

# Dokumentation der Herleitung und Vollständigkeit von mit SAS erstellten CDISC SDTM-Datensätzen mit Hilfe des Tools Doxygen

Jörg Sahlmann  
iOMEDICO AG  
Hanferstraße 28  
79108 Freiburg

Joerg.Sahlmann@iomedico.com

Renate Scheiner-Sparna  
iOMEDICO AG  
Hanferstraße 28  
79108 Freiburg

Renate.Scheiner-Sparna@iomedico.com

## Zusammenfassung

CDISC SDTM-Datensätze sind vom Aufbau hochstandardisiert, nichtsdestotrotz gibt es bei der Befüllung der Datensätze häufig Interpretationsspielraum. Um zwischen unterschiedlichen Studien hier eine Einheitlichkeit herzustellen, ist eine gute Dokumentation der SAS-Programme, die die SDTM-Datensätze erstellen, notwendig. Ausgehend von den SDTM-Definitionen und einer Liste der in der elektronischen Patientendokumentation (e-CRF) verwendeten Variablen zeigt dieser Beitrag auf, wie die Dokumentation auf der Basis des Tools doxygen assistiert erstellt werden kann.

Schrittweise wird geprüft,

- ob auf alle eCRF Variablen lesend zugegriffen und die Informationen genutzt werden
- ob allen SDTM-Variablen, die "required" sind, Werte zugewiesen werden

Fehlende Nutzungen und Zuweisungen werden in getrennten Tabellen gespeichert sowie bei der Dateibeschreibung farblich kodiert angezeigt. Vorhandene Zuweisungen können an die passende Stelle der Dokumentation der Variablen kopiert werden.

Es werden weitere Dokumentationsmöglichkeiten mit Doxygen aufgezeigt, z.B. verwaiste Makros, die nicht für die aktuelle Auswertung genutzt werden. Diese Optionen dienen der Verschlinkung der Programmverzeichnisse, um wirklich nur die Dateien vorzuhalten, die für die aktuelle Auswertung gebraucht werden.

**Schlüsselwörter:** CDISC, SDTM, Doxygen

## 1 Was haben wir, was wollen wir?

Wir haben gerade mit der Programmierung einer Studiauswertung begonnen. Uns liegen die aktuellen SAS-Programme, die SDTM-Datensätze, eine Übersicht der SDTM-Variablen und eine Übersicht der im eCRF benutzten Variablen vor. Wir wollen die Herleitung der Variablen dokumentieren und die Datensätze auf Vollständigkeit der Variablen prüfen.

Diese Prüfungen können auch mit Tools wie Pinnacle21 durchgeführt werden, das in vielen Aspekten mehr leistet, aber für unsere Dokumentationszwecke nicht die passenden Schnittstellen bietet.

## **1.1 SDTM-Variablen**

Aktuell verwenden wir CDISC SDTM in der Version 1.4 zusammen mit dem Implementation Guide in der Version 3.2. Von CDISC wurde eine Excel-Tabelle `sdtm-3-2-excel.xls` auf der CDISC-Homepage zum Download angeboten. Diese enthält in tabellarischer Form die Metadaten, wie sie domänenweise in dem Dokument Study Data Tabulation Model Version 1.4 beschrieben werden.

Zur weiteren Verarbeitung wird diese Excel-Tabelle in einen SAS-Datensatz und für die Verarbeitung mit Python in eine SQLite3-Datenbank umgewandelt.

## **1.2 eCRF-Variablen**

Unser eCRF stellt die Daten auf Grundlage des CDISC ODM (Operational Data Model) bereit.

Die entsprechenden Schlüsselvariablen sind:

- ClinicalData – StudyOID
- SubjectData – SubjectKey
- StudyEventData – StudyEventOID
- FormData – FormOID
- ItemGroupData – ItemGroupOID
- ItemData – ItemOID

Als Variablenbezeichnung im engeren Sinne dieses Beitrages wird die ItemOID bezeichnet. Als Variablenbezeichnung im weiteren Sinne wird die Kombination aus ItemGroupOID und ItemOID bezeichnet.

Zur weiteren Verarbeitung werden diese ItemOID und die Kombination von ItemGroupOID und ItemOID genau wie die SDTM-Variablen in einen SAS-Datensatz und in eine SQLite3-Datenbank umgewandelt.

