

SAS® PROC CDISC Model=ODM Anwendung und Limitierungen

Irene Droll	Elena Valkanova
XClinical GmbH	Biostat International, Inc.
Siegfriedstraße 8	14506A University Point Place
80803 München	Tampa, Fl, 33613
id@xclinical.com	evalkanova@biostatinc.com

Zusammenfassung

Der CDISC ODM (Operational Data Model) Standard hat sich neben anderen bekannten und in der klinischen Forschung verwendeten CDISC Datenstandards bewährt. Der ODM Standard ist – wie jeder XML Standard – ein plattformunabhängiges Format für den Austausch elektronischer Daten aus verschiedenen Datenquellen, wie zum Beispiel CRFs (case report forms), elektronischen Patiententagebüchern oder administrativen Daten mit Informationen zu Prüfzentren. Der ODM Standard definiert die Struktur der CRF Metadaten, als auch die erfassten klinischen Daten (=Snapshot Daten) sowie die zu jedem erfassten Datenpunkt gehörenden Audit Trial Einträge (=Transaktionale Daten).

In diesem Beitrag konzentrieren wir uns auf den Import einer ODM XML Snapshot Datei in SAS unter Anwendung von PROC CDISC Model=ODM (ab SAS 8.2). Der Fokus liegt auf den Parametern und Optionen von PROC CDISC Model=ODM und gibt Hinweise zu den bekannten Limitierungen von PROC CDISC, die zur Generierung von vollständigen und korrekten SAS Datasets berücksichtigt werden sollten. Der Beitrag geht außerdem auf die Anforderungen ein, welche die ODM XML Datei erfüllen muss, damit diese von PROC CDISC fehlerfrei eingelesen werden kann. Darüber hinaus wird ein SAS Makro vorgestellt, das mit einem einzigen Aufruf alle SAS Datasets aus der ODM XML Datei generiert.

Als Vorwissen wird ein Verständnis der CDISC ODM Struktur bestehend aus SubjectProtocol, Events, Forms, ItemGroups, Items, Codelists und MeasurementUnits vorausgesetzt.

Schlüsselwörter: CDISC, ODM, PROC CDISC, XML, Import, SAS Clinical Standards Toolkit

1 Einleitung

XML ist *der* Industriestandard für den plattformunabhängigen Datenaustausch. Die SAS Prozedur PROC CDISC wird in der klinischen Forschung hauptsächlich für die Validierung von SAS Datasets im ODM oder SDTM Standard genutzt. Mit dem neuen SAS Clinical Standards Toolkit (CST) wurde das Model=SDTM von PROC CDISC für die Validierung und Erstellung von SAS SDTM Datasets abgelöst.

Um nach dem ODM Standard strukturierte XML Daten in SAS einzulesen und SAS Datasets nach dem ODM Standard zu validieren, ist die Benutzung der SAS Prozedur

PROC CDISC immer noch das Mittel der Wahl [1]. Die Validierung der Daten gegen deren ODM Metadaten ist in die Prozedur integriert. Durch die Kombination verschiedener Parameter hat der Anwender die Möglichkeit, bestimmte Informationen aus der ODM XML-Datei zusätzlich in die zu erzeugenden SAS Datasets zu schreiben. Zum Beispiel können beim Import der XML ODM Datei gleichzeitig klinische als auch administrative Daten (Informationen zum Zentrum und zum Prüfarzt) in die SAS Datasets eingelesen werden. Die entsprechenden Parameter sind im Kapitel 3 aufgelistet. Dort ist auch ein Beispiel-Aufruf der PROC CDISC MODEL=ODM mit häufig verwendeten Optionen für den Import einer XML ODM Datei zu finden.

Beim Import einer XML ODM Datei kann es zu verschiedenen `ERRORS` und `WARNINGS` im SAS Log kommen, welche man von vorne herein vermeiden kann. Ursachen und Lösungen zu diesen SAS Log Meldungen werden ebenfalls behandelt.

