

## Timing Variables In Clinical Trials: Avoiding Common Mistakes And Dealing With Unforeseen Issues

Gregory Weller, Rho<sup>®</sup>

### ABSTRACT

Timing variables are an often overlooked part of clinical trials data collection and mapping, particularly in the SDTM and ADaM standards. However, failure to properly plan timing variable collection and mapping before study data is collected can have major consequences later on.

This paper will review good practices for collecting and mapping timing information in clinical trials, from clinical database setup through SDTM and ADaM. Examples from real studies will be used to highlight the challenges that can happen when either: a) insufficient timing information is collected, or b) unforeseen issues arise in the collection of timing variables.

### INTRODUCTION

In clinical trials, data collection focuses primarily on the results of the tests and examinations performed. This includes everything from efficacy data relating to endpoints, to safety data such as vital signs and adverse events. While it is very important to make sure that data is accurate, it is also critically important to describe precisely when a measurement was taken or an event happened. Too often the collection and mapping of timing variables is not properly planned from the beginning of a study, and this can lead to major problems later on.

The SDTM standard offers many different timing variables to describe when information is collected or when events happen. The most frequent issues arise in the most commonly used variables: VISITNUM and the time components of --DTC/--STDTC/--ENDTC.

### VISIT AND VISITNUM

In the experience of this author, no timing variables have provided more problems than the mapping of VISIT and VISITNUM in SDTM datasets. In a clinical trial, the protocol always provides a visit schedule of when subjects need to come in for assessments, and what exact assessments they will have completed at that time. For example, in Table 1 below you can see that at the "Screening" visit a subject will need to: sign informed consent, be assessed for whether they meet the inclusion criteria, have vital signs taken, etc. Assuming they pass the screening criteria, they will then come in for the "Baseline", "Week 1", etc. visits. VISIT and VISITNUM provide a way to identify which visit an observation was taken at. VISIT is usually populated with text similar to what is in the protocol (e.g. "Week 1"), and VISITNUM is an ordinal 1-1 mapping with VISITNUM.

|                              | Screening | Baseline | Week 1 |
|------------------------------|-----------|----------|--------|
| Informed Consent             | X         |          |        |
| Inclusion/Exclusion Criteria | X         |          |        |
| Vital Signs                  | X         | X        | X      |
| Medical History              | X         |          |        |
| Study Drug Dosing            |           | X        |        |
| ...                          |           |          |        |

Table 1: Example Visit schedule

The TV dataset provides the mapping of VISITNUM and VISIT for all planned visits, and the SV dataset provides a record of what visits actually existed.

It is important to note that like most SDTM variables, information about VISIT and VISITNUM should only come from the CRF itself. This may differ from the visit that an observation is assigned for analysis purposes, which are the variables AVISIT and AVISITN in ADaM datasets. These are usually derived based on study day windowing rules, and while there AVISIT is usually the same as VISIT there are usually numerous observations within a study where that is not the case. For example, if the rule in the protocol for a "Week 4" visit states it must happen at study day 28 +/- 3 days, a subject who comes in for this visit at day 32 will have VISIT = "Week 4" in the SDTM datasets since that is what is recorded. However, since this occurred outside the acceptable window AVISIT in the ADaM datasets might be something other than "Week 4". Too often programmers and statisticians are tempted to treat the CRF-reported visit information using analysis visit rules, but it should be kept in mind that the goal of SDTM is to map the clinical data "as-is" and analysis visits should be reserved for ADaM datasets.

The primary issue that arises in mapping VISIT and VISITNUM occurs when dealing with unscheduled visits. Occasionally a subject will need to come in for an assessment outside of the scheduled protocol-defined visits, and it is important to be able to identify each of these unique occurrences as a separate visit. Below are 3 examples from real studies illustrating how VISIT and VISITNUM are mapped for unscheduled visits, along with an explanation of their pros and cons.

#### VISIT AND VISITNUM EXAMPLE 1

In this example, unscheduled visits have been built into the CRF as separate unique visits between planned visits. Table 2 below describes a snapshot of how this works.

| VISITNUM | VISIT                |
|----------|----------------------|
| ...      | ...                  |
| 4        | Week 4               |
| 4.1      | Week 4 Unscheduled 1 |
| 4.2      | Week 4 Unscheduled 2 |
| 6        | Week 6               |
| ...      | ...                  |

**Table 2: Unscheduled Visits Built Into the CRF**

If a subject had their "Week 4" visit but needed to come in for an unscheduled visit before the next planned visit at Week 6, there is already a "Week 4 Unscheduled Visit 1" built into the CRF for the investigators to fill out. Here, there are always two unscheduled visits built into the CRF between each set of planned visits. This setup is ideal since it provides a highly detailed description of when an unscheduled visit actually occurred. The major downside to this setup is that it will take a lot more work to build into the CRF for a feature that might not be used frequently.