## **1.3 SAS-Programme**

Die Programme für die Studienauswertung sind gemäß ihrem Zweck auf die Ordner IO, SDTM, ADS, Tables, Listings, Figures und andere aufgeteilt.

Die SAS-Programme für die Erstellung der SDTM-Datensätze besitzen den Doxygen-Header und den auskommentierten SAS-Code, wie er bei Bachert [2] beschrieben wird. Der Name der Programme ergibt sich aus dem Präfix `SDTM_` und den zwei Buchstaben der Domäne (z. B. `SDTM_DM.sas`).

## **1.4 SDTM-Datensätze**

Die SDTM-Datensätze sind normale SAS-Datensätze, die sich im Aufbau an den Spezifikationen der SDTM V1.4 orientieren.

Die Studiendaten sind gemäß ihrem Bearbeitungsstatus auf die Dateiodner RawData, SDTM und ADS aufgeteilt.

Der Ordner SDTM enthält alle generierten SDTM-Datensätze. Die Datensätze sind mit den zwei Buchstaben der Domäne benannt (DM.sas7bdat). Die Ergänzungsdatensätze sind entsprechend mit dem Präfix supp und den zwei Buchstaben der Domäne benannt (supp\_DM.sas7bdat).

## 2 Ergebnislisten und ihre Herleitung

Die Ergebnisse der unterschiedlichen Verarbeitungsschritte werden als ASCII-Datei gespeichert und dann als freie Seite in Doxygen eingebunden (Tab: Related Pages) oder in den Header des einzelnen SAS-Programmes integriert.

### 2.1 Definition der SDTM-Variablen

Die Definition der SDTM-Variablen wird mit im Header des entsprechenden SAS-Programmes abgelegt, um auf der zugehörigen Dokumentationsseite für das Programm die Definitionen greifbar zu haben.

Die drei Variablenkategorien **Required**, **Permitted** und **Expected** werden in dieser Liste durch unterschiedliche Farbtöne dargestellt.

**Required:** Eine Variable, die unbedingt notwendig für die Identifikation des Datensatzes ist oder die dem Datensatz seine Bedeutung gibt. Sie muss immer enthalten sein und darf für einzelne Beobachtungen nicht leer sein.

**Expected:** Eine Variable, die einen Datensatz im Rahmen einer bestimmten Domäne Bedeutung gibt. Sie kann für einzelne Beobachtungen leer sein. Wenn sie für alle Beobachtungen leer ist, muss sie dennoch enthalten sein.

**Permitted:** Eine Variable, die genutzt wird, wenn entsprechende Daten gesammelt worden sind. Wenn sie für alle Beobachtungen leer ist, kann sie in der Domäne enthalten sein.

Der Inhalt dieser Liste (Abbildung 1) ergibt sich aus der Datei sdtm-3-2-excel.xls<sup>1</sup>, die auf die entsprechende Domäne gefiltert worden ist.

---

<sup>1</sup> Download von der CDISC Homepage [www.cdisc.org](http://www.cdisc.org) nach Registrierung

INGE-B

Main Page	<b>Related Pages</b>	Modules	Files
-----------	----------------------	---------	-------

**CDISC variables of domain EX**

**Definition of variables according to CDISC**

SDTM-Variable	Label	Core
STUDYID	Study Identifier	Req
DOMAIN	Domain Abbreviation	Req
USUBJID	Unique Subject Identifier	Req
EXSEQ	Sequence Number	Req
EXGRPID	Group ID	Perm
EXREFID	Reference ID	Perm
EXSPID	Sponsor-Defined Identifier	Perm
EXLNKID	Link ID	Perm
EXLNKGRP	Link Group ID	Perm
EXTRT	Name of Treatment	Req
EXCAT	Category of Treatment	Perm
EXSCAT	Subcategory of Treatment	Perm
EXDOSE	Dose	Exp

**Abbildung 1:** Ausschnitt aus der Variablendefinition der Domäne EX

## 2.2 Herleitung der SDTM-Variablen

Die Herleitung der SDTM-Variablen wird mit im Header des entsprechenden SAS-Programmes abgelegt, um auf der zugehörigen Dokumentationsseite für das Programm die Herleitungen nachlesen zu können. Leider können die Herleitungen (noch) nicht vollständig automatisch extrahiert werden, es werden aber Programmteile als Schnipsel gesucht und angeboten. Für alle nachfolgend beschriebenen Herleitungsarten gilt:

Die Suche überspringt den Header und fängt im auskommentierten SAS-Code an.