Der Beitrag gliedert sich in vier Bereiche:

- PROC CDISC Model=ODM und SAS Datasets
- PROC CDISC Statement Parameter
- Anforderungen an die externe ODM XML Snapshot Datei
- SAS Makro zum Import einer ODM XML Snapshot Datei

2 PROC CDISC Model=ODM und SAS Datasets

Die ODM XML Datei, die unter Verwendung von PROC CDISC in SAS eingelesen werden soll, muss sowohl klinische Daten (Snapshot) als auch die CRF Metadaten enthalten.

Falls die CRF Metadaten während der Datenerfassung um ein oder mehrere Amendments erweitert wurden und die Studiendaten nun in unterschiedlichen Metadatenversionen vorliegen, müssen alle in den Daten vorhandenen Metadatenversionen ebenfalls in der XML ODM Datei enthalten sein.

Im ODM Format liegen die Daten in einer hierarchischen XML-Struktur vor. PROC CDISC horizontalisiert diese Struktur auf ItemGroup Ebene. D.h. jede *ItemGroup* (= inhaltlich zusammengehöriger Datenblock) in den XML ODM Metadaten wird in einem SAS Dataset abgebildet.

Abbildung 1 zeigt die Struktur der ODM Metadaten.

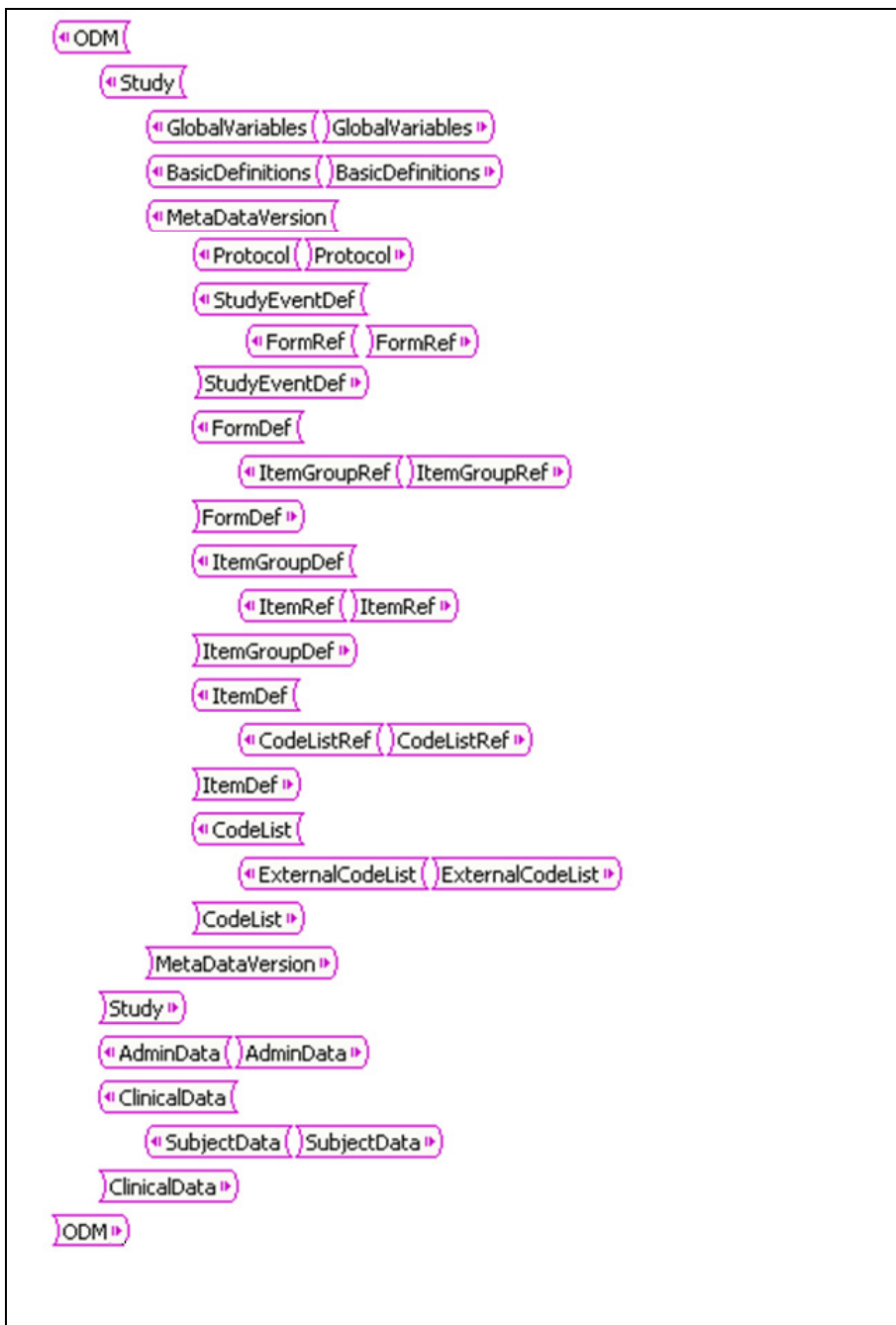


Abbildung 1

Der Name der zu erstellenden SAS Datasets und SAS Variablen (Spalten) kann in den optionalen SAS Attributen der XML ODM Metadaten angegeben werden:

- ODM Attribut **SASDatasetName** bei der Metadatendefinition einer *ItemGroup*:

```

- <ItemGroupDef OID="VISIT" Name="Visit" Repeating="No" SASDatasetName="VISIT">
  <ItemRef ItemOID="VISITDT" OrderNumber="1" Mandatory="Yes" />
</ItemGroupDef>

```

Abbildung 2

- ODM Attribut **SASFieldName** bei der Metadatendefinition eines *Items*:

```
- <ItemDef OID="VISITDT" Name="Visitdate" DataType="date" Length="10" SASFieldName="VISDT">  
