------|------|--------|----------|
| 0001    | Pulse Rate (bpm) | 1        | Month 1 | 1       | Month 1      | 60   |        |          |
| 0001    | Pulse Rate (bpm) | 2        | Month 2 | 2       | Month 2      | 70   |        |          |
| 0001    | Pulse Rate (bpm) | 3        | Month 3 | 3       | Month 3      | 80   | Y      |          |
| 0001    | Pulse Rate (bpm) | 3        | Month 3 | 4       | End of Study | 80   |        | LOCF     |
| 0001    | Pulse Rate (bpm) |          |         | 99      | Endpoint     | 80   |        | Endpoint |

**Table 5.1 Missing one value**

Table 5.2: Month 2 is the last non-missing post-baseline value available

| USUBJID | PARAM            | VISITNUM | VISIT   | AVISITN | AVISIT       | AVAL | LVOTFL | DTYPE    |
|---------|------------------|----------|---------|---------|--------------|------|--------|----------|
| 0002    | Pulse Rate (bpm) | 1        | Month 1 | 1       | Month 1      | 60   |        |          |
| 0002    | Pulse Rate (bpm) | 2        | Month 2 | 2       | Month 2      | 70   | Y      |          |
| 0002    | Pulse Rate (bpm) | 2        | Month 2 | 3       | Month 3      | 70   |        | LOCF     |
| 0002    | Pulse Rate (bpm) | 2        | Month 2 | 4       | End of Study | 70   |        | LOCF     |
| 0002    | Pulse Rate (bpm) |          |         | 99      | Endpoint     | 70   |        | Endpoint |

**Table 5.2 Missing two values**

The reason why the “LOV” is suggested not to be used for the Endpoint DTYPE in Table 5.1 or Table 5.2 is that the last observed value “LOV” is not crystal clear for the derived endpoint rows like the previous examples Table 4.1 and Table 4.2 after “LOCF” has been applied to the last missing assessment value for the scheduled visit. If you examine the data closer, you can see the current last analysis visit (AVISIT) is not the same as the last collected visit (VISIT). In Table 5.1, the last available analysis value now AVISIT is “End of Study” instead of “Month 3”; and in Table 5.2, the last available analysis value now AVIST is also “End of Study” instead of “Month 2” because “LOCF” has been applied to both derived records. The important concept here is that the DTYPE “Endpoint” has to be consistent with the last available AVISIT value if the “LOV” is used for the terminology.

Table 5.3: There's no missing value, End of Study visit is the last non-missing post-baseline value

| USUBJID | PARAM            | VISITNUM | VISIT        | AVISITN | AVISIT       | AVAL | LVOTFL | DTYPE    |
|---------|------------------|----------|--------------|---------|--------------|------|--------|----------|
| 0003    | Pulse Rate (bpm) | 1        | Month 1      | 1       | Month 1      | 60   |        |          |
| 0003    | Pulse Rate (bpm) | 2        | Month 2      | 2       | Month 2      | 70   |        |          |
| 0003    | Pulse Rate (bpm) | 3        | Month 3      | 3       | Month 3      | 80   |        |          |
| 0003    | Pulse Rate (bpm) | 4        | End of Study | 4       | End of Study | 90   | Y      |          |
| 0003    | Pulse Rate (bpm) |          |              | 99      | Endpoint     | 90   |        | Endpoint |

**Table 5.3 No missing values**

“LOCF” is also suggested not to be used for the Endpoint DTYPE as you can see that some subjects, like “0003” in Table 5.3, have their trial completed successfully and there’s no need to apply LOCF at all. If “LOCF” is assigned to the DTYPE for subjects “0001” in Table 5.1 and “0002” in Table 5.2, obviously it’s not appropriate for the subjects like “0003” in Table 5.3 who has successfully completed all scheduled visits. Besides, the endpoint itself is not a scheduled visit while LOCF is defined to be used for missing a scheduled visit or scheduled visits. To use “LOCF” here could end up misleading the reviewers to think that all the endpoint values have been derived from LOCF method, which should be avoided.

## CONCLUSION

This paper has presented a review of experiences implementing endpoint DTYPE in ADaM BDS datasets, discussed comprehensive utilization of DTYPE terminology such as “LOV”, “LOCF” and “Endpoint”. The examples table 5.1, Table 5.2 and Table 5.3 have not been covered in ADaM IG however they have been developed within the scope of ADaM IG. In the development of ADaM BDS endpoint values, when both “LOV” and “LOCF” are involved, the best presentation of DTYPE is to use “Endpoint” which has been proved to be the accurate way to express the endpoint derivation value and to benefit the reviewers for less confusion.

The purpose of this paper is focused on the discussion of the endpoint DTYPE terminology per CDISC ADaM IG. The details regarding the concept of what is CDISC ADaM are not covered in this paper.

## REFERENCES

- [1] Analysis Data Model (ADaM) Implementation Guide Version 1.0 < CDISC Analysis Data Model Team >, pages 47, 48.
- [2] Analysis Data Model (ADaM) Implementation Guide Version 1.0 < CDISC Analysis Data Model Team >, page 35.
- [3] Analysis Data Model (ADaM) Implementation Guide Version 1.0 < CDISC Analysis Data Model Team >, page 06.
- [4] Analysis Data Model (ADaM) Implementation Guide Version 1.0 < CDISC Analysis Data Model Team >, page 46.
- [5] Analysis Data Model (ADaM) Implementation Guide Version 1.0 < CDISC Analysis Data Model Team >, page 70.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Maggie Ci Jiang  
 Enterprise: Teva Pharmaceuticals  
 Address: 905 Airport Rd  
 City, State, ZIP: West Chester, PA 19380  
 Work Phone: 610-883-5790  
 E-mail: Maggie.jiang@tevapharm.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.