ISO 8601 – An International Standard for Date and Time Formats

Shi-Tao Yeh, GlaxoSmithKline, King of Prussia, PA

ABSTRACT

ISO 8601 is an International Standard for date and time representations referenced by W3C® HTML recommendation and adopted by CDISC as the standard for date and time interchange formats.

This paper discusses the elements of the ISO 8601 Standard that are supported by SAS®. Some sample code and output are included in the paper to illustrate how to implement the Standard.

The SAS products used in this paper are SAS BASE® 9.1.3 on a PC Windows® platform.

INTRODUCTION

ISO stands for International Organization for Standardization. ISO 8601, "Data elements and interchange formats – information interchange – Representation of dates and times", is a profile to specify standard representations of date and time. ISO 8601 supersedes the previous ISO standards: 2014, 2015, 2711, 3307, and 4031. ISO 8601 is referenced by the World Wide Web Consortium (W3C) HTML recommendation. Clinical Data Interchange Standards Consortium (CDISC) also adopts ISO 8601 as the standards for date and time interchange formats.

Different countries have different date and time notations. A date notation like 04/05/06 has six different interpretations. This standard helps to avoid confusion in date and time presentations caused by many different notations.

The Standard defines a wide range of notations of dates, times, and time intervals. ISO 8601 provides six features as follows. 1) All values are organized from most to least significant digits. 2) Each value has a fixed number of digits which must be padded with a leading zero. 3) ISO 8601 provides two types of notation format; basic and extended formats. A basic format is a notation with a minimal number of characters. An extended format is a notation with separators to enhance human readability. The standard permits a dash separator between date elements and a colon between time elements. An optional "T" or blank space is allowed between date and time elements. 4) Partial date and time are allowed. Any number of fields may be dropped from the representation, but the least significant fields must be dropped first. 5) It allows an optional 'time zone' indicator. Without a time zone indicator, the context of the value is local time. 6) The standard supports the additional fraction to the smallest time unit, where higher precision is needed.

EXAMPLES OF ISO 8601 NOTATION

This section shows the standard representations of date, time, date/time and period of time formats.

Date Formats

- Calendar date: YYYY-MM-DD (2003-10-01)
- Ordinal date Formats: The day number within a given year; YYYY-vvv (2005-045)
- Week date: YYYY-Wvv-D (2003-W14-3)

Time Formats

- Local time of the day hh:mm:ss (23:57:59)
- Fractional local time of day hh:mm:ss,f (23:57:30,7) (given one decimal place)
- Coordinated Universal Time (UTC), by appending the symbol "Z" without spaces to any of the local time or fractional local time formats given above. (23:50:30Z)
- Offset between local and UTC times: +/- hh:ss (+13:00)

Combined Date/Time Formats

- The symbol "T" is used to separate the date and time parts of the combined representation. YYYY-MM-DDThh:mm:ss (2005-06-02T15:10:16)

Period of Time

- Specific Start and Specific End. A slash "/" is used to separate the two time values. (20050513T161016/20050605T091112)
- Periods of time, no specific start or end. The value starts with "P" and is followed by a list of periods, each appended by a single letter designator: "Y" for years, "M" for months, "W" for weeks, "D" for days, "H" for hours, "M" for minutes, and "S" for seconds. (P12Y7M11DT9H20M8S)
- Period with specific start: (20050319T171016/P12Y5M)
- Period with specific end: (P12Y5M/20050319T171016)

ISO 8601 DEFINED IN CDISC ODM

The CDISC Operational Data Model (ODM) provides a format for representing the study metadata, study data and administrative data associated with a clinical trial. It is written in XML. The current version of the ODM standard is available at http://www.cdisc.org/standards/index.html. The specification of ISO date/time standard is shown as follows:

```
<xs:simpleType name="date">
  <xs:restriction base="xs:date">
    <xs:pattern   value="[0-9][0-9][0-9][0-9]-[0-1][0-9]-[0-3][0-9]"/>
  </xs:restriction>
</xs:simpleType>

<xs:simpleType name="time">
  <xs:restriction base="xs:time">
    <xs:pattern   value="[0-2][0-9]:[0-5][0-9]:[0-5][0-9](\.[0-
9]+)?((\+|-)[0-2][0-9]:[0-5][0-9])?"/>
  </xs:restriction>
</xs:simpleType>

<xs:simpleType name="datetime">
  <xs:restriction base="xs:dateTime">
    <xs:pattern   value="[0-9][0-9][0-9][0-9]-[0-1][0-9]-[0-3][0-9]T[0-
2][0-9]:[0-5][0-9]:[0-5][0-9](\.[0-9]+)?((\+|-)[0-2][0-9]:[0-5][0-9])?"/>
  </xs:restriction>
</xs:simpleType>
```

ISO SUPPORTED BY SAS

SAS supports some of elements of the ISO 8601 standard. Table 1 is a summary of SAS formats and informats which support ISO 8601.