- <Question>  
  <TranslatedText xml:lang="en">Visit Date</TranslatedText>  
</Question>  
</ItemDef>
```

Abbildung 3

- ODM Attribut **SASFormatName** bei der Metadatendefinition einer *Codelist*:

```
- <CodeList OID="Gender" Name="Gender" DataType="text" SASFormatName="$SEX">  
- <CodeListItem CodedValue="M">  
  - <Decode>  
    <TranslatedText xml:lang="en">Male</TranslatedText>  
  </Decode>  
</CodeListItem>  
- <CodeListItem CodedValue="F">  
  - <Decode>  
    <TranslatedText xml:lang="en">Female</TranslatedText>  
  </Decode>  
</CodeListItem>  
</CodeList>
```

Abbildung 4

Die in den ODM Metadaten definierten SAS Attribute werden von PROC CDISC als Dataset-, Variablen- und Formatnamen benutzt. Falls diese optionalen SAS Attribute in den ODM Metadaten nicht definiert wurden, verwendet SAS PROC CDISC die ersten acht Zeichen des optionalen Attributes **Name**. Wenn auch das Attribut **Name** in den Metadaten nicht definiert wurde, verwendet SAS PROC CDISC die ersten acht Zeichen des Attributes **OID** (Object Identifier)[2].

Obwohl die OID in den ODM Metadaten eindeutig definiert sein muss, kann durch das Abschneiden auf die ersten acht Zeichen der eindeutige SAS Dataset-, Variablen- oder Formatname verloren gehen. Dadurch kann es beim Einlesen der ODM XML Datei zum Überschreiben von Datasets, Variablen und Formaten kommen, worauf SAS mit einem `ERROR` im Log hinweist. Außerdem können die SAS Attribute schon mit CDASH bzw. SDTM Domain- und Variablennamen hinterlegt werden, so dass eine spätere Generierung der SDTM Datasets erleichtert wird.

