



SAS
ユーザー総会
2017

SASを利用したPinnacle21 Communityの セミオートメーション化による業務効率化

○片山 雅仁^{1†}、勘場 大^{1†}、神谷 亜香里¹、
江川 広明¹、角谷 伸一¹

(¹塩野義製薬株式会社 解析センター)

Semi-automation of Pinnacle21 Community using SAS

Masahito Katayama, Masaru Kamba, Akari Kamitani,
Hiroaki Egawa, Shinichi Kakutani
Biostatistics Dept., SHIONOGI CO., LTD.

† : Equal contribution

要旨:

PMDAへの申請電子データ提出が開始され, Pinnacle21 Communityを利用してバリデーションする機会が多くなっている.
SASを利用した効率よく間違いのない実行方法を紹介する.

キーワード: PMDA, FDA, SDTM, ADaM, Define.xml, Pinnacle21 Community, Validation, Command Line Interface

Disclaimer

- 本発表で紹介している技術は発表者のPC, ソフトウェアの環境で動作を確認しているものであり, 使用する際はご自身のPCの環境を確認の上, ご自身の責任の下で実施ください



背景

- 2016年10月よりPMDAの申請電子データ(SDTM, ADaM等)提出が開始されたことにより, 申請電子データをバリデーションする機会が増えた
- バリデーションを繰り返すことにより, 手間になることがある

プログラムによる自動化により業務効率化を考えた



バリデーションツールについて

PMDA Web Siteより

申請電子データ提出に関する技術情報(FAQ、データカタログ等)

申請電子データ提出に関する技術情報を掲載しております。

PMDAではPinnacle21 Enterprise 3.0.5が標準

- バージョンアップ等により当該ソフトウェアを更改する際には、予め本HP上でご案内する予定です。
- [バージョンアップ等により当該ソフトウェアを更改する際には、予め本HP上でご案内する予定です。](#)
- [申請電子データに関するFAQ](#)
- [TSVファイルの作成・編集方法](#)
- PMDAは、平成28年10月から承認申請される品目に添付される申請電子データのバリデーションには、『Pinnacle 21 Enterprise 3.0.5』を利用します。

Compatibility with Enterprise

Pinnacle21 Web Siteより

Pinnacle 21 has executed a series of tests on Pinnacle 21 Community v2.1.3 and Pinnacle 21 Enterprise v.3.0.5 in the Pinnacle 21 test environment, and the testing showed compatibility