Category	Pre-fix Expression		Post-fix	Time Zone	SAS Statement	Notation
Date	Basic	ND	8601DA	No	Informat	*YYYYMMDD*
	Extended	IS			Format/Informat	*YYYY-MM-DD*
Time	Extended	IS	8601LZ	Yes	Format/Informat	*hh:mm:ss*[*.fffff*][Z][[+|-]*hh:mm*]
Time	Basic	ND	8601TM	No	Informat	*hhmmss*
	Extended	IS			Format/Informat	*hh:mm:ss*[*.fffff*]

Time	Basic	ND	8601TZ	Yes	Informat	*hhmmss*[.*fffff*][Z][[+|-]*hhmm*]
	Extended	IS			Format/Informat	*hh:mm:ss*[.*fffff*][Z][[+|-]*hh:mm*]
Date/ time	Basic	ND	8601DN	No	Informat	*YYYYMMDD*
	Extended	IS			Format/Informat	*YYYY-MM-DD*
Date/ time	Basic	ND	8601DT	No	Informat	*YYYYMMDDThhmmss*[.*fffff*]
	Extended	IS			Format/Informat	*YYYY-MM-DDThh:mm:ss*[.*fffff*]
Date/ time	Basic	ND	8601DZ	Yes	Informat	*YYYYMMDDThhmmss*[.*fffff*][Z][[+|-]*hhmm*]
	Extended	IS			Format/Informat	*YYYY-MM-DDThh:mm:ss*[.*fffff*][Z][[+|]*hh:mm*

Table 1. SAS formats and informats support ISO standard.

ISO 8601 does not require dashes to separate date parts nor colons to separate time parts; however, the are required for the CDISC SDTM V3.1. This requirement eliminates the usage of SAS ISO Basic formats in all clinical study settings.

IMPLEMENT SAS ISO FORMATS

SDTM V3.1 has eliminated SAS numeric date/time variables and uses ISO 8601 character date string. In order to use dates for computations or graphical presentation, SAS extended ISO formats are used. The following sample codes illustrate how to use the SAS extended ISO formats.

Sample Code I: Combining Data and Time Elements

SAS requires a full, valid date. SAS treats partial date and missing date or time element as invalid data. Sample code I below shows an input data with date and time elements.

```
data w1 ;
      input year month day hour temp min sec;
      datetime = dhms( mdy( month, day, year ),
hour, min, sec );

   datalines;
   05 10 16 21  61 11 15
   04 10 17  0  56 09 03
   06 10 17  3  53 38 40
   98 10 17  6  54  .  .
   99 10 17  9  65  .  .
   02 10 17 12  72 51 55
   ;
data w1;
  set w1;
   format datetime is8601dt.;
   keep datetime;
proc print ;run;
```

The output from Sample code I is shown as follows:

Obs	datetime
1	2005-10-16T21:11:15
2	2004-10-17T00:09:03
3	2006-10-17T03:38:40
4	.
5	.
6	2002-10-17T12:51:55

Sample Code II: Expression of Date

This sample code uses date element only.

```
data w1 ;
  input evtdate is8601da.;

datalines;
2006-08-20
2006-06-11
2006-07
2006
   ;
data w2;
  set w1;
  informat evtdate is8601da.;
  format evtdate is8601da.;

proc print ;run;
```

The output is shown below.

Obs	evtdate
1	2006-08-20
2	2006-06-11
3	.
4	.

Sample Code III: Computation of Duration

This sample code defines date field as character field for displaying the text string of the date. SAS IS8601DA format is used to define the same data field as numeric field for computation of duration between start date and end date.

```
data w1 ;
  input @1 aestdt is8601da. @1 aestdtc $10.
        @12 aeeddt is8601da. @12 aeeddtc $10.;
  datalines;
2006-03-20 2006-04-16
2006-02-11 2006-04-11
2006-01    2006-02-22
2006       2006-01-30
   ;
 data w2;
  set w1;
  duration = aeeddt - aestdt + 1;
 run;
proc print ;run;
```

The output is shown below.

Obs	aestdt	aestdtc	aeeddt	aeeddtc	duration
1	16880	2006-03-20	16907	2006-04-16	28
2	16843	2006-02-11	16902	2006-04-11	60
3	.	2006-01	16854	2006-02-22	.
4	.	2006	16831	2006-01-30	.

CONCLUSION

ISO 8601 provides unambiguous representation of date and time formats. SAS extended ISO formats support these standards. SAS requires a full time element. If a record field involves partial date or time, an extra character field is needed for displaying the partial date or time. The benefits of ISO formats are:

- Adopted by CDISC 3.1 and W3C as date/time formats.
- Platform and language independent.
- Allows reduced precision (partial date/time).
- Larger units are written in front of smaller units for consistency.
- Easily readable, writeable, comparable, and sortable.
- Avoid confusion; such as 01/02/03.
- Allows basic format (20060607) and extended format (2006-06-07).

REFERENCES

[1] CDISC website: http://www.cdisc.org/

[2] CDISC: *ODM 1.2.1* http://www.cdisc.org/standards/index.html

[3] ISO: *Numeric representation of Dates and Time*
http://www.iso.org/iso/en/prods-services/popstds/datesandtime.html

[4] Korpela J*: Info on ISO 8601, the date and time representation standard.*
http://www.cs.tut.fi/~jkorpela/iso8601.html

[5] Kuhn M.: *A summary of the international standard date and time notation.*
http://www.cl.cam.ac.uk/~mgk25/iso-time.html

[6] SAS Institute Inc.: *SAS 9.1 XML LIBNAME Engine User's Guide,* SAS Institute Inc., Cary NC, USA. 2004

[7] Wolf M. and C. Wicksteed: *Date and Time Formats*. http://www.w3.org/TR/NOTE-datetime

[8] W3C: *Use International date format (ISO)*; http://www.w3.org/QA/Tips/iso-date

TRADEMARKS

AUTHOR CONTACT INFORMATION

Shi-Tao Yeh, Ph. D.
(610)787-3856 (W)
E-mail: shi-tao.yeh@gsk.com