Die Ergebnisse zu einem Variablennamen werden jeweils zusammengefasst, Dubletten werden gelöscht.

Aus der Zusammenfassung dieser Vorschläge kann dann die eigentliche Herleitung ermittelt werden.

### 2.2.1 Zuweisen einer Konstanten an eine SDTM-Variable

Die einfachste Variante der Herleitung einer SDTM-Variablen ist die Zuweisung eines konstanten Wertes (String oder Zahl).

Hierzu werden die Programme nach dem Variablennamen, gefolgt von keinen bis mehreren Leerzeichen, gefolgt von einem Gleichheitszeichen, gefolgt von keinem bis meh-

renen Leerzeichen, gefolgt von einer Stringkonstanten oder einer numerischen Konstanten durchsucht und die Suchergebnisse werden als Liste in den Header geschrieben. Die Beschreibung dieser Suche drängt einem geradezu einen regulären Ausdruck auf. So wird es dann mit Python auch umgesetzt.

Für eine numerische Konstante wird der reguläre Ausdruck

```
\s*SDTMVAR\s*=\s*(\-[0,1][0-9]+\.[0,1][0-9]*\s*);
```

benutzt, wobei SDTMVAR durch den entsprechenden Variablennamen ersetzt wird.

Aus der Codezeile

```
CMSEQ = 100;
```

wird der entsprechende Eintrag für CMSEQ in der Kategorie 1\_NUMKONST eingetragen (siehe vierte und fünfte Tabellenzeile in Abbildung 2).

Für eine Stringkonstante wird der reguläre Ausdruck

```
^\s*SDTMVAR\s*=\s*["\']([a-z0-9_\-/\ ]+)\s*["\']
```

benutzt, wobei SDTMVAR wiederum durch den entsprechenden Variablennamen ersetzt wird.

Aus der Codezeile

```
STUDYID = "INGE_B";
```

wird der entsprechende Eintrag für STUDYID in der Kategorie 2\_STRKONST eingetragen (siehe erste Tabellenzeile in Abbildung 2).

## Specification of variables according to current scripts

Domain	SDTM-Variable	Category	Specification	Core
CM	STUDYID	2_STRKONST	INGE_B	Req
CM	DOMAIN	2_STRKONST	CM	Req
CM	USUBJID	7_ELSEWHERE	Possible hints see below	Req
CM	CMSEQ	1_NUMKONST	100	Req
CM	CMSEQ	1_NUMKONST	200	Req
CM	CMSEQ	4_CALCULATION	100 + INPUT(ID_GL_IGSEQUNO, 3.)	Req
CM	CMSEQ	4_CALCULATION	200 + INPUT(ID_GL_IGSEQUNO, 3.)	Req
CM	CMGRPID	7_ELSEWHERE	Possible hints see below	Perm
CM	CMSPID	7_ELSEWHERE	Possible hints see below	Perm
CM	CMTRT	3_VARIABLES	ID_CM_CMPRIORHORMTRT	Req
CM	CMTRT	3_VARIABLES	ID_CM_CMPRIORANTRACHEMTRT	Req

**Abbildung 2:** Auszug aus der Kategorisierung von Ableitungen der Domäne CM

Für Abwandlungen, die if-Bedingungen oder else if-Bedingungen enthalten, gibt es die folgenden regulären Ausdrücke.

```
^\s*(if\s*[a-z0-9_=\"\']+\s*then\s*SDTMVAR\s*=\s*[\\"'\][a-z0-9_\-/ ]+[\\"'\'])  
^\s*else\s*(if\s*[a-z0-9_=\"\']+\s*then\s*SDTMVAR\s*=\s*[\\"'\][a-z0-9_\-/ ]+[\\"'\'])
```

Für numerische Konstanten und Stringkonstanten wird angenommen, dass jeweils eine Anweisung in einer Codezeile steht.

### 2.2.2 Zuweisen einer Variablen an eine SDTM-Variable

Einer SDTM-Variablen kann auch direkt eine andere Variable zugewiesen werden.

Hierfür wird der reguläre Ausdruck

```
^\s*SDTMVAR\s*=\s*([a-z0-9_]+);
```

benutzt, wobei SDTMVAR wiederum durch den entsprechenden Variablennamen ersetzt wird.

