

BIOTRIAL



Retour d'expérience sur l'interprétation des SDTM IG et ADaM IG

***GUF CDISC,
04 Octobre 2016***

*Karine PROVOST, Data Management Project Leader
Hélène TRIBODET, Biostatistics Project Leader*



Avertissement

Le contenu de cette présentation est fourni à titre d'information uniquement, est non exhaustif et ne constitue en aucun cas une recommandation, mais notre expérience.

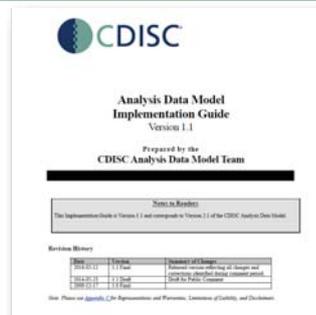
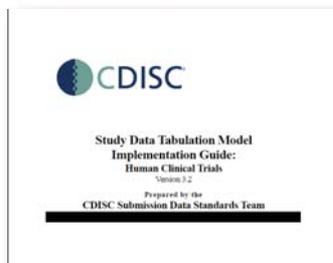


Agenda

Différents process dans une CRO

Interprétations des SDTM IG

Interprétations des ADaM IG



Agenda

Différents process dans une CRO

Interprétations des SDTM IG

Interprétations des ADaM IG



Différents process dans une CRO



Sponsors ont une équipe de gouvernance des standards CDISC et nous fournissent leur guidelines



Sponsors n'ont pas de guidelines et nous utilisons nos conventions



Sponsors ont une base de données dans leur propre format et nous faisons le mapping au format SDTM puis ADaM.



Agenda

Différents process dans une CRO

Interprétations des SDTM IG

Interprétations des ADaM IG



Interprétations des SDTM IG



Interprétations des SDTM IG

Custom domains



- Utiliser le code présent dans la version la plus récente des SDTM Domain Abbreviation Control Terminology s'il existe.

Cod	Load	Load	Code	Code	Code	Code	Code	Code	Code
Cod	st	Extens	Code	Code	Code	Code	Code	Code	Code
	Extens	Code	Code	Code	Code	Code	Code	Code	Code
	Code	Code	Code	Code	Code	Code	Code	Code	Code
C66734		Yes	SDTM Domain Abbreviation	DOMAIN	SDTM Domain Abbreviation	A unique, 2-character domain code used in the regulatory submission process. The domain abbreviation is used consistently throughout the submission, i.e. in the dataset name, as the value of the domain variable within the dataset, and as a prefix for most variable names in the dataset. (CDISC Glossary)	CDISC SDTM Submission Domain Abbreviation Terminology		
C49604	C66734		SDTM Domain Abbreviation	ML	Meal Data	Information regarding the subject's meal consumption, such as fluid intake, amount, form (solid or liquid state), frequency, etc., typically used	Meal Data Domain		

- Créer un code de domaine commençant par les lettres X, Y ou Z. Une lettre ou un numéro peut être utilisé dans la deuxième position.

Interprétations des SDTM IG

CTRL terminology extensible



- Lorsqu'une codeliste est extensible ("Yes" dans la colonne "Codelist Extensible") et que le code recherché n'existe pas alors un nouveau code est créé

↓

Code	Codelist Code	Codelist Extensible (Yes/No)	Codelist Name	CDISC Submission Value	CDISC Synonym(s)	CDISC Definition
C74456		Yes	Anatomical Location	LOC	Anatomical Location	Terminology codelist used for anatomical location within CDISC.
C120521		Yes	Anti-Viral Outcome of Treatment	AVOUTTRT	Anti-Viral Outcome of Treatment	A terminology codelist to describe the anti-viral outcome of treatment.
C124300		Yes	Biospecimen Characteristics Test Code	BSTESTCD	Biospecimen Characteristics Test Code	Terminology relevant to the test codes that describe the characteristics of biospecimens and extracted samples.
C124299		Yes	Biospecimen Characteristics Test Name	BSTEST	Biospecimen Characteristics Test Name	Terminology relevant to the test names that describe the characteristics of biospecimens and extracted samples.
C124297		Yes	Biospecimen Events Dictionary Derived Term	BEDECOD	Biospecimen Events Dictionary Derived Term	A terminology codelist based on the verbatim or preprinted term for the biospecimen events that occur.
C125923		Yes	BRIDG Activity Mood	BRDGMOOD	BRIDG Activity Mood	The modes that may be applied to an entity as it passes through various phases of a life cycle or business process.
C101859		Yes	Cardiac Procedure Indication	CVPRCIND	Cardiac Procedure Indication	The reason that a subject had a cardiac procedure.
C101866		Yes	Cardiac Rhythm Device Failure Manifestation	CRYDFMAN	Cardiac Rhythm Device Failure Manifestation	The effect that a cardiac rhythm device malfunction had on the ability to stimulate the heart.



Interprétations des SDTM IG

CTRL terminology extensible



- Créer un nouveau code en concaténant le code de la codeliste suivi d'un chiffre de 1 à n.