Für jede *ItemGroup*, die in den ODM Metadaten definiert wurde, wird ein SAS Dataset erzeugt. Um die Anzahl Datasets und somit die Anzahl benötigter SAS *merge* Schritte für eine Studie zu minimieren, kann man sich an folgende Tipps halten:

- Wenn möglich, *ItemGroups* als Repeating definieren, z.B. CM (concomitant medication)
- Wenn möglich, *ItemGroup* Definitionen in verschiedenen Events/Visiten wiederverwenden, z.B. VS (vital signs)
- Wenn sinnvoll, alle inhaltlich zusammengehörenden Items in eine *ItemGroup* zusammenfassen, z.B. AE (adverse events)

Abbildung 5 zeigt das Beispiel einer ODM Metadaten-Definition für ein CRF (links), den dazugehörigen Screenshot aus einem EDC User Interface mit eingegebenen klinischen Daten (Mitte) und die daraus mit PROC CDISC generierten SAS Datasets sowie den dazugehörigen SAS Format Katalog (rechts).

Abbildung 5 zeigt die ODM Metadaten-Definition für ein CRF (links), den dazugehörigen Screenshot aus einem EDC User Interface mit eingegebenen klinischen Daten (Mitte) und die daraus mit PROC CDISC generierten SAS Datasets sowie den dazugehörigen SAS Format Katalog (rechts).

ODM Metadaten-Definition (links):

- Screening*
 - Basedata*
 - Visit* - SAS: "VISIT"
 - Visitdate* - SAS: "VISDT" (date 10)
 - Demographics* - SAS: "DM"
 - Gender* - SAS: "SEX" (text 1)
 - Gender - SAS: "\$SEX" (text)
 - Birthdate* - SAS: "BRTHDT" (date 10)
 - Height* - SAS: "HGHT" (integer 3)
 - inch
 - cm
 - Vitals* - SAS: "VS"
 - Systolic* - SAS: "SYSBP" (integer 3)
 - Diastolic* - SAS: "DIABP" (integer 3)
 - Concomittant* - SAS: "CM"
 - Startdate* - SAS: "STARTDT" (date 10)
 - Enddate - SAS: "STOPDT" (date 10)
 - Term* - SAS: "CMTERM" (text 100)
 - Dose* - SAS: "DOSE" (float 5,2)
 - Unit* - SAS: "UNIT" (integer 1)
 - Unit - SAS: "UNIT" (integer)

EDC User Interface Screenshot (Mitte):

Visit	
Date of Visit	01/07/2009

Demographics	
Gender	male
Birthdate	15/07/1969
Height	180 cm

Vital Signs	
Systolic Blood Pressure	120
Diastolic Blood Pressure	80

Concomittant Medication				
Startdate	Enddate	Term	Dose	Unit
01/02/2008	01/07/2009	Aspirin	200	mg
01/07/2009		Paracetamol	100	mg

SAS Datasets und Format Katalog (rechts):

- Visit
- Dm
- Sasmacr
- Vs
- Cm
- Formats

Abbildung 5

3 PROC CDISC Statement Parameter

3.1 Technische Voraussetzungen

- Betriebssystem: Windows, UNIX und z\OS.
- SAS/Base – ab Version 8.2.
- Installation der neuesten PROC CDISC Version 2.15.62. (PROC CDISC version; RUN; gibt die aktuell installierte Version im SAS Log aus)

Anmerkung: Falls eine ältere Version von PROC CDISC installiert ist, kann es sein, dass manche der hier gelisteten und beschriebenen Parameter nicht verfügbar sind und ein ERROR im SAS Log ausgegeben wird.

3.2 Einlesen einer ODM XML Datei mit PROC CDISC

Jedes ODM Element hat ein OID Attribut, das das Element eindeutig identifiziert. OIDs, die in die SAS Dataset Spalten/Daten geschrieben werden, nennt man auch „key-set fields“. Die meisten dieser „keyset fields“ werden zur genauen Zuordnung eines Wertes zu einem Patienten, einer Visite und einem Formular benutzt: StudyOID, MetaDataVersionOID, SubjectKey, StudyEventOID, StudyEventRepeatKey, FormOID, FormRepeatKey, ItemGroupOID, ItemGroupRepeatKey und TransactionType (diese Variable hat immer den Wert “insert”).