Pinnacle21 Community 2.1.3と互換性がある

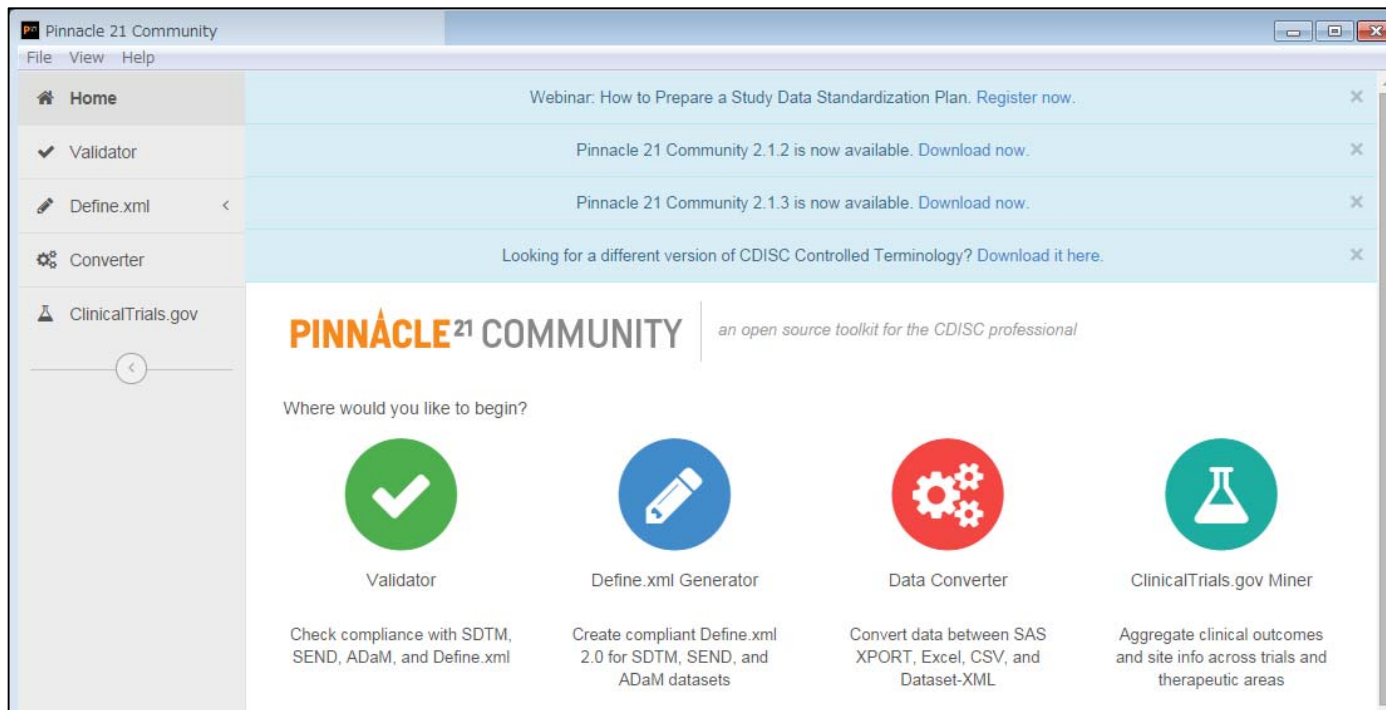
Enterprise vs. Community

	Pros.	Cons.
Enterprise	<ul style="list-style-type: none"> ➤ Rejectのリスクを避ける ➤ ユーザーサポートがある ➤ 毎回の設定が不要 	<ul style="list-style-type: none"> ➤ 有償版である ➤ Zipファイルの作成
Community	<ul style="list-style-type: none"> ➤ 無償版である ➤ 手軽に利用できる 	<ul style="list-style-type: none"> ➤ 毎回の設定が必要 ➤ 正確を期すには不安

現時点でCommunity版を利用している実務担当者も大勢いると予想される

Community版の利用を想定してプログラムを開発した

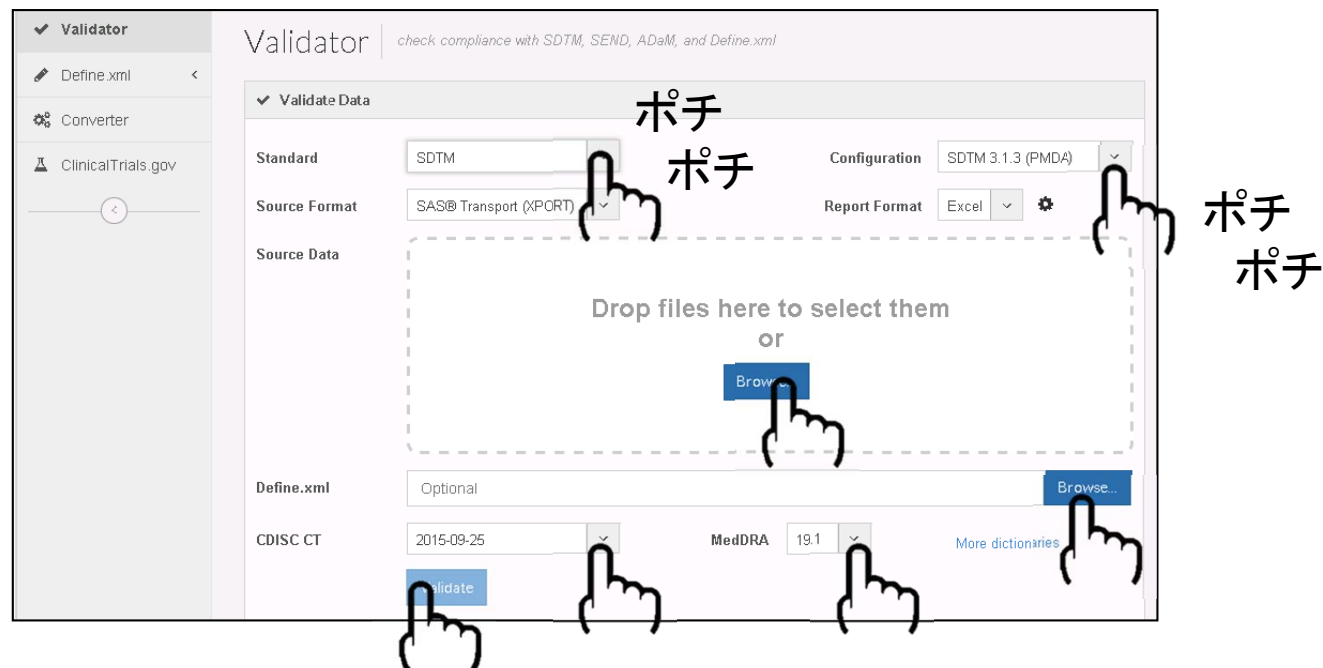
Pinnacle21 Community



Pinnacle21 CommunityはGUI# で誰でも簡単に操作することができるツールであるため、非常に便利である

GUIのデメリット

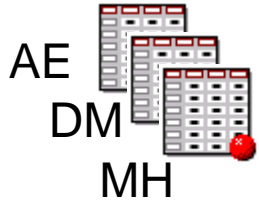
- 手作業で必要項目を指定する必要
 - とくに複数回のバリデーションをすることになると、その操作に時間を要する