#### VISIT AND VISITNUM EXAMPLE 2

In this example, there are 6 planned study visits which are given VISITNUM values of 1 to 6. When a subject comes in for an unscheduled visit, the investigators trigger an "Unscheduled Visit 1" form on the CRF. When a subject comes in for another unscheduled visit, there is an "Unscheduled Visit 2" form triggered for the investigator to fill out, and so on. VISITNUM for these unscheduled visits is mapped as 101, 102, 103, etc. Table 3 below illustrates this mapping schema.

| VISITNUM | VISIT               |
|----------|---------------------|
| 1        | Screening           |
| 2        | Baseline            |
| ...      | ...                 |
| 6        | Week 20             |
| 101      | Unscheduled Visit 1 |
| 102      | Unscheduled Visit 2 |

**Table 3: Unscheduled Visits Generated As Needed**

This method of mapping unscheduled visits still provides a way to identify them as unique occurrences, but requires far less effort to set up the CRF than in example 1. The only downside is that VISITNUM will not provide a true ordinal sort of the data in the order it actually occurred, but in the author's opinion this is more of an aesthetic feature which is secondary to being able to uniquely identify an unscheduled visit.

### VISIT AND VISITNUM EXAMPLE 3

In this example, the client set up the CRF so that any observation recorded at an unscheduled visit would just be recorded as "Unscheduled" and we assigned VISITNUM=99 for all such observations. Here, there is no way to uniquely identify separate unscheduled visits. For example, if a subject has an unscheduled visit marked in their lab data on day 23 and then has an unscheduled vital signs assessment on day 24, should those be considered separate unscheduled visits or would they be part of the same unscheduled visit that spans multiple days?

There were two approaches considered for mapping VISIT and VISITNUM for this study. The first approach was to keep VISITNUM=99 and VISIT = "Unscheduled" for all unscheduled observations. Per the SDTM Implementation Guide v3.2 section 4.1.4.5, this approach should be avoided. The second approach, which is the approach that was eventually adopted, was to create pseudo VISITNUM values for each unique date of an unscheduled visit for a subject. The first unique date for a subject was assigned VISITNUM = 99.01, the second date was given 99.02, etc.

Overall, the approach highlighted in example 2 is likely to be optimal for most clinical trials. Even though the SDTM data sorted by VISITNUM will not be in the same order as if it were sorted by date, this will ensure that all unscheduled visits are uniquely and properly recorded. The approach in example 1, while ideal, the additional benefit derived is unlikely to justify the added time and money to set up the CRF for many companies. The approach highlighted in example 3 should be avoided at all costs.

## TIME COMPONENTS OF DATES

Whenever possible, the time component of an observation should be collected in addition to the date for any on-study assessments. Time is a critical piece of information for everything from baseline flag derivation to assessing whether procedures are done when the protocol specifies they should be. Dealing with time values, however, can leave a study team with some unexpected challenges as highlighted in the two examples below.

### TIME COMPONENTS EXAMPLE 1

A clinical trial is conducted at ten sites in the USA, and the study drug is in the form of an injection administered by the investigator at the study site. A subject who meets the inclusion criteria is randomized, then administered the study drug shortly thereafter. The date/time of randomization comes from the IWRS, which was using Greenwich Mean Time (GMT). The date/time of study drug injection is recorded in the CRF by the investigator using the time local to the site. The result is that all subjects are reported as starting treatment before they are randomized (since GMT is at least four hours ahead of any USA time), which is a situation that should never happen. This oversight is not caught until approximately 10 subjects are randomized when the safety monitors notice that all subjects are reported as having taken study drug several hours before they were randomized.

## TIME COMPONENTS EXAMPLE 2

Subjects are provided an electronic diary so they can record their level of pain in-between study visits. Every morning and evening, the subjects answer a few questions along the lines of "On a scale of 1 to 10 how are you feeling?" This study is also conducted only at sites in the USA, and the information received by the company administering the diary was recorded using that subject's local time. What is not accounted for is the fact that subjects can travel to different time zones, sometimes crossing many time zones in the process. This leads to situations such as "morning" observations being recorded at 3:00 p.m. and "evening" observations at 4:00 a.m.

In each of the examples above, the correct local times of the observations of interest are able to be recovered from the clinical data. One potential option to handle Example 1 is to derive the correct randomization time in the SDTM datasets using the GMT time and the known time zone of the site. However, this is sub-optimal due to the potential for error caused by a programmer's lack of knowledge about time zones randomization information should only be pulled in directly from the IWRS system.

## CONCLUSION

The collection and mapping of timing variables in clinical trials is an important piece that needs to be planned from the very beginning. While there will always be unforeseen issues that arise, proper planning can minimize the amount and severity of issues that occur during a trial.

## REFERENCES

CDISC Submission Data Standards Team, 2013-11-26, Study Data Tabulation Model Implementation Guide: Human Clinical Trials Version 3.2: Clinical Data Interchange Standards Consortium, Inc.

## CONTACT INFORMATION

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

Gregory Weller  
Rho, Inc.  
919-595-6436  
greg\_weller@rhoworld.com