FMD

Database Design
VS
SAS Dataset Structure

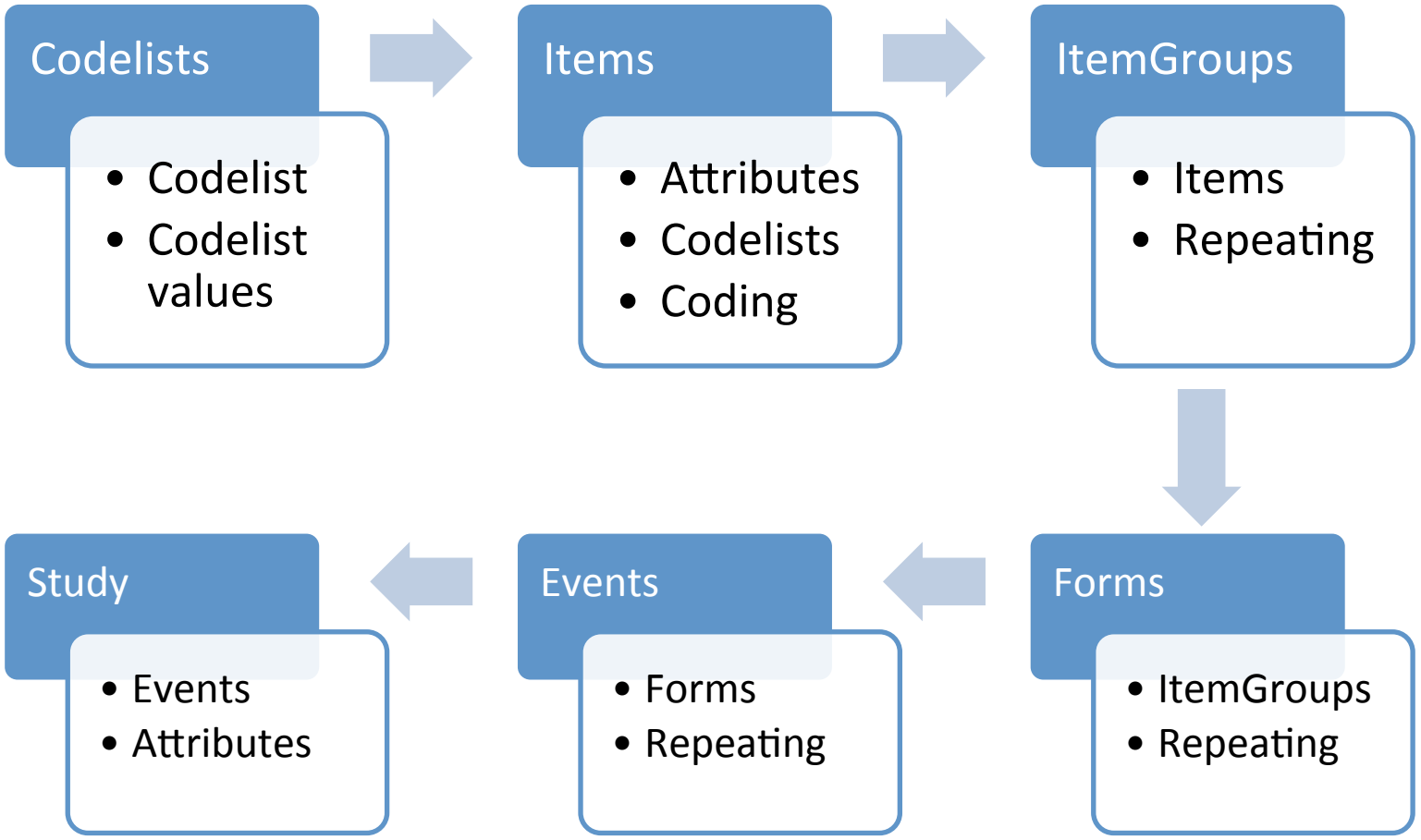
Anran Xie, FMD K&L

Nov.2016

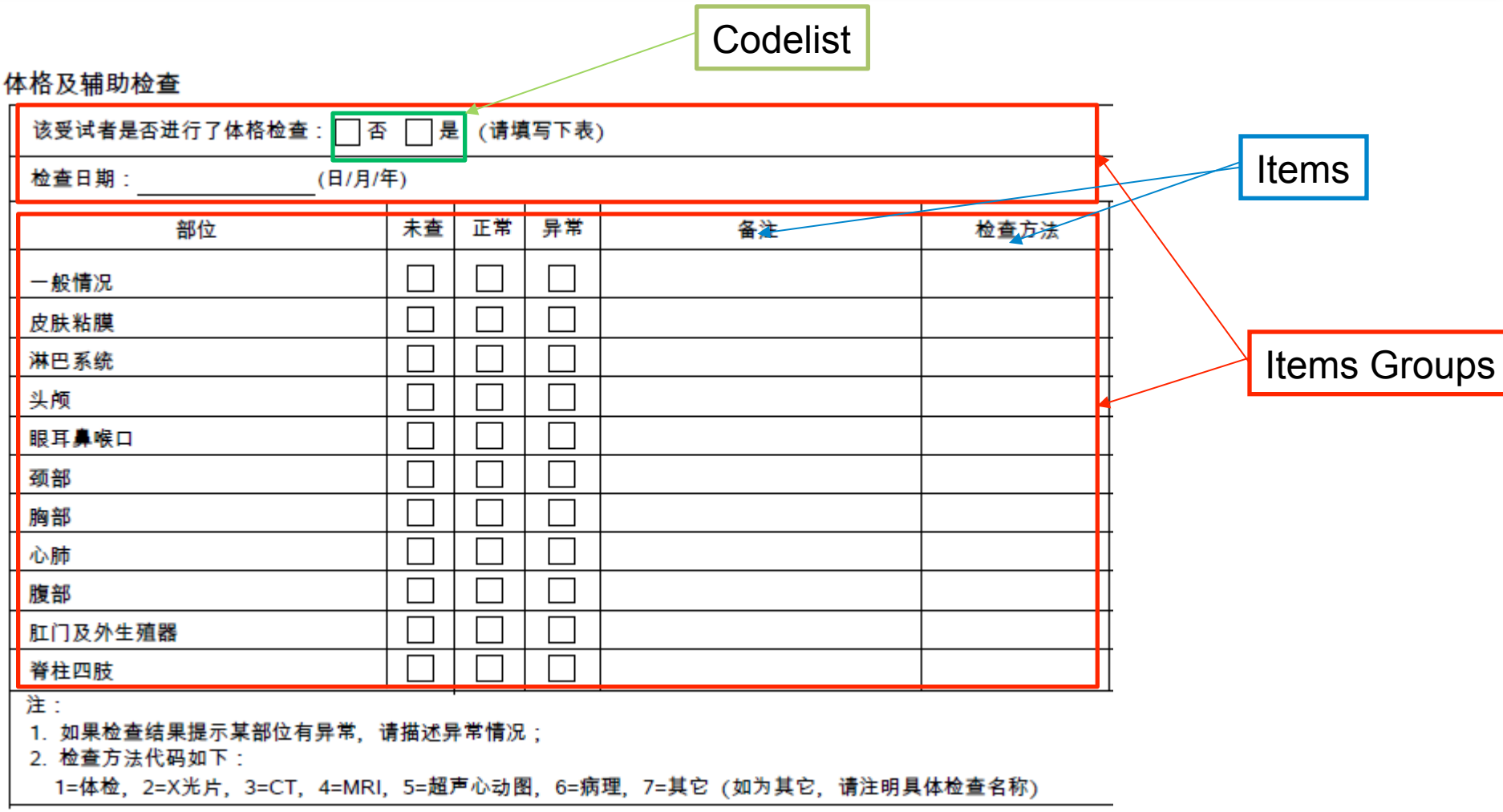
Agenda

- Database design basics
  - DB design process
  - Form structures VS. Dataset Num
- How datasets are generated
  - Item Attributes VS. SAS variable
  - Database underlying table structures
  - Dataset generation
- Dataset common & extra variables
  - Common Variables
  - Extra variables