Aus der Codezeile

```
CMTRT = ID_CM_CMPRIORANTRACHEMTRT;
```

wird der entsprechende Eintrag für CMTRT in der Kategorie 3\_VARIABLES eingetragen (siehe letzte Tabellenzeile in Abbildung 2).

### 2.2.3 Zuweisen einer Berechnung an eine SDTM-Variable

Eine weitere Variante der Herleitung einer SDTM-Variablen ist die Zuweisung einer Berechnung.

Hierzu werden die Programmteile gesucht, bei denen die SDTM-Variable links von einem Gleichheitszeichen steht und rechts von dem Gleichheitszeichen nicht nur eine Stringkonstante oder eine numerische Konstante stehen, sondern auch Operatoren benutzt werden.

Für eine berechnete SDTM-Variable wird der reguläre Ausdruck

```
\s*SDTMVAR\s*=\s*(.*?[\+|-|/].*?);
```

benutzt, wobei SDTMVAR durch den entsprechenden Variablennamen ersetzt wird. Der reguläre Ausdruck ist relativ grob und muss ggf. an eigene Belange angepasst werden. Dieses gilt für alle in diesem Artikel angegebenen regulären Ausdrücke.

Aus der Codezeile

```
CMSEQ = 100 + input(ID_GL_IGSEQUNO, 3.);
```

wird der entsprechende Eintrag für CMSEQ in der Kategorie 4\_CALCULATION eingetragen (siehe sechste Tabellenzeile in **Abbildung 2**).

Für berechnete SDTM-Variablen wird angenommen, dass jeweils eine Anweisung in einer Codezeile steht.

## 2.2.4 Direkte Übernahme einer eCRF-Variablen als SDTM-Variable

Eine weitere Variante der Herleitung einer SDTM-Variablen ist die direkte Übernahme einer eCRF-Variablen als SDTM-Variable, wenn die eCRF-Variable bereits den notwendigen Namen besitzt.

Diese Variante ist nicht direkt aus dem Programm heraus zu ermitteln, da die eCRF-Variable nicht explizit benannt werden muss.

Wenn man sie jedoch explizit mit KEEP anspricht, gibt es dafür auch eine Referenz im Programm, die ausgelesen werden kann.

Der reguläre Ausdruck

```
;\s+set\s+[\&a-z0-9\.\_\s]+\(\s*keep\s*=\s*([a-z0-9_\s]+)
```

bietet eine Möglichkeit, sich dem Suchergebnis anzunähern.

Er findet die Struktur

```
;\nset dataset (keep = variable1 variable2 variable3
```

Durch entsprechende Wahl der Optionen bei der Auswertung des regulären Ausdrucks kann sich die Struktur auch über mehrere Zeilen erstrecken.

Aus den Codezeilen

```
set &ds1._template IO.ig_common (\n  keep = ID_COMMON_STUDYID ID_COMMON_SITEID USUBJID\n  rename = (\n    ID_COMMON_STUDYID = STUDYID\n  )\n);
```

wird der entsprechende Eintrag für USUBJID in der Kategorie 5\_KEEP eingetragen (siehe dritte Tabellenzeile in Abbildung 3).

Domain	SDTM-Variable	Category	Specification	Core
DM	STUDYID	2_STRKONST	INGE-B	Req
DM	DOMAIN	2_STRKONST	DM	Req
DM	USUBJID	5_KEEP	USUBJID	Req

**Abbildung 3:** Auszug aus der Kategorisierung von Ableitungen der Domäne DM

## 2.2.5 Umbenennen als SDTM-Variable

Noch eine Variante der Herleitung einer SDTM-Variablen ist die Umbenennung einer eCRF-Variablen in eine SDTM-Variable.

Hierzu werden die Programmteile gesucht, die einen rename-Block enthalten, bei denen die SDTM-Variable auf der rechten Seite eines Gleichheitszeichens steht.

**Der reguläre Ausdruck**

```
rename[\s|\r\n]=\s*\ (.*(\s([a-z]+)\s*=\s*SDTMVAR\s*)\.*?\);
```

bietet eine Möglichkeit, die folgende Struktur zu finden.

```
rename = ( var1 = SDTMVAR1 var2 = SDTMVAR2)
```

Durch entsprechende Wahl der Optionen bei der Auswertung des regulären Ausdrucks kann sich die Struktur auch über mehrere Zeilen erstrecken.