Par exemple:

C74456.000001 pour le premier nouveau code de la codeliste C74456

- Identifier les nouveaux codes par une étoile (Extended Value)



Interprétations des SDTM IG

Planned Time Point Number (--TPTNUM)



■ Trier de 1 à n

Par exemple:

--TPT	--TPTNUM
PREDOSE	1
1 HOUR POST DOSE	2
4 HOURS POST DOSE	3

■ Utiliser les timepoints

Par exemple:

--TPT	--TPTNUM
PREDOSE	0
1 HOUR POST DOSE	1
4 HOURS POST DOSE	4



Interprétations des SDTM IG

Unplanned visits (VISITNUM / VISIT)



■ Attribuer un code unique pour toutes les visites non prévues au protocole avec la variable VISIT renseignée ou non.

Par exemple:

VISIT	VISITNUM
UNSCHEDULED	99
	99

■ Incrémenter le code visite de la visite prévue au protocole de (+.1 à +.n) - Utiliser pour trier les tables dans l'ordre chronologique avec la variable VISIT renseignée ou non.

Par exemple:

VISIT	VISITNUM
WEEK 2	3
WEEK 2 UNSCHEDULED	3.1
	3.2
WEEK 3	4



Interprétations des SDTM IG

Retest/Valeur de contrôle



- Incrémenter le code visite de la visite prévue au protocole de (+.1 à +.n ou +.01 à +.0n) en fonction du nombre de contrôles par visite et décrire dans la variable SVUPDES (Description of unplanned visit) du domaine SV les données retestées.

Par exemple:

Domain EG

VISIT	VISITNUM	EGDTC
WEEK 2	3	2016-06-07T08:00
WEEK 2 UNSCHEDULED	3.1	2016-06-07T08:05

Domain SV

VISIT	VISITNUM	SVUPDES
WEEK 2	3	
WEEK 2 UNSCHEDULED	3.1	ECG Test



Interprétations des SDTM IG

Retest/Valeur de contrôle



- Attribuer le code visite de la visite prévue au protocole et ajouter un flag « Unscheduled test » dans le domaine SUPP--.

Par exemple:

Domain EG

EGSEQ	VISIT	VISITNUM	EGDTC
9	WEEK 2	3	2016-06-07T08:00
15	WEEK 2	3	2016-06-07T08:05

Domain SUPPEG

IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
EGSEQ	15	UNSCHEDULED	Unscheduled Test	Y



Interprétations des SDTM IG

Retest/Valeur de contrôle



- Attribuer le code visite de la visite prévue au protocole et distinguer par la variable --TPTNUM/--TPT.

Par exemple:

Domain EG

EGSEQ	VISIT	VISITNUM	EGTPTNUM	EGTPT	EGDTC
9	WEEK 2	3			2016-06-07T08:00
15	WEEK 2	3	99	UNSCHEDULED	2016-06-07T08:05

- Attribuer le code visite de la visite prévue au protocole et distinguer par la variable --SPID.

Par exemple:

Domain LB

LBSEQ	VISIT	VISITNUM	LBSPID	LB DTC
9	WEEK 2	3	1	2016-06-07T08:00
15	WEEK 2	3	2	2016-06-07T14:00



Interprétations des SDTM IG

Screen failure



- ARM reste vide dans le domaine DM (selon Study Data Technical conformance guide - July 2016)
- ARM complété à Screen Failure (SDTM IG v3.2 et les versions précédentes)



Interprétations des SDTM IG



Re-enrollement:
USUBJID must be unique for each subject



First Enrollment 0030001
Screen fail



Second Enrollment 0030002
Subject enrolled



Interprétations des SDTM IG



Re-enrollement:
USUBJID must be unique for each subject

- 2 enregistrements avec un même USUBJID et SUBJID différent – Ajout de SUBJID dans SUPP-- (domaine ≠ DM)

Par exemple.

Domain DM

USUBJID	SUBJID
STUDY1-0030001	0030001
STUDY1-0030001	0030002

Domain SV

USUBJID	VSSEQ
STUDY1-0030001	1
STUDY1-0030001	50

Domain SUPPVS

USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
STUDY1-0030001	VSSEQ	1	SUBJID	Subject Identifier for the Study	0030001
STUDY1-0030001	VSSEQ	50	SUBJID	Subject Identifier for the Study	0030002



Interprétations des SDTM IG



Re-enrollement:
USUBJID must be unique for each subject

- Le dernier enrôlement dans le domaine parent et les informations du premier enrôlement dans le domaine SUPP--



Interprétations des SDTM IG



Re-enrollement:
USUBJID must be unique for each subject

Par exemple:

Domain DM

USUBJID	SUBJID
STUDY1-0030002	0030002

Domain SUPPDM

USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
STUDY1-0030002			PSUBJID	Previous Subject Identifier for the Study	0030001

Domain DS

USUBJID	DSSEQ
STUDY1-0030002	1

Domain SUPPDS

USUBJID	IDVAR	IDVARVAL	QNAM	QLABEL	QVAL
STUDY1-0030002	DSSEQ	1