DB design process



DB design process



Different ItemGroup structures

- Non-repeat (single-repeat)
- Repeat
- Repeat with default values

Different ItemGroup structures

- Non-repeat (single-repeat)

人口统计学资料

出生日期：_____(日/月/年),

如果不能获得请填写年龄：_____ 岁

性别	☐ 男 ☐ 女
婚否	☐ 已婚 ☐ 未婚
民族	☐ 汉族 ☐ 其他 其他：_____

- Export One dataset

Different ItemGroup structures

- Repeat

既往病史或伴随疾病

该受试者是否有既往疾病史/伴随疾病：☐否 ☐是 （请填写下表）

编号	疾病或异常名称	开始日期 (日/月/年)	结束日期 (日/月/年)	是否 持续
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否
				☐是 ☐否

Different ItemGroup structures

- Repeat with default values

体格及辅助检查

该受试者是否进行了体格检查：☐否 ☐是 （请填写下表）

检查日期：_____（日/月/年）

部位	未查	正常	异常	备注	检查方法
一般情况	☐	☐	☐		
皮肤粘膜	☐	☐	☐		
淋巴系统	☐	☐	☐		
头颅	☐	☐	☐		
眼耳鼻喉口	☐	☐	☐		
颈部	☐	☐	☐		
胸部	☐	☐	☐		
心肺	☐	☐	☐		
腹部	☐	☐	☐		
肛门及外生殖器	☐	☐	☐		
脊柱四肢	☐	☐	☐		

Form structure VS. Dataset Num

Form Structure	Dataset Num	Description
Non-repeat only	1	
Repeat only	1	
Non-repeat + Repeat	1	Merged Non-repeat with Repeat in two ways: 1). Non-repeat value display as Record 0 2). Non-repeat value displays in each record
	2	All non-repeat item groups as one dataset Each repeat item group as one dataset

Note: Form with minor differ(Subsets) may have more datasets per different EDC systems

How datasets are generated

- Item Attributes VS. SAS Variable
- Database underlying table structures
- Dataset generation

Item Attributes VS. SAS Variable

Items (In database)	Variables (SAS Dataset)
Item Name	Variable Name
Data Type Char, Integer, Float, Datetime, Boolean Length Precision	Format
Item label	Label
SAS Var Name	Variable Name
SAS Format	Format
SAS Label	SAS Label

Database underlying table structures

- Horizon (SDTM like)
- Vertical (One datapoint per line)

	PT	CPEVENT	SUBEV...	REPEATSN	BRTHDAT	AGE	AGEU	SEX
1	T1	SCREEN	0	1	19900302	25	周岁	M
2	T10	SCREEN	0	1	19590319	56	周岁	M
3	T101	SCREEN	0	1	19900101	22	周岁	M
4	T102	SCREEN	0	1	19900101	25	周岁	M
5	T103	SCREEN	0	1	19900101	22	周岁	M
6	T104	SCREEN	0	1	19900101	22	周岁	F
7	T105	SCREEN	0	1	19900101	22	周岁	M
8	T107	SCREEN	0	1	20110101	1	周岁	M
9	T108	SCREEN	0	1	20110201	1	周岁	M
10	T11	SCREEN	0	1	19900909	25	周岁	M
11	T110	SCREEN	0	1	19900101	39	周岁	M
12	T111	SCREEN	0	1			周岁	F

Database underlying table structures

- Horizon (SDTM like)
- Vertical (One datapoint per line)