Die Paare var = SDTMVAR werden dann auf die gesuchte SDTM-Variable untersucht.

**Aus den Codezeilen**

```
set io.centers (
  rename = (
    CENTERNAME = INVNAM
    CENTERLOCATIONID = INVID
    ID = SITEID
  )
);
```

werden die entsprechenden Einträge für INVID und INVNAM in der Kategorie 6\_RENAME eingetragen (siehe Abbildung 4).

Domain	SDTM-Variable	Category	Specification	Core
DM	...	...	...	...
DM	INVID	6_RENAME	CENTERLOCATIONID INVID	Perm
DM	INVNAM	6_RENAME	CENTERNAME INVNAM	Perm
DM	...	...	...	...

**Abbildung 4:** Auszug aus der Kategorisierung von Ableitungen der Domäne DM (Fortsetzung)

**2.2.6 Kennzeichnung nicht referenzierter SDTM-Variablen**

Die SDTM-Variablen, die auf die vorstehenden Arten nicht referenziert worden sind, werden entsprechend in der Liste in der zugehörigen Übersichtsseite der Domäne gekennzeichnet (ein Beispiel für die fehlende Variable EXRFTDTC in der Domäne EX siehe **Abbildung 5**, wie sie dann in der Übersicht der Domäne EX dargestellt wird).

Für eine bessere Übersicht werden sie dann noch einmal auf einer Extraseite zusammengefasst (siehe Abbildung 6).

Diese fehlenden Variablen müssen dann ggf. noch ergänzt werden.

**SDTM-Variables not used**

Domain	SDTM-Variable	Category	Specification	Core
EX	EXRFTDTC	8_NOTMENTIONED		Perm

**Abbildung 5:** Kennzeichnung einer fehlenden Variablen in der Domäne EX



## List of missing SDTM variables

Domain	SDTM-Variable	Core
AE	AEACNOTH	Perm
AE	AECAT	Perm
AE	AEDUR	Perm
AE	AEENRF	Perm
AE	AEENRTPT	Perm
AE	AEENTPT	Perm
AE	AEGRPID	Perm
AE	AELOC	Perm
AE	AEMODIFY	Perm
AE	AEPATT	Perm
AE	AEPRESP	Perm
AE	AEREFID	Perm
AE	AESCAN	Perm
AE	AESCAT	Perm
AE	AESEV	Perm
AE	AESOD	Perm
AE	AESPID	Perm
EX	EXRFTDTC	Perm

Abbildung 6: Separate Seite mit Übersicht aller fehlenden SDTM-Variablen

### 2.3 Nutzung der eCRF-Variablen (itemoids)

Die eCRF-Variablen (itemoids), die in den SAS-Programmen für die Erstellung der SDTM-Datensätze referenziert werden, werden in der entsprechenden Übersicht für die Domäne gelistet (siehe Abbildung 7).

### Usage of itemoids according to current scripts

Domain	Itemoid	Usage
AE	ID_AE_AEACN	if ID_AE_AEACN = '5' then AEACN = 'DOSE NOT CHANGED';
AE	ID_AE_AEACN	else if ID_AE_AEACN = '1' then AEACN = 'DRUG WITHDRAWN';
AE	ID_AE_AEACN	else if ID_AE_AEACN = '3' then AEACN = 'DOSE REDUCED';
AE	ID_AE_AEACN	else if ID_AE_AEACN = '4' then AEACN = 'DOSE INCREASED';
AE	ID_AE_AEACN	else if ID_AE_AEACN = '2' then AEACN = 'DRUG INTERRUPTED';
AE	ID_AE_AEACN	else if ID_AE_AEACN = '6' then AEACN = 'UNKNOWN';
AE	ID_AE_AECONCMEDCAUS	AERELNST = ID_AE_AECONCMEDCAUS;
AE	ID_AE_AECONCMEDCAUS	if ID_AE_AECONCMEDCAUS ne ";
AE	ID_AE_AECONCMEDCAUS	if ID_AE_AECONCMEDCAUSYN = 'y' then AEOTHERMED = 'Y';
AE	ID_AE_AECONCMEDCAUS	else if ID_AE_AECONCMEDCAUSYN = 'n' then AEOTHERMED = 'N';

Abbildung 7: Gebrauch der itemoids in der Domäne AE

Für eine umfassende Darstellung werden dann noch einmal alle itemoids und die nicht genutzten itemoids auf einer Extraseite zusammengefasst (siehe Abbildung 7 und Abbildung 8).

