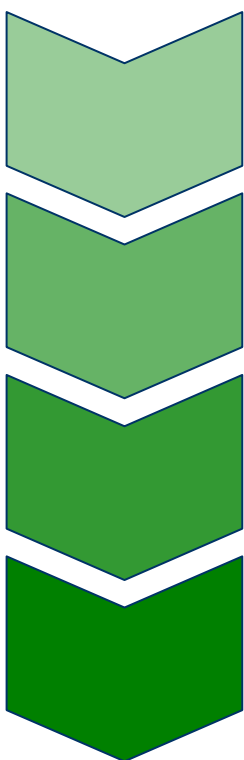




Retour d'expérience sur la vérification de package CDISC pour soumission à la FDA

GUF 2025 - Marie Ober

Emergence d'un besoin



Externalisation

Les CROs préparent des packages CDISC pour des sponsors

Manque d'expertise

Absence de ressources internes pour contrôler la qualité

Besoin de validation

Nécessité de vérifier l'**acceptabilité par les autorités**

Notre réponse

Création d'un département spécialisé en validation de packages CDISC

Que valide-t-on ?

- Respect du standard
 - SDTM
 - ADaM
 - Sponsor
- Cohérence des données et de la documentation
 - dans chaque livrable (dataset, acrf, define.xml, xDRG)
 - entre les livrables
- Respect d'une *guidance* spécifique (FDA, PMDA, NMPA)

Les erreurs les plus courantes

- Datasets
 - Clés
 - Baseline flags & EPOCH
 - *Controlled Terminology* (CT)
 - SUPP
 - Cohérence inter-dataset
- Documentation
 - acrf
 - define
 - csdrg & adrg
- Pinnacle 21 Validation

Les clés

- Bon indicateur de la qualité d'un dataset
 - Important de la chercher sans a priori
 - Représentative du acrf (dans la mesure du possible)
- Bonne clé :
 - Clé naturelle
 - Peu de variables
 - Cohérente avec le define
- Clé suspecte :
 - Contient des variables de type identifier (--SPID, --SEQ)
 - Trop de variables
 - Incohérente avec le define.xml