	PATIENT	CLIN_PLAN_EVE_N...	NAME	SUBSET	REPEAT_SN	QUESTION_NAME	VALUE_TEXT	EXCEPTION_VALUE_TEXT
1	S2001	CYCLE1D1	PE_V1	1	1	PETEST	甲状腺	(null)
2	S2001	CYCLE1D1	PE_V1	1	1	PEORRES	正常	(null)
3	S2001	CYCLE1D1	PE_V1	1	1	PEDESC	(null)	(null)
4	S2001	CYCLE1D1	PE_V1	1	2	PETEST	淋巴结	(null)
5	S2001	CYCLE1D1	PE_V1	1	2	PEORRES	正常	(null)
6	S2001	CYCLE1D1	PE_V1	1	2	PEDESC	(null)	(null)
7	S2001	CYCLE1D1	PE_V1	1	3	PETEST	头、颈	(null)
8	S2001	CYCLE1D1	PE_V1	1	3	PEORRES	正常	(null)
9	S2001	CYCLE1D1	PE_V1	1	3	PEDESC	(null)	(null)
10	S2001	CYCLE1D1	PE_V1	1	4	PETEST	耳、鼻、眼、喉	(null)
11	S2001	CYCLE1D1	PE_V1	1	4	PEORRES	正常	(null)
12	S2001	CYCLE1D1	PE_V1	1	4	PEDESC	(null)	(null)
13	S2001	CYCLE1D1	PE_V1	1	5	PETEST	心血管，肺	(null)
14	S2001	CYCLE1D1	PE_V1	1	5	PEORRES	正常	(null)
15	S2001	CYCLE1D1	PE_V1	1	5	PEDESC	(null)	(null)
16	S2001	CYCLE1D1	PE_V1	1	6	PETEST	皮肤粘膜	(null)
17	S2001	CYCLE1D1	PE_V1	1	6	PEORRES	正常	(null)
18	S2001	CYCLE1D1	PE_V1	1	6	PEDESC	(null)	(null)

Database underlying table structures

- Vertical to Horizon (Views)

```
CREATE VIEW CD19CART$STABLE.PEPE2 AS
SELECT RDCM.SITE INVSITE,
  RDCM.INVESTIGATOR INV,
  RDCM.PATIENT PT,
  RDCM.CLIN_PLAN_EVE_NAME CPEVENT,
  RDCM.VISIT_NUMBER VISIT_NUMBER,
  to_number(RDCM.VISIT_NUMBER
  ||'.'
  || ltrim(TO_CHAR(RDCM.SUBEVENT_NUMBER,'09'))) ACTEVENT,
  RDCM.SUBEVENT_NUMBER SUBEVENT_NUMBER,
  RDCM.DOCUMENT_NUMBER DOCNUM,
  DCMS.NAME DCMNAME,
  DCMS.DCM_SUBSET_SN SUBSETSN,
  RDCM.QUALIFYING_VALUE QUALIFYING_VALUE,
  RDCM.LAST_DATA_CHANGE_TS LSTCHGTS,
  R.REPEAT_SN REPEATSN,
  MAX( DECODE(r.dcm_question_id,623201,SUBSTRB(R.VALUE_TEXT,1,200),NULL)) PETEST,
  MAX( DECODE(r.dcm_question_id,623201,SUBSTRB(DVG_VAL(10701,'DVG_LONG_VALUE', R.VALUE_TEXT,-1,-1,NULL, -1),1,199),NULL)) PETEST_DLV,
  MAX( DECODE(r.dcm_question_id,623201,NVL(R.EXCEPTION_VALUE_TEXT, R.VALUE_TEXT),NULL)) PETEST_FUL,
  MAX( DECODE(r.dcm_question_id,623101,SUBSTRB(R.VALUE_TEXT,1,30),NULL)) PEORRES,
  MAX( DECODE(r.dcm_question_id,623101,SUBSTRB(DVG_VAL(5701,'DVG_LONG_VALUE', R.VALUE_TEXT,-1,-1,NULL, -1),1,199),NULL)) PEORRES_DLV,
  MAX( DECODE(r.dcm_question_id,623101,NVL(R.EXCEPTION_VALUE_TEXT, R.VALUE_TEXT),NULL)) PEORRES_FUL,
  MAX( DECODE(r.dcm_question_id,623301,SUBSTRB(R.VALUE_TEXT,1,200),NULL)) PEDESC,
  MAX( DECODE(r.dcm_question_id,623301,NVL(R.EXCEPTION_VALUE_TEXT, R.VALUE_TEXT),NULL)) PEDESC_FUL
FROM rdcms_view rdcm,
  dcms,
  responses_view r
WHERE rdcm.clinical_study_id = 18001
AND r.received_dcm_id        =rdcm.received_dcm_id
AND r.response_entry_ts     <= rdcm.as_of_ts
AND r.end_ts                 > rdcm.as_of_ts
AND dcms.dcm_subset_sn       = rdcm.dcm_subset_sn
AND dcms.dcm_layout_sn       = 1
AND ((dcms.dcm_id            = 57901
AND rdcm.dcm_id              = 57901
AND r.dcm_question_group_id IN (91401)))
GROUP BY 'CD19CART'
  || NULL,
  RDCM.SITE,
```