### List of all itemoids

Itemoid
ID_AE_AEACN
ID_AE_AEAPPLINGR
ID_AE_AEARGUSCASEID
ID_AE_AECONCMEDCAUS
ID_AE_AECONCMEDCAUSYN
ID_AE_AEENDAT
ID_AE_AEHOSP
ID_AE_AEIDENTIFIER
ID_AE_AELSTSTUD
ID_AE_AEOUT

Abbildung 8: List of all itemoids (Auswahl der ersten zehn Elemente)

## List of unused itemoids

Itemoid
ID_AE_AEAPPLINGR
ID_AE_AEARGUSCASEID
ID_AE_AEIDENTIFIER
ID_AE_AERELTESTLBORNRI
ID_AE_AERELTESTLBORNRO
ID_AE_RELDIAGDAT
ID_AE_RELDIAGINV
ID_AE_RELDIAGRES
ID_AE_RELDIAGUNIT
ID_CM_AETHERTRT

**Abbildung 9:** List of unused itemoids (Auswahl der ersten zehn Elemente)

### 3 Zweck von Doxygen

Doxygen wird genutzt, um die erstellten Listen zusammenzufassen und ansprechend zu einem Paket zu verpacken. Das ist dann die Grundlage für die verfügbaren Listen für das weitere Vorgehen sowie am Schluss für die finale Dokumentation.

### 4 Beispielsitzung

In den folgenden Absätzen wird eine kurze Beispielsitzung skizziert, um das Arbeiten mit den Programmen zu verdeutlichen.

#### 4.1 Pfade setzen

Wir benötigen den Pfad für den lesenden Zugriff auf die SDTM Programme, den Pfad für den schreibenden Zugriff auf die SQLite3-Datenbank für die SDTM Definitionen und die temporären Tabellen und den Pfad für den schreibenden Zugriff für die Ausgabedateien, die mit dem Python-Programm generiert und dann per Doxygen in die Ausgabe eingebunden werden.

#### 4.2 Hauptprogramm ausführen

Das Programm main.py ist das Hauptprogramm und ruft nacheinander alle notwendigen Programme auf. Nach der Ausführung dieses Programmes gibt es für jede SDTM Domäne eine Datei XX.sas, wobei XX für die entsprechende Domäne steht, für die ein entsprechendes SAS-Programm vorliegt.

Die Datei XX.sas enthält den entsprechenden HTML-Code, der dann von Doxygen dargestellt wird. Der Verweis auf die jeweilige Datei findet sich auf der Registerkarte „Related Pages“ (siehe Abbildung 10).

Auf dieser Registerkarte finden sich ebenfalls die beiden Übersichtsseiten, die die Vollständigkeit der Nutzung der SDTM Variablen zusammenfassen und die Nutzung der eCRF Variablen auflisten. Zusätzlich gibt es eine Seite, die alle Makros listet und die aufführt, die nicht genutzt werden.

INGE-B

Main Page Related Pages Modules Files Search

### Related Pages

Here is a list of all related documentation pages:

<a href="#">CDISC variables of domain AE</a>	
<a href="#">CDISC variables of domain CM</a>	
<a href="#">CDISC variables of domain DM</a>	
<a href="#">CDISC variables of domain DS</a>	
<a href="#">CDISC variables of domain EX</a>	
<a href="#">CDISC variables of domain LB</a>	
<a href="#">CDISC variables of domain MI</a>	
<a href="#">List of itemoids (unused, all)</a>	
<a href="#">Summary of completeness of SDTM variables</a>	
<a href="#">CDISC variables of domain PR</a>	
<a href="#">CDISC variables of domain QS</a>	
<a href="#">CDISC variables of domain RS</a>	
<a href="#">CDISC variables of domain TU</a>	
<a href="#">List of macros (unused, all)</a>	
<a href="#">CDISC variables of domain VS</a>	
<a href="#">Todo List</a>	

Generated on Tue Jan 23 2018 09:33:30 for INGE-B by **doxygen** 1.8.3

**Abbildung 10:** Auswahl der Zusatzseiten (Automatische Sortierung der Seiten durch Doxygen ist noch nicht optimal)

## 4.2 Ergebnisse einarbeiten

Jetzt werden die bisher vorliegenden Dokumentationsseiten für die Domänen um die Ergebnisse der Programmausführung ergänzt.