Baseline flags & EPOCH

Baseline

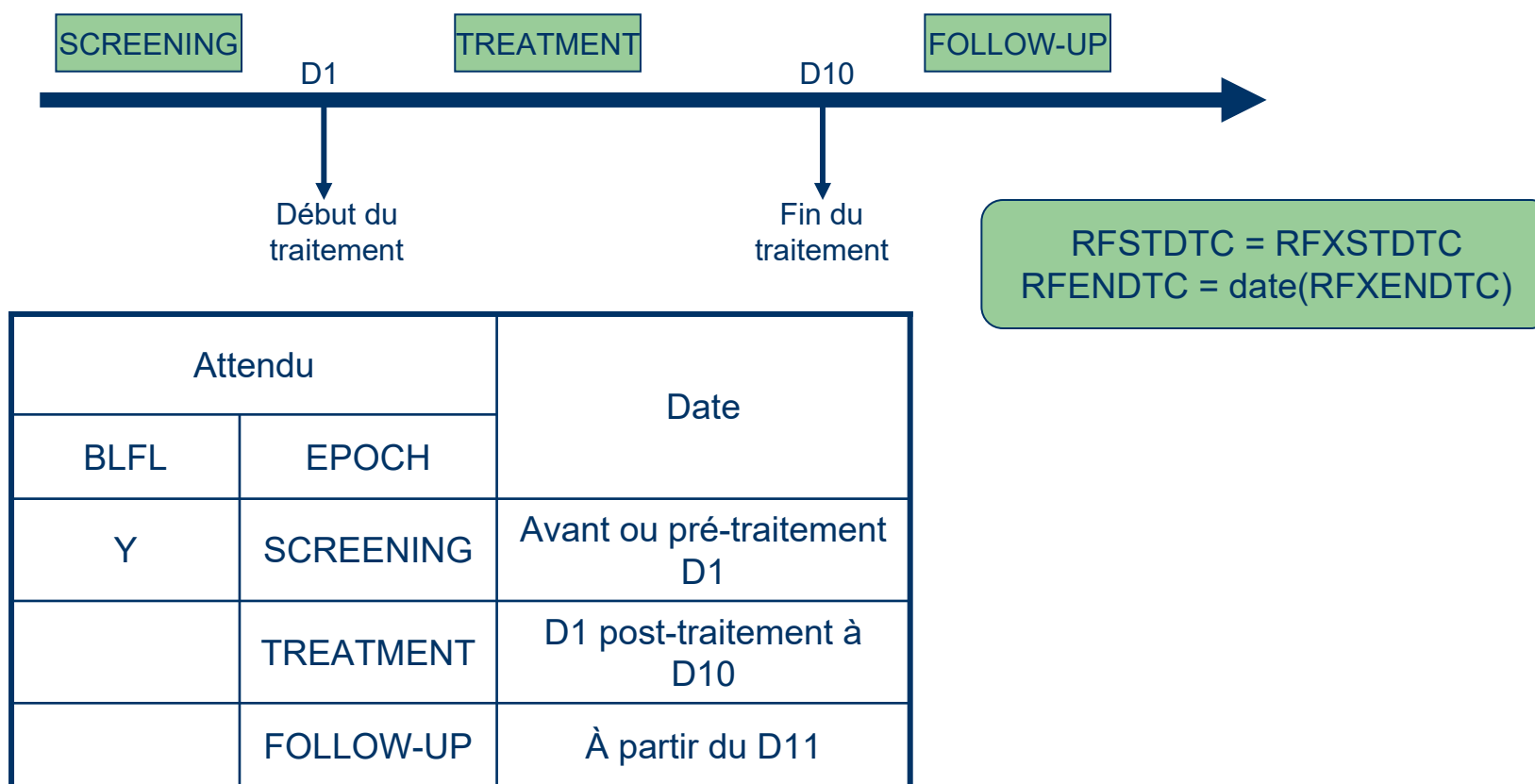
- IG 3.2:
 - BLFL : *Indicator used to identify a baseline value.*
 - Date de référence : RFSTDTC ou RFXSTDTC
- IG 3.3
 - LOBXFL : *Operationally-derived indicator used to identify the last non-missing value prior to RFXSTDTC.*
 - Variable de référence: LOBXFL
 - Date de référence : RFXSTDTC

Baseline flags & EPOCH

EPOCH

- SDTM : *Epoch associated with the start date/time of the observation, or the date/time of collection if start date/time is not collected*
- Comment on dérive EPOCH ?
 - SDTM-IG 3.2 : instructions floues (visite, date de référence)
 - SDTM-IG 3.3 :
 - Clarification section 4.1.3.1 EPOCH Variable Guidance
 - Findings : assignation basée sur la variable --DTC
 - Interventions & event : assignation basée sur la variable --STDTC
- Importance des dates de référence et de SE

Baseline flags & EPOCH



Baseline flags & EPOCH



RFSTDTC = RFXSTDTC
RFENDTC = date(RFXENDTC)

Vu		Date
BLFL	EPOCH	
Y	SCREENING	Avant D1
Y	TREATMENT	Pré traitement au D1
Y	FOLLOW-UP	Avant D1
	FOLLOW-UP	D10

correct

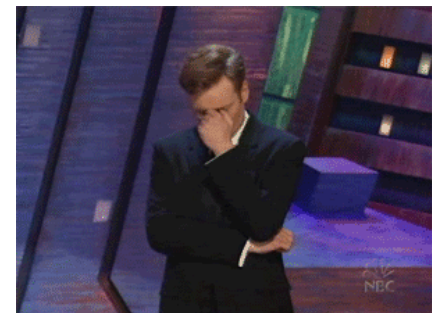
?? TPT = Pre-Dose, EPOCH dérivé avec les visites

?? EPOCH ou dérivation de RFSTDTC & RFENDTC fausse

?? EPOCH ou dérivation RFENDTC ne suit pas la demande du sponsor

La CT (Controlled Terminology)