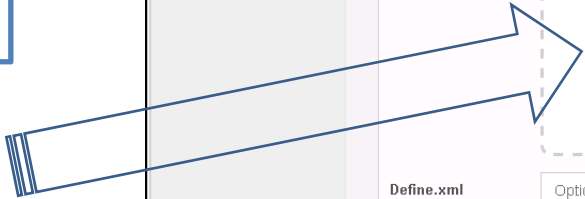
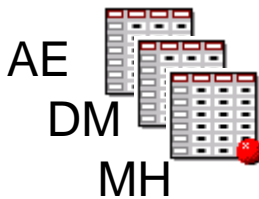
Pinnacle 21あるある その1

CDISC標準は、どの試験でもファイル名がほぼ同じなので、バリデーションしたい試験以外を選択してバリデーションしてしまう

バリデーションしたい試験



間違って選択した試験



Pinnacle 21あるある その2

前回の設定が保存されるため、複数試験の場合階層の移動が面倒

技術的ガイド/Technical Conformance guideで決められているフォルダ構造で各種成果物が格納されている階層は深い

前回の設定でウィンドウが展開

フォルダ間の移動

薬剤Aのフォルダ 薬剤Bのフォルダ

The image shows a side-by-side comparison of folder structures for two studies, labeled '薬剤Aのフォルダ' (left) and '薬剤Bのフォルダ' (right). Both structures are identical, starting with 'm5' and 'datasets', followed by a study-specific folder '[study id / iss / ise]'. Underneath, there are sub-folders for 'analysis', 'adam' (with sub-folders 'datasets' and 'programs'), 'adam_j', 'legacy' (with sub-folder 'datasets'), 'procedures', 'misc', 'tabulations', 'legacy', 'sdtrm', and 'sdtrm_j'. A double-headed arrow between the two structures is labeled 'フォルダ間の移動' (Folder movement). To the right, a screenshot of the Pinnacle 21 interface shows a 'Browse...' button with a hand cursor pointing to it. A callout bubble points to this button, stating '前回の設定でウィンドウが展開' (The window is expanded with the previous settings). Another callout bubble points to the folder structure, stating '技術的ガイド/Technical Conformance guideで決められているフォルダ構造で各種成果物が格納されている階層は深い' (The folder structure is deep because various results are stored according to the folder structure defined in the Technical Conformance Guide).

薬剤Aのフォルダ 薬剤Bのフォルダ

Pinnacle 21あるある その3

前回の設定が保存されるため、間違ったバージョンを選ぶリスクがある

ほげえ！



Validator | *check compliance with SDTM, SEND, ADaM, and Define.xml*

✓ Validate Data

Standard: SDTM Configuration: SDTM 3.1.3 (PMDA)

Source Format: SAS® Transport (XPORT) Report Format: Excel

Source Data: Drop files here to select them or Browse...

Define.xml: Optional Browse...

CDISC CT: 2015-09-25 MedDRA 19.1 More dictionaries

Validate

Pinnacle 21あるある その4

- GUIでを使用した場合、バリデーション結果がPinnacle21 Communityの所定のサブフォルダに保存される
- また、そのフォルダにアクセス権が適切に設定されていない場合、セキュリティ上好ましくない場合がある
- Pinnacle21の結果が出力されたフォルダからアクセス制御されたフォルダへファイルの移動をすることが好ましいが、手作業になるため、操作ミスを引き起こす可能性がある

出力されるサブフォルダ

名前
pinnacle21-report-2016-12-14T21-31.xlsx
pinnacle21-report-2016-12-15T16-59.xlsx
pinnacle21-report-2016-12-22T05-24.xlsx
pinnacle21-report-2016-12-22T05-26.xlsx
pinnacle21-report-2016-12-27T04-23.xlsx
pinnacle21-report-2017-01-03T23-00.xlsx
pinnacle21-report-2017-01-13T16-10.xlsx
pinnacle21-report-2017-02-07T19-19.xlsx
pinnacle21-report-2017-03-19T03-22.xlsx
pinnacle21-report-2017-03-20T16-22.xlsx
pinnacle21-report-2017-04-14T18-24.xlsx
pinnacle21-report-2017-04-14T19-00.xlsx

- 意図的に削除しないと過去のファイルがどんどん溜まる
- そのため、どのファイルを作成したかがわかりにくい

こまったにゃー



Pinnacle 21 Communityのインストール場所

	Pros.	Cons.
代表的なフォルダに1つインストール	<ul style="list-style-type: none"> ➤ IQ/OQ/PQがとりやすい 	<ul style="list-style-type: none"> ➤ Pinnacle21のフォルダと、各臨床試験のフォルダの距離が離れる ➤ アクセス権の問題
各品目又は各プロトコルフォルダにインストール	<ul style="list-style-type: none"> ➤ ファイルの移動距離が短い(又は移動不要) 	<ul style="list-style-type: none"> ➤ インストールした分IQ/OQ/PQが必要 ➤ サーバーの容量を消費する



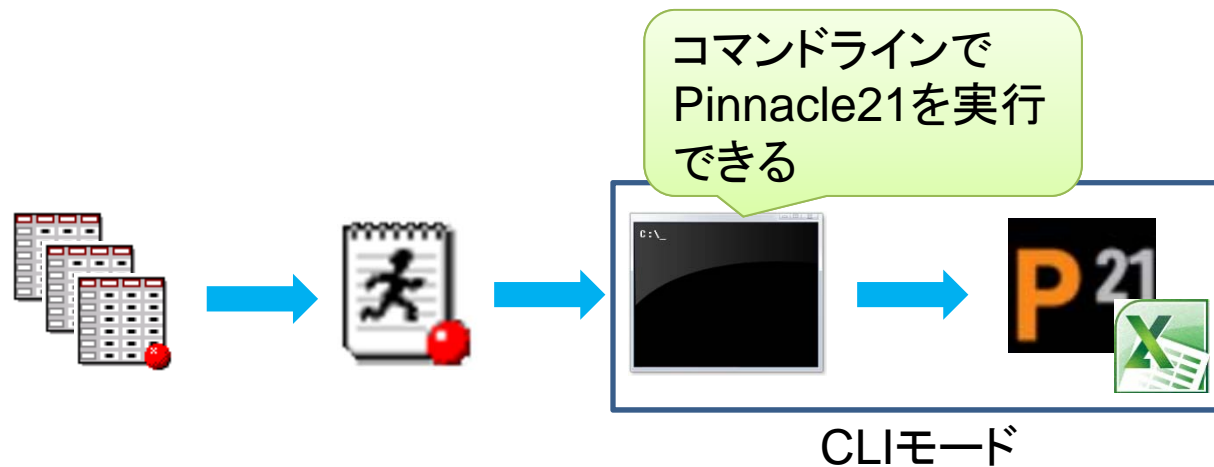
代表的なフォルダに1つ



各品目又はプロトコル

CLI#モード

- Pinnacle21の動作に必要な所定のパラメータをコマンドラインに渡すことで、Define.xml作成やバリデーションなどの機能を利用できる
- SASでコマンドラインを動かすことができるので、プログラムで事前にパラメータを設定しておけば、該当試験のバリデーション時に繰り返し利用可能となる



GUIがCLIモードになると？

- データセットのバリデーション
- Define.xml Generator
- Define.xmlのバリデーション
- Data Converter

データセットのバリデーション

データセットのバリデーションに必要なパラメータ(一部抜粋)

Parameter Type	Parameter	Valid Values	Description
	-type	SDTM, ADaM, SEND, Define, Custom	Data Standard/Structure to validate
Source Data	-source	<path>	ソースデータのpath
	-source:type	SAS Delimited	ファイルのフォーマット
Configuration	-config	<path>	Configのpath
	-config:define	<path>	Define.xmlのpath
	-config:cdisc	<version>	CTのバージョン
	-config:meddra	<version>	MedDRAのバージョン
Report	-report	<path>	結果の出力先
	-report:type	Excel, CSV, XML	結果のフォーマット
	-report:overwrite	yes, no	結果の上書き

データセットのバリデーション

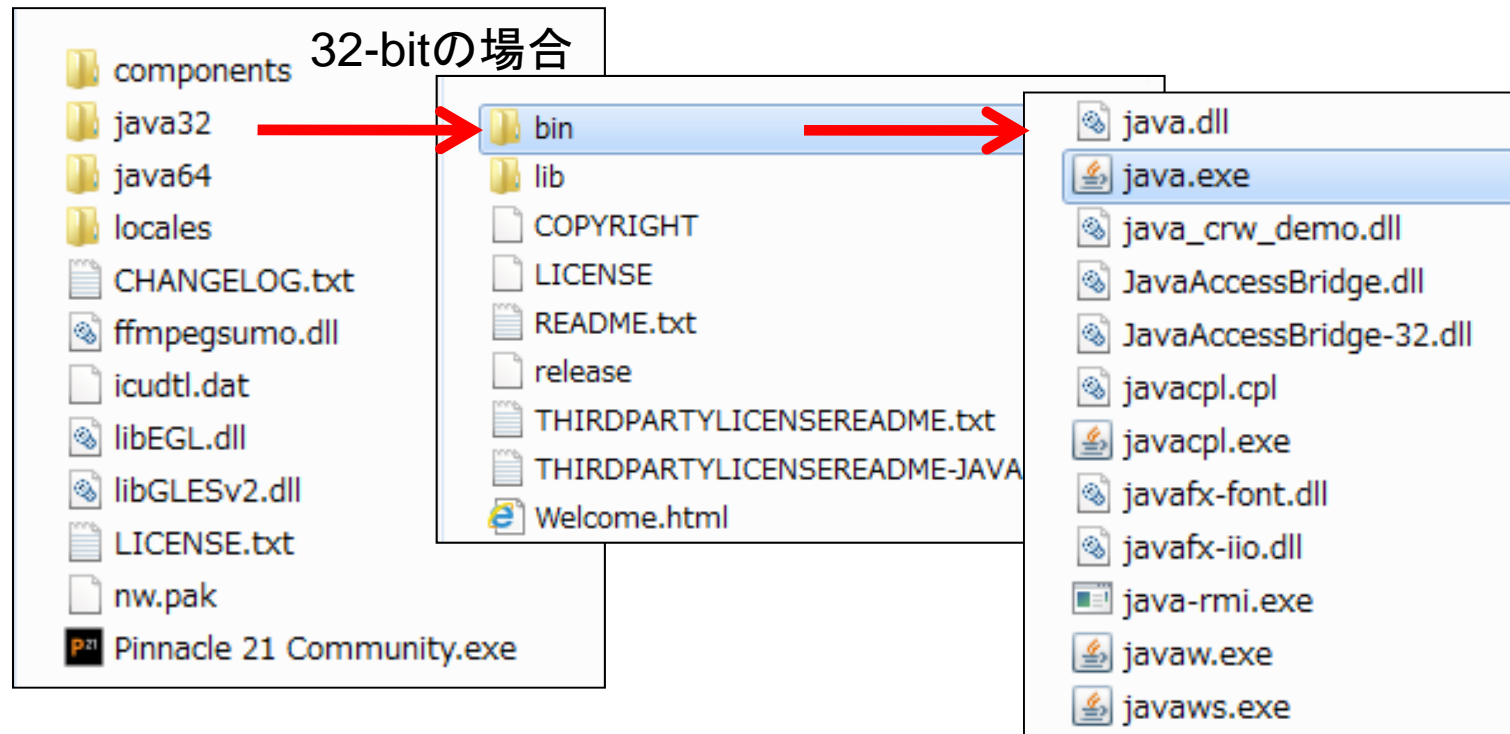
チェンジディレクトリ (cd)
を指定しないとCLIは利
用できない



```
x "cd &path_pinnacle.¥pinnacle21-community_2.1.3¥java32¥bin";  
x java -Xms256m -Xmx1024m  
-jar "&path_pinnacle.¥pinnacle21-community_2.1.3¥components¥lib¥validator-cli-2.1.3.jar"  
-type=sdtm  
-source:type=sas  
-source="&path_study.¥tabulations¥sdtm¥*.xpt"  
-config="&path_pinnacle.¥pinnacle21-community_2.1.3¥components¥config¥SDTM 3.1.3  
(PMDA).xml"  
-config:cdisc=2015-09-25  
-config:meddra=19.1  
-config:define="&path_study.¥tabulations¥sdtm¥define.xml"  
-report="&path_report.¥pinnacle21-report-&date..xlsx"  
-report:type=excel  
-report:overwrite=yes ;
```

結果の出力先を指定できる！
ファイルの移動が不要！！
安全！！

Pinnacle21でのJava



Pinnacle21 CommunityにはJavaが実装されているため、OpenCDISC Validatorの時代のように、PCにインストールされているJavaのバージョンを気にする必要はない

Pinnacle21の各機能の保存先

Pinnacle21 Community 2.1.3の場合

components
java32
java64
locales
CHANGELOG.txt
ffmpegsumo.dll
icudtl.dat
libEGL.dll
libGLESv2.dll
LICENSE.txt
nw.pak
Pinnacle 21 Community.exe

config
lib
logs
reports

properties
clinicaltrials-miner-2.0.3.jar
commons-codec-1.9.jar
commons-codec-1.10.jar
commons-io-2.4.jar
commons-lang-2.6.jar
commons-lang3-3.1.jar
commons-lang3-3.4.jar
data-converter-1.0.6.jar
data-parsing-1.1.3.jar
define-generator-1.1.4.jar
desktop-updater-1.1.2.jar
fastutil-6.6.3.jar
guava-19.0.jar
ipc-bridge-1.0.3.jar
joda-time-2.7.jar
jsap-2.1.jar
jsr305-3.0.1.jar
log4j-1.2.17.jar
log4j-api-2.1.jar
log4j-core-2.1.jar
minimal-json-0.9.4.jar
opencsv-2.3.jar
poi-3.13.jar
poi-ooxml-3.13.jar
poi-ooxml-schemas-3.13.jar
protobuf-java-2.6.1.jar
rxjava-1.0.7.jar
saxon-8.7.jar
saxon-dom-8.7.jar
slf4j-api-1.7.13.jar
slf4j-log4j12-1.7.13.jar
stax-api-1.0.1.jar
validator-api-2.1.3.jar
validator-cli-2.1.3.jar
validator-core-2.1.3.jar

これらのファイルをコマンド
ラインで呼び出す

データセットのバリデーションの注意点

- SDTMのFAドメイン
 - コマンドプロンプトに以下のメッセージが現れ、Hangしてしまう...

```
The ItemGroupDef named FA conflicts with a configuration prototype (oid: IG.FA).  
This shouldn't happen!
```

- Pinnacle21のForumでも取り上げられた問題だが未解決(参考資料3)

Define.xmlのバリデーション

データセットのバリデーションよりいくつか少ないパラメータですが...

```
x "cd &path_pinnacle.¥pinnacle21-community_2.1.3¥java32¥bin";  
x java -Xms256m -Xmx1024m  
-jar "&path_pinnacle.¥pinnacle21-community_2.1.3¥components¥lib¥validator-cli-2.1.3.jar"  
-type=Define  
-source="&path_study.¥tabulations¥sdm¥define.xml"  
-config="&path_pinnacle.¥pinnacle21-community_2.1.3¥components¥config¥Define.xml  
(PMDA).xml"  
-config:define="&path_study.¥tabulations¥sdm¥define.xml"  
-config:cdisc=2015-09-25  
-report="&path_report.¥sdm¥pinnacle21-report_&date..xlsx"  
-report:type=excel  
-report:overwrite=yes ;
```

実はCLIではDefine.xmlのバリデーションはできない...
Pinnacle21のForumにできないことが書いてある...
(参考資料4, 5)

Define.xml Generator

Parameter	Valid Values	Description
<Template>	<path>	Template Excelのpath
<Destination>	<path>	xml出力先のpath

```
options noxsync noxwait;
x "mkdir &path_study.¥tabulations¥sdtm¥define.xml";
x "cd &path_pinnacle.¥pinnacle21-community_2.1.3¥java32¥bin";
x echo Y | java -Xms256m -Xmx1024m -jar "&path_pinnacle.¥pinnacle21-
community_2.1.3¥components¥lib¥define-generator-1.1.4.jar"
  "&path_spec.¥sdtm¥sdtm_metadata.xlsx"
  "&path_study.¥tabulations¥sdtm"
;
```

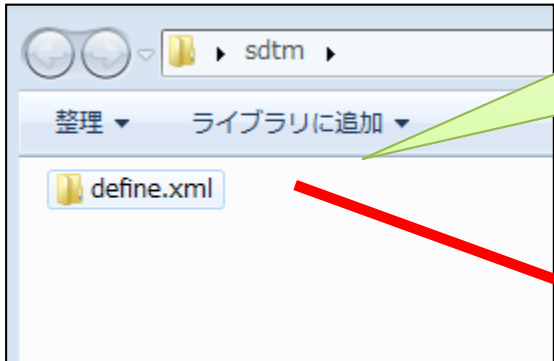
事前に[define.xml]という
フォルダを作成しないと
define.xmlを作成できない



echo Y | を指定することで(は
いいえ)の質問で、「はい」を
自動的に入力できる



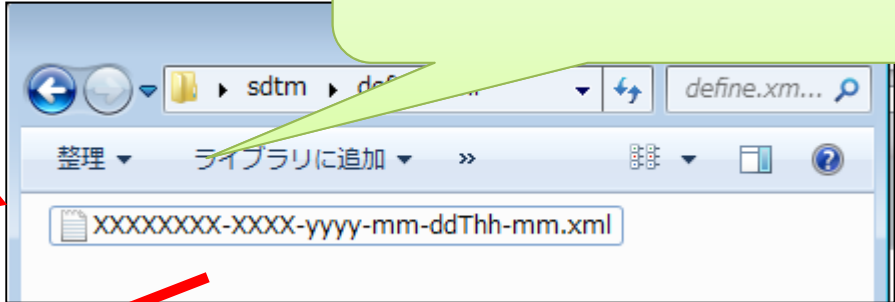
Define.xmlの作成先とファイル操作(SAS)