In Tabelle 1 sind alle Statements und Parameter gelistet, die für den Import einer ODM XML Datei in SAS verwendet werden können [3]:

Tabelle 1

PROC CDISC*	Parameter	Wert	Beschreibung/Erklärung
	MODEL*	ODM/ SDTM	Name des angewandten CDISC Modells.
	READ*	FILENAME	Speicherstelle der Quelldatei, die importiert werden soll.
	FORMATACTIVE	YES/NO	Falls YES, wird aus den in den ODM Metadaten definierten Codelisten ein SAS Format Katalog erstellt und den darauf referenzierenden SAS Variablen zugewiesen. Der Standardwert ist NO.
	FORMATNO-REPLACE	YES/NO	Falls YES, werden Formate mit dem gleichen SAS Namen nicht ersetzt. Diese Option kann nur verwendet werden, wenn FORMATACTIVE = YES. Der Standardwert ist NO.
	FORMAT-LIBRARY	LIBRARY	Hiermit wird eine permanente SAS Format Bibliothek angelegt. Diese Option kann nur verwendet werden, wenn FORMATACTIVE = YES.
	LANGUAGE+	EN, DE, etc.	Falls mehr als eine Sprache in den ODM Metadaten definiert ist, kann man mit diesem Parameter die Sprache definieren, in der die SAS Labels generiert werden sollen. Der Defaultwert ist EN.

ODM*	ODMVERSION*	1.2	Version des CDISC ODM Modells.
ODM	ODMMINIMUM-KEYSET	YES/NO	<p>Falls NO, werden alle “keyset fields” (wie oben beschrieben) in das SAS Dataset geschrieben. Falls YES, wird aus den “keyset fields” nur der SubjectKey in das SAS Dataset geschrieben;</p> <p><u>Achtung:</u> Sobald eine ItemGroup Definition von mehreren Stellen in den ODM Metadaten referenziert wird (Verwendung in mehreren Visiten), muss der Parameter auf NO gesetzt werden.</p> <p>Der Standardwert ist NO.</p>
ODM	ODMMAXIMUM-OIDLENGTH	Zahl (z.B. 8, 20, ...)	<p>Falls dieser Parameter gesetzt ist, wird die OID aus den ODM Metadaten auf die angegebene Zeichenanzahl abgeschnitten.</p> <p>Falls dieser Parameter nicht gesetzt ist, kann die OID bis zu 100 Zeichen lang sein.</p>
ODM	USERNAMEASLABEL+	YES/NO	<p>Falls YES, wird das ODM Attribut <i>Name</i> als SAS Variablen Label benutzt.</p> <p>Der Standardwert ist NO.</p>
ODM	LONGNAMES+	YES/NO	<p>Falls YES, wird anstatt der ODM SAS Attribute das ODM Attribut <i>Name</i> als SAS Variablenname verwendet. Der SAS Variablenname wird auf 32 Zeichen gekürzt, Leerzeichen werden mit Unterstrichen (_) ersetzt.</p> <p>Der Standardwert ist NO.</p>
ODM	ORDERNUMBER	YES (USE) / NO (IGNORE)	<p>Falls YES, werden die ODM OrderNumber Attribute der Item-Referenzierungen auf Vollständigkeit und Zulässigkeitsbereich validiert. (z.B. wird ein ERROR im SAS Log angezeigt, wenn die OrderNumber die tatsächliche Anzahl Items in der <i>ItemGroup</i> überschreitet)</p> <p>Der Standardwert ist YES.</p>

CLINICALDATA *	OUT	TEXT	Name des zu erstellenden SAS Datasets.
CLINICALDATA	SASDATASET- NAME	TEXT	Der Wert dieses Parameters muss dem ODM <i>SASDatasetName</i> entsprechen.
CLINICALDATA	NAME+	TEXT	Falls LONGNAMES=YES verwendet wird, muss der Name, der hier angegeben wird, dem ODM <i>Name</i> Attribut der <i>ItemGroup</i> entsprechen.
CLINICALDATA	SITEREf+	YES/NO	Falls YES, wird eine zusätzliche SAS Variable "LocationOID" mit der ID des Zentrums in das SAS Dataset geschrieben.
CLINICALDATA	INVESTIGATOR REF+	YES/NO	Falls YES, wird eine zusätzliche SAS Variable "UserOID" mit der ID des Prüfarztes in das SAS Dataset geschrieben.
CONTENTS+			Mit dieser Option werden die Contents des erzeugten CDISC ODM SAS Datasets in das SAS Log ausgegeben.
DATASETS+			Diese Option gibt eine Auflistung der in der ODM XML Datei definierten SAS Datasets in das SAS Log aus.

Note: + Optionaler Parameter, * Erforderlicher Parameter

3.3 Beispiel

In Abbildung 6 wird das Beispiel eines PROC CDISC Aufrufs gezeigt, unter Verwendung einiger der Parameter, die in Tabelle 1 aufgelistet sind.

```

1  LIBNAME OUT 'C:\MYPROJECT\DATA\SAS';
2  FILENAME XMLIMP 'C:\MYPROJECT\DATA\XML\MYFILE.XML';

3  PROC CDISC MODEL=ODM
      READ=XMLIMP
      FORMATACTIVE=YES
      FORMATNOREPLACE=NO
      LANGUAGE="EN";

4  ODM      ODMVERSION="1.2"
           ODMMAXIMUMOIDLENGTH=20
           ODMMINIMUMKEYSET=NO
           USENAMEASLABEL=YES;

```



```

5   CLINICALDATA OUT=OUT.DSET
      SASDATASETNAME = "DSET"
      SITEREF=YES
      INVESTIGATORREF=YES;

RUN;

```

Abbildung 6

1. Hier wird das Verzeichnis spezifiziert, in welches das mit PROC CDISC erzeugten SAS Datasets abgelegt werden sollen.
2. Das FILENAME Statement definiert eine Referenz auf den physikalischen Speicherplatz der ODM XML Datei.
3. Der READ Parameter von PROC CDISC MODEL=ODM benutzt diese Referenz.
4. Die ODM Parameter sind so definiert, dass alle "keyset fields" im zu erzeugenden SAS Dataset DSET enthalten sein werden. Die Länge aller "keyset fields" ist mit maximal 20 Zeichen definiert. Wenn der Parameter ODMMAXIMUMOIDLENGTH nicht angegeben wäre, würden die OIDs in den "key fields" erst ab 100 Zeichen (Standardwert) abgeschnitten werden.
5. Die CLINICALDATA Parameter sind in diesem Beispiel so gesetzt, dass das SAS Dataset „DSET“ mit jeweils einer zusätzlichen Spalte für das Rekrutierungszentrum und die Prüfarztinformation erzeugt wird.

4 Anforderungen an die externe ODM XML Snapshot Datei

Um ein vollständiges und fehlerfreies Mapping der ODM XML Datei in SAS Datasets zu gewährleisten, sollte die ODM XML Datei auf bestimmte Voraussetzungen geprüft werden.

Die Auswirkungen dieser „Fehler“ können in drei Gruppen eingeteilt werden:

- ODM XML Code Teile, die im SAS Log ERROR Meldungen ausgeben.
- ODM XML Code Teile, die im SAS Log WARNING Meldungen ausgeben.
- ODM XML Code Teile, die beim Import mit PROC CDISC nicht berücksichtigt werden.

In Tabelle 2 und Tabelle 3 sind häufige SAS ERRORS und WARNINGS mit entsprechenden Erklärungen gelistet.