- *Raw data* utilisées directement sans mapping :
 - LBMETHOD
 - CMFRQ
 - TULOC
- Mauvaise utilisation des valeurs de soumission
 - LBORRESU
 - LBSTRESU
- Codetable non utilisée :
 - EGTESTCD & EGSTRESC
 - QSCAT & QSTESTCD



SUPP

- Trop de SUPP, tue la SUPP
 - Indique généralement un mapping déficient
 - *raw data* sont directement intégrées sans mapping
 - Méconnaissance du standard
 - SUPPVS.VSFAST (Fasting Status)
 - acrf: Was the subject fasting?
 - IG 3.2 et plus: variable principale, “generally not to be used”
 - Max : 201 SUPPs dans 35 datasets

AE	CM	DM	DS	DV	EC	EG	EX	FA	HO	LB	MH	PE	PR	RP	RS	SU	TR	TU	VS
17	17	3	8	2	30	1	20	33	8	7	1	1	19	7	11	7	1	10	1

SUPP

- Variables commentaires en SUPP
 - FDA & CDISC n'en veulent pas
 - Commentaires vont dans CO
 - Exemple
 - RS.ASCOM : Assessment Summary Comments
 - RS.IQCOM : Comments on Image Quality
 - RS.TIMCOM : Timepoint response: Override Comments
 - EG.EGTESTCD = COVAL1
 - EG.EGTESTCD = COVAL2

SUPP

- Deux SUPPs, une notion
 - 2 noms, 2 labels mais le contenu est identique
 - Exemples :
 - PROGSITE :
 - label : What is the Site of Progression ?
 - data : Distant, Local\Regional or Both
 - define : origin CRF
 - PROGTYP :
 - label : Progression Type
 - data : Distant Progression or Locoregional Progression
 - define : origin eDT

Solution : une variable avec QORIG et QEVAL différents et VLM dans le define.xml

La cohérence inter-dataset

- **DM** : tous les sujets sont-ils bien dans DM ?
- **Les visites** : SV, TV, tous les datasets avec VISITNUM
- **DS** : cohérent avec RFSTDTC, RFENDTC, RFPENDTC & DTHDTC
- **Temporalité & design** : SE, DM, EX, TA, TE
- ***Eligibility criteria*** : TI, IE, DM, DS
- **AE-CM** : AE, CM, RELREC
- **La mort** : AE, DS, DM, SS, DD, SE
- **Les tumeurs** : TU, TR, RS

acrf

- Double table des matières manquante
 - par visite
 - par module
- Annotation manquante
 - formulaire sur 2 pages, dataset annoté que sur la 1ère page
 - variable présente sur le acrf & absente des données
 - data : STORTIM
 - acrf : STORTIME
 - variable absente du acrf avec une origine CRF dans le define.xml
- Cohérence de couleur non respectée

define

SDTM

- Codelist incohérente avec les data
- Variables absentes des datasets
- Peu de VLM

ADaM

- Dérivation incompréhensible
- Dérivation fausse
- Codelist incohérente avec les data
- *Origin* erronée
 - *Derived*
 - ADSL.RACE: DM.RACE
 - ADSL.HGTBL: VS.VSTESTCD
 - *Predecessor*
 - ADLB.AVALC : LBSTRESN
 - ADEG.AVAL : EGSTRESC

Reviewer Guides

- csdrg:
 - 2.1 Protocol Number and Title
 - 3.1 Overview
 - 3.3 SDTM Subject Domain
- adrg
 - 4.2 Data Dependencies
 - 5.2 Analysis Datasets



Pinnacle 21

- Erreurs non comprises
 - SD1120 : Comments in SUPPQUAL domain
 - “The respective PE comment is provided in SUPPPE”
 - “Image Quality and Assessment Summary comments related to independant assessor readings have been captured in SUPPRS.”
 - SD0042: --STAT does not equal 'NOT DONE' when --PRESPE='Y' and --OCCUR is null
 - “Populated as per external data”

Pinnacle 21

- Explications minimaliste et non pertinente
 - CT2002: <variable> value not found in 'Unit' extensible codelist
 - “Values mapped per sponsor request.”
 - “Units as collected are added to the extensible codelist”
 - SD1076: Model permissible variable added into standard domain
 - “As sponsor request”

Fini...