Database underlying table structures

- Vertical to Horizon (Views)

	INVSITE	PT	CPEVENT	DCMNAME	SUBSETSN	REPEATSN	PETEST	PEORRES	PEDESC	PEDESC_FUL
▸1	03	0313	VISIT 7	PE_V1	1	1	皮肤	正常		
2	03	0313	VISIT 7	PE_V1	1	2	淋巴结	正常		
3	03	0313	VISIT 7	PE_V1	1	3	眼耳鼻喉口	异常	口腔溃疡	口腔溃疡
4	03	0313	VISIT 7	PE_V1	1	4	呼吸系统	正常		
5	03	0313	VISIT 7	PE_V1	1	5	心血管系统	正常		
6	03	0313	VISIT 7	PE_V1	1	6	消化系统	正常		
7	03	0313	VISIT 7	PE_V1	1	7	内分泌和代谢系统	未查		
8	03	0313	VISIT 7	PE_V1	1	8	泌尿生殖系统	未查		
9	03	0313	VISIT 7	PE_V1	1	9	骨骼肌肉系统	正常		
10	03	0313	VISIT 7	PE_V1	1	10	神经精神系统	正常		
11	03	0313	VISIT 7	PE_V1	1	11	其它	未查		

Dataset generation

- From tables (Views) to SAS dataset

```
/********************** View  DM ***************************/
proc sql;
connect to odbc(dsn=OCPROD1 uid=nnras1  pwd=xxx);
create table result.DM as select
STUDY label='Clinical Study' format $15.
,INVSITE label='Site' format $10.
,INV label='Investigator' format $10.
,PT label='Patient' format $10.
,CPEVENT label='CPE Name' format $20.
,VISIT_NUMBER as VISITNUM label='Visit Number' format 10.
,SUBEVENT_NUMBER as SUBEVENT label='Sub Event' format 10.
,QUALIFYING_VALUE as QUALIFYV label='Qualifying Value' format $70.
,REPEATSN label='Repeat #' format 3.
,BRTHDAT label='出生日期' format $8.
,AGE label='年龄' format 2.
,AGEU label='年龄单位' format $60.
,SEX label='性别' format $30.
,AGED label='Derived Age' format 4.1
,RACE label='Race' format $30.
,RACEOTH label='Other Race' format $200.
,HEIGHT label='Height' format 8.
,HEIGHTU label='Unit of height' format $30.
,WEIGHT label='Weight' format 9.2
,WEIGHTU label='Unit of weight' format $30.
from connection to odbc (
SELECT * FROM ASAP$CURRENT.DM order by pt,visit_number,subevent_number,qualifying_value,repeatsn);
disconnect from odbc;
quit;
```

Variable Name
Label
Format

Dataset common & extra variables

- Common Variables

	OC/RDC	Rave	Inform
Study Name	STUDY	Project	
Site Name	INVSITE	Site	SITENAME
Investigator Name	INV		
Subject ID	PT	Subject	SUBJECTNUMBERSTR
Event Name	CPENAME	InstanceName	VISITMNEMONIC
Event RepeatNo		InstanceRepeatNumber	VISITINDEX
Form Name	DCMNAME	DataPageName	FORMMNEMONIC
Form RepeatNo	SUBEVENT	PageRepeatNumber	FORMINDEX
ItemGroup Name	QUESTIONGROUP		
ItemGroup RepeatNo			
Item Name			