結果を出力したいフォルダ

ここにdefine.xmlというサブフォルダを作る

試験名, 日付, 時間入りのdefine.xmlができる



define.xmlフォルダ (一時フォルダ)



最終的な姿

- ①最終フォルダへxmlファイルの移動
- ②一時フォルダの削除
- ③ファイル名をdefine.xmlへリネーム

Define.xml Generator

コード フルバージョン

```
options noxsync noxwait;  
x "mkdir &path_spec.¥sdtm¥define.xml";  
x "cd &path_pinnacle.¥pinnacle21-community_2.1.3¥java32¥bin";  
x echo Y | java -Xms256m -Xmx1024m -jar "&path_pinnacle.¥pinnacle21-  
community_2.1.3¥components¥lib¥define-generator-1.1.4.jar"  
  "&path_spec.¥sdtm¥sdtm_metadata.xlsx"  
  "&path_study.¥tabulations¥sdtm"  
;  
  
/* Move xml -> Remove Folder -> Rename xml */  
data _null_;  
  call system("move &path_study.¥tabulations¥sdtm¥define.xml¥*.xml  
&path_study.¥tabulations¥sdtm");  
  call system("rd &path_study.¥tabulations¥sdtm¥define.xml");  
  call system("ren &path_study.¥tabulations¥sdtm¥*.xml define.xml");  
run;
```

Xコマンドの代わりに
Call systemルーチン
でも代用可

Data Converter

Parameter	Valid Values	Description
-s	source	Data source のpath
-i	source type	Data sourceのフォーマット (Default:xpt)
-o	output	Output fileのpath
-e	output type	Output fileのフォーマット (Default:xlsx)
-d	define	define.xmlのpath (Dataset-XMLのみ)
-c	config	configのpath (Dataset-XMLのみ)

```
x "cd &path_pinnacle.¥pinnacle21-community_2.1.3¥java32¥bin";
x java -Xms256m -Xmx1024m
-jar "&path_pinnacle.¥pinnacle21-community_2.1.3¥components¥lib¥data-converter-1.0.6.jar"
-s="&path_study.¥tabulations¥sdtm"
-o="&path_study.¥excel"
;
```

SDTMをExcelにアウトプット
できるので、非常に便利！



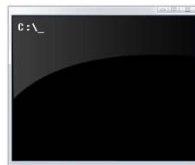
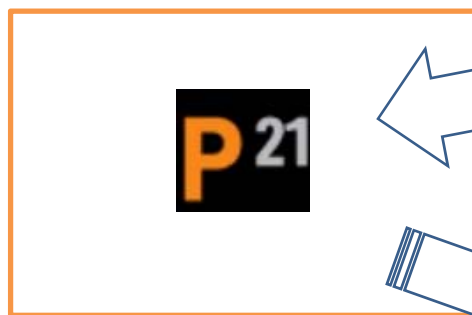
Pinnacle21 Community実行マクロ

- SASでDefine.xml作成, 申請電子データ(SDTM, ADaMデータセット)のバリデーション用にPinnacle21 CommunityをCLI実行するSASマクロを作成した

Pinnacle21 Community実行SASマクロ

品目/試験ごとに設定されている
アクセス権設定済みの場所

Pinnacle21のインストール場所



各プロセス(SDTM
Validation等)でプ
ログラムを準備し、
試験名などは自動
取得する



実行結果

SASマクロをアクセス権が適切に付与された環境下から実行することで、
Pinnacle 21のインストール場所のサブフォルダを介すことなく、
所定の環境に結果のファイルを作成できる
セキュアなバリデーションプロセスを検討した