### SDTM\_EX.sas File Reference

Study Data Table generating programs

---

Creates SDTM.EX - Exposure. [More...](#)

[Go to the source code of this file.](#)

---

#### Detailed Description

Creates SDTM.EX - Exposure.

This script creates SDTM.EX

Definition of source according to CDISC:

The Exposure domain model records the details of a subject's exposure to protocol-specified study treatment. Study treatment may be any intervention that is prospectively defined as a test material within a study, and is typically but not always supplied to the subject.

#### Derivation of variables

- STUDYID: Set to "INGE\_B"
- DOMAIN: Set to "EX"
- USUBJID: NOT YET DONE
- EXSEQ: NOT YET DONE
- EXGRPID: NOT YET DONE
- EXREFID: NOT YET DONE
- EXSPID: NOT YET DONE
- EXLNKID: NOT YET DONE
- EXLNKGRP: NOT YET DONE
- EXTRT: NOT YET DONE
- EXCAT: NOT YET DONE

#### Specification of variables according to current scripts

Domain	SDTM variable	Category	Specification	Core
EX	STUDYID	2_STRKONST	INGE_B	Req
EX	DOMAIN	2_STRKONST	EX	Req
EX	USUBJID	7_ELSEWHERE	Possible hints see below	Req
EX	EXSEQ	3_VARIABLES	EXSEQRANKED	Req
EX	EXGRPID	7_ELSEWHERE	Possible hints see below	Perm
EX	EXREFID	7_ELSEWHERE	Possible hints see below	Perm
EX	EXSPID	7_ELSEWHERE	Possible hints see below	Perm
EX	EXLNKID	7_ELSEWHERE	Possible hints see below	Perm
EX	EXLNKGRP	7_ELSEWHERE	Possible hints see below	Perm
EX	EXTRT	2_STRKONST	PALBOCICLIB	Req
EX	EXTRT	2_STRKONST	LETROZOLE	Req
EX	EXTRT	2_STRKONST	FULVESTRANT	Req
EX	EXTRT	2_STRKONST	ANASTROZOLE	Req
EX	EXTRT	2_STRKONST	EXEMESTANE	Req
EX	EXCAT	7_ELSEWHERE	Possible hints see below	Perm

Abbildung 11: Übertrag von Hilfstabelle auf Dokumentationsseite

Dazu werden die Herleitungen in das eigentliche SAS Programm übertragen, die Farbkodierungen werden zurückgesetzt und ggf. fehlende SDTM Variablen werden im Programm ergänzt. Bei der Bearbeitung kann es durchaus mehrere Reviewzyklen geben.

## **5 Schlussfolgerung**

Die beschriebene Sammlung von Programmen ermöglicht es, die Nutzung von SDTM Variablen effektiver und konsistenter zu dokumentieren. Die Python-Programme können über die Emailadresse [requestscripts@joerg-sahlmann.de](mailto:requestscripts@joerg-sahlmann.de) und dem Stichwort „KSFE2018“ in der Betreffzeile bis zur kommenden KSFE 2019 angefordert werden.

Aufbauend auf diesen Programmen können die Seiten mit den Zusatzdokumentationen in Doxygen individuell erweitert werden. Aktuell gibt es eine Seite, die alle benutzten Makros auflistet und diejenigen gesondert aufführt, die nicht benutzt werden.

## **Literatur**

- [1] Doxygen, Generate documentation from source code, <http://www.doxygen.org> [28.12.2017].
- [2] A. Bachert, SAS-Programme mit Doxygen automatisch dokumentieren, <https://communities.sas.com/t5/CoDe-SAS-German/SAS-Programme-mit-Doxygen-automatisch-dokumentieren/gpm-p/221854> [28.12.2017].
- [3] Pinnacle21, <https://www.pinnacle21.com/> [28.12.2017].
- [4] CDISC. Study Data Tabulation Model Version 1.4, [https://www.cdisc.org/system/files/members/standard/foundational/sdtm/study\\_data\\_tabulation\\_model\\_v1\\_4.pdf](https://www.cdisc.org/system/files/members/standard/foundational/sdtm/study_data_tabulation_model_v1_4.pdf) [28.12.2017].