Extra Variables

- OC/RDC

列名	类型	长度	输出格式
STUDY	文本	11	$15.
INVSITE	文本	10	$10.
INV	文本	10	$10.
PT	文本	10	$10.
CPEVENT	文本	20	$20.
VISITNUM	数字	8	10.
SUBEVENT	数字	8	10.
QUALIFYV	文本	70	$70.
REPEATSN	数字	8	3.
INJNO	数字	8	1.
INJNO_FUL	文本	200	$200.
CMTRTCN	文本	200	$200.
CMTRTCN_FUL	文本	200	$200.
CMROUTE	文本	200	$200.
CMROUTE_FUL	文本	200	$200.
CMROUOTH	文本	200	$200.
CMROUOTH_FUL	文本	200	$200.
CMINDC	文本	200	$200.
CMINDC_FUL	文本	200	$200.
CMSTDAT	文本	8	$8.
CMSTDAT_FUL	文本	200	$200.
CMONGO	文本	6	$6.
CMONGO_FUL	文本	200	$200.
CMENDAT	文本	8	$8.
CMENDAT_FUL	文本	200	$200.

- Item
  - Conform data
- Item_FUL
  - Raw data (including non-conform data)
- Item_DLV
  - Codelist label

Extra Variables

- Rave (Regular)

列名	类型	长度	输出格式
DataPageName			
PageRepeatNumber			
RecordDate			
RecordId			
RecordPosition			
MinCreated			
MaxUpdated			
SaveTS			
ICDAT			
ICDAT_RAW			
ICDAT_INT			
ICDAT_YYYY			
ICDAT_MM			
ICDAT_DD			
ICYN			
ICYN_STD	数字	8	1.
BRTHDAT	数字	8	DATETIM...
BRTHDAT_RAW	文本	30	$20.
BRTHDAT_INT	数字	8	DATETIM...
BRTHDAT_YYYY	数字	8	4.
BRTHDAT_MM	数字	8	2.
BRTHDAT_DD	数字	8	2.
AGE	数字	8	3.
AGE_RAW	文本	12	$8.
RACE	文本	6	$4.
RACE_STD	数字	8	1.
RACEOTH	文本	600	$400.
SEX	文本	3	$2.
SEX_STD	数字	8	1.
CBPYN	文本	3	$2.
CBPYN_STD	数字	8	1.
HEIGHT	数字	8	3.

- Item
  - Conform data
- Item_RAW
  - Raw data
  - Applies to Num & Datetime
- Item_STD
  - Codelist coded value
  - Applies to Text with codelist

Extra Variables

- Inform

列名	类型	长度	输出格式
CD_COUNT	数字	8	
CREATEDBYUSERID	数字	8	
CREATEDDATETIME	数字	8	DATETIM...
DELETED	文本	4	$4.
DELETEDFORM	文本	4	$4.
DOV	数字	8	DATETIM...
FIRSTDATATIME	数字	8	DATETIM...
FORMCOMMENTTEXT	文本	4000	$4000.
FORMID	数字	8	
FORMINDEX	数字	8	
FORMMNEMONIC	文本	80	$80.
FORMREV	数字	8	
ITMAGE	数字	8	
ITMAGE_ND	文本	256	$256.
ITMBMI	数字	8	
ITMBMI_ND	文本	256	$256.
ITMBMI_U	文本	36	$36.
ITMBRTHDTC	数字	8	DATETIM...
ITMBRTHDTC_DTS	文本	76	$76.
ITMBRTHDTC_ND	文本	256	$256.
ITMFEOTHSP	文本	800	$800.
ITMFESP	文本	136	$136.
ITMFESP_C	数字	8	
ITMFESTAT	文本	112	$112.
ITMFESTAT_C	数字	8	
ITMHEIGHT	数字	8	
ITMHEIGHT_ND	文本	256	$256.
ITMHEIGHT_U	文本	36	$36.
ITMRACE	文本	56	$56.
ITMRACE_C	数字	8	

- Item
  - Item data
- Item_ND
  - Intentionally marked blank
- Item_C
  - Codelist coded value
- Item_U
  - Item unit

FMD

方恩医药发展有限公司

FOUNTAIN MEDICAL DEVELOPMENT LTD.

Thank you