SASマクロ作成のポイント

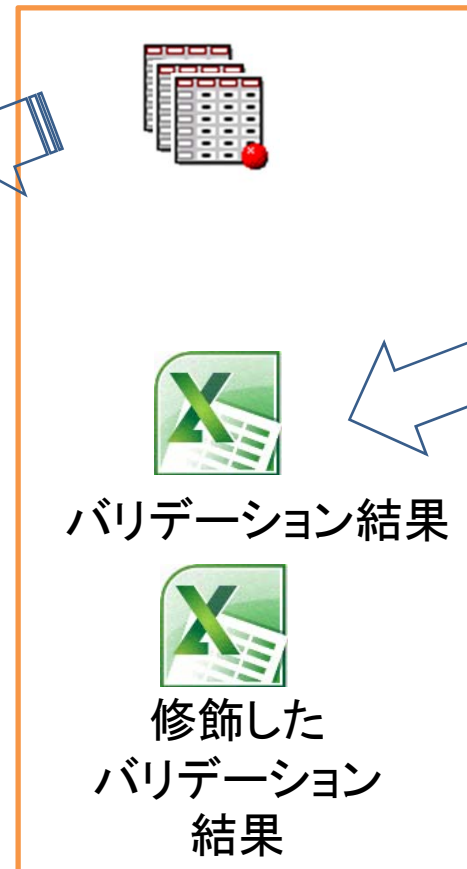
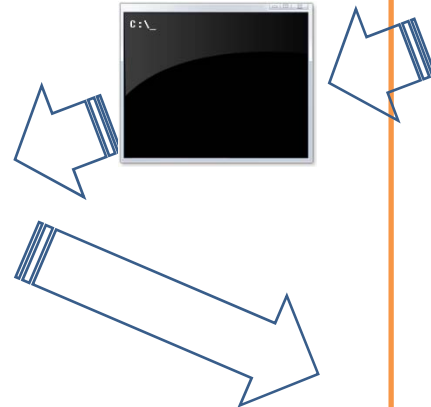
- あらかじめ決めておいた方がよい情報
 - Pinnacle21のフォルダ
 - Pinnacle21のバージョン
 - 試験データの環境のフォルダ構造
 - 利用するファイルの命名規則
 - 結果ファイルの出力場所
 - 出力するファイルのフォーマット
- マクロパラメータにすべき情報
 - Configのバージョン
 - MedDRAのバージョン
 - CT (CDISC Controlled Terminology) のバージョン
 - その他の辞書 (SNOMED, UNII, NDF-RT)

パラメータをなるべく少なくして、使いやすく！

Pinnacle21実行SASマクロの拡張

品目/試験ごとに設定されている
アクセス権設定済みの場所

Pinnacle21の
インストール場所



社内ルールファイル：
Warningが出るものが
予め分かっているルール

Pinnacle21実行SASマクロの拡張

Domain	Record	Count	Variables	Values
AE			VARIABLE	EPOCH
AE	9		AEENDTC	null
AE	20		AEENDTC	null
AE	29		AEENDTC	null

(例) Validation RuleにはEPOCHも含まれるが、そもそもSDTMIGにはEPOCHは定義されていないため、このWarningが出力されたとしても、検討する必要は無い

Pinnacle 21 ID	Publisher ID	Message	Consistency	Warning
SD1076		Model permissible variable added into sta		
SD0021		Missing End Time-Point value		
SD0021		Missing End Time-Point value		
SD0021		Missing End Time-Point value		

Excelの色フィルタを利用

大量に得られるバリデーション結果についても、事前に取り決めた社内ルールと結合することで、試験途中でのデータ確認を簡略化できる

まとめとさらなる展望

- まとめ
 - Pinnacle21 CommunityをCommand Line Interfaceから実行できる方法と結果の出力先をコントロールできるGUIにはないメリットを紹介した
 - Pinnacle21 Communityの実行をするSASマクロと効率よく、間違いのない方法を紹介した
 - バリデーション結果に社内ルールと結合することで、データ確認の作業を簡略化する方法を紹介した
- さらなる展望
 - Pinnacle21 Community版でのCLIの改良
 - Xコマンドを利用したSASと他のツールの連携

参考資料

1. Sirichenko, S., DiGiantomasso, M., and Collopy, T. (2015). Usage of OpenCDISC Community Toolset 2.0 for Clinical Programmers. PharmaSUG 2015 – Paper HT04
2. Pinnacle21 <https://www.pinnacle21.com/using-opencdisc-validator-cli>
3. Pinnacle21 Forum <https://www.pinnacle21.com/forum/hang-itemgroupdef-conflicts-configuration-prototype>
4. Pinnacle21 Forum <https://www.pinnacle21.com/forum/issues-cli-command-validating-definexml-21-vs-202>
5. Pinnacle21 Forum <https://www.pinnacle21.com/forum/cli-functionality-wrong-version-number-reported-no-definexml-support>
6. Holland C, Shostak J. Implementing CDISC Using SAS®: An End-to-End Guide. SAS Institute INC (NC); 2012.
7. PMDA 申請電子データ提出に関する技術情報(FAQ、データカタログ等) <https://www.pmda.go.jp/review-services/drug-reviews/about-reviews/p-drugs/0028.html>

ご静聴ありがとうございました