Tabelle 2

Error message in SAS Log	Reason
<p>ERROR: Invalid member name for file WORK.CON.DATA</p>	<p>Die SAS Attribute in den ODM Metadaten dürfen keine Namen verwenden, die vom Betriebssystem verwendet werden. Diese sind z.B.: AUX, CON, NUL, PRN, LPT1 - LPT9, COM1 - COM9.</p>
<p>ERROR: ItemGroupDef OID=ig.PSA has no matching ItemGroupRef.</p>	<p>ODM Elemente (<i>ItemGroups</i> oder <i>Items</i>), die an keiner Stelle referenziert werden, müssen aus den ODM XML Metadaten entfernt werden.</p>
<p>ERROR: ItemData ItemOID = "i.dummy". Corresponding ItemDef not found.</p>	<p>SAS PROC CDISC kann nicht mit „impliziten XML closing tags“ umgehen, sondern benötigt „explizite XML closing tags“. D.h. <ItemGroupName OID="ig.test"/> führt zu einem ERROR im SAS Log, obwohl dies eine korrekte XML Syntax ist. Die XML Syntax, mit der PROC CDISC umgehen kann, muss <ItemGroupNameOID="ig.test"></ItemGroupName> lauten. Automatisiert kann man dies mit einer XSL Transformation der ODM XML Datei erreichen.</p>
<p>ERROR: ItemDef "Birthdate" has an invalid Length attribute value. ERROR: 1 ItemDef elements had incorrect Length or SignificantDigits attributes</p>	<p>Items mit Datentyp = "date" müssen in den ODM Metadaten immer mit Länge = 10 definiert sein. Items mit Datentyp = "time" müssen in den ODM Metadaten immer mit Länge = 5 definiert sein. Items mit Datentyp = "text" sollten in den ODM Metadaten mit einer maximalen Länge von 200 definiert sein. Ein Wert der mehr als 200 Zeichen hat wird beim Import nach 200 Zeichen abgeschnitten. Die aktuelle Version von PROC CDISC fügt nicht definierte Nachkommastellen („Significant Digits“) automatisch als Metadatenattribut hinzu.</p>
<p>ERROR: Some code points did not transcode.</p>	<p>Zeichen, die von einer fremden Tastatur eingegeben werden, können ebenfalls zu ERRORS im SAS Log führen, z.B. Kyrillisches C.</p>

Tabelle 3

Warning message in SAS Log	Reason
WARNING: Variable DATE already exists on file WORK.CM	SASFieldNames müssen innerhalb einer <i>ItemGroup</i> (SAS Dataset) eindeutig sein. Bei Duplikaten von SASFieldNames werden die SAS Variablen innerhalb eines SAS Datasets einfach überschrieben.
WARNING: SAS character format names must begin with a \$. SEX changed to SAS compliant format name \$SEX	Wenn eine ODM Codelist Text statt Nummern als Code benutzt, sollte der SASFormatName mit \$ beginnen (e.g. \$SEX). Ansonsten fügt SAS dieses \$ automatisch an den Format Namen und gibt ein WARNING in das SAS Log aus.
WARNING: ItemRef OID="Height" OrderNumber="4" outside range. Ignoring OrderNumber.	Wenn die SAS Option ORDERNUMBER=IGNORE/NO nicht benutzt wird und die <i>OrderNumber</i> die Anzahl der Items in der <i>ItemGroup</i> überschreitet, wird ein WARNING in das SAS Log ausgegeben.

Andere ERRORS, die ein unvollständiges oder fehlerhaftes Mapping der ODM XML Daten in SAS Datasets zur Folge haben, können mit folgenden Punkten zusammenhängen:

- Alle ODM SAS Attribute müssen mit einem Buchstaben oder Unterstrich beginnen (keine Zahlen). Falls das erstes Zeichen des ODM SAS Attributs eine Zahl ist, stellt PROC CDISC einen Unterstrich vor das SAS Attribut und schneidet dafür ein Zeichen ab (aus 8DATE wird somit _8DAT).
- Alle ODM SAS Attribute sollten maximal 8 Zeichen lang sein. Falls der SAS Name länger ist, schneidet PROC CDISC den Namen auf 8 Zeichen ab. Dadurch kann es passieren, dass gleiche SAS Dataset- oder SAS Variablennamen erzeugt und beim Einlesen überschrieben werden.
- PROC CDSIC ignoriert die ODM Metadaten Referenz auf MeasurementUnits. Ein Item kann jedoch auf mehrere MeasurementUnits referenzieren, z.B. das Item Gewicht auf „kg“ und „Pfund“. Das SAS Log gibt keine Hinweise darauf, dass MeasurementUnits von SAS ignoriert wurden. Der Anwender ist also selbst in der Verantwortung die Einheit, in der ein Wert gemessen wurde, an SAS zu übergeben. Dies kann z.B. mit einer XSL Transformation erreicht werden, mit der man für alle Items, die auf eine MeasurmentUnit referenzieren, ein zusätzliches Item und somit eine zusätzliche SAS Variable anlegt, die diese Einheit enthält.
- Wenn ein Item in der ODM XML Datei eine externe Codelist (Konzept der ODM Struktur) referenziert, muss diese externe Codelist entweder manuell in die ODM

XML Metadaten eingefügt werden oder es muss die Option `FORMATOPTION=NO` gesetzt sein und die Codes müssen mit zusätzlichen `PROC FORMAT` Schritten dekodiert werden.

5 SAS Makro zum Import einer ODM XML Snapshot Datei

Mit einem `PROC CDISC` Aufruf wird für eine *ItemGroup* aus der ODM Datei ein SAS Dataset erzeugt. Die ODM Metadaten einer Studie enthalten jedoch viele *ItemGroups*.

Abbildung 7 zeigt ein SAS Makro, mit dem man automatisiert alle *ItemGroups* aus einer ODM XML Datei einlesen kann und das somit alle dazugehörigen SAS Datasets mit einem Aufruf erzeugt.

Dafür wird die SAS Tabelle `DICTIONARY.TABLES` benutzt, die in jeder SAS Session zur Verfügung steht.

```
LIBNAME OUT 'C:\TEST';
FILENAME XMLIMP 'C:\EXAMPLES\TEST.XML';
LIBNAME XMLIMP XML XMLTYPE=CDISCODM;

%MACRO ALLSETS (LIB);
PROC SQL NOPRINT;
SELECT UNIQUE MEMNAME INTO :DSETS SEPARATED BY '|'
  FROM DICTIONARY.TABLES
  WHERE UPCASE(LIBNAME)="&LIB" ;
  %PUT DATASETS: &DSETS;
QUIT;
%LET NUM=1;
%LET DSET=%SCAN(%QUOTE(&DSETS), &NUM, |);
  %DO %UNTIL (%QUOTE(&DSET)=%STR());
    PROC CDISC MODEL=ODM READ=XMLIMP
      FORMATACTIVE=YES
      FORMATNOREPLACE=NO
      LANGUAGE="EN";

      ODM          ODMVERSION="1.2"
                  ODMMAXIMUMOIDLENGTH=20
                  ODMMINIMUMKEYSET=NO
                  USENAMEASLABEL=YES;

      CLINICALDATA OUT=OUT.&DSET
                  SASDATASETNAME = "&DSET"
                  SITEREF=YES
                  INVESTIGATORREF=YES;
```

```
RUN;  
%LET NUM=%EVAL (&NUM+1) ;  
%LET DSET=%SCAN (%QUOTE (&DSETS) , &NUM, |) ;  
%END;  
%MEND ALLSETS;  
  
%ALLSETS (XMLIMP) ;
```

Abbildung 7

6 Fazit

PROC CDISC ist im Moment das einzige Tool, das von SAS[®] zur Verfügung gestellt wird, um Daten im ODM XML Format in SAS einlesen zu können.

Ein korrekter und vollständiger Import der Daten aus einer ODM XML Datei kann erreicht werden, indem man sich der Einschränkungen von PROC CDISC bewusst ist und die ODM XML Metadaten so vorbereitet, dass mögliche `ERRORS` und `WARNINGS` beim Einlesen von vorne herein vermieden werden.

Das vorgestellte SAS Makro hilft dabei, die Daten im ODM XML Format „in einem Aufwasch“ einlesen zu können.

Literatur

- [1] Spezifizierung des Operational Data Models (ODM), CDISC
<http://www.cdisc.org/models/odm/v1.2/ODM1-2-0.html>
- [2] “The CDISC Procedure for SAS[®] Software, Release 8.2 and Later”,
SAS Institute, Inc.
<http://support.sas.com/rnd/base/xmlengine/proccdisc/TW8774.pdf>
- [3] <http://www.sas.com/industry/pharma/cdisc/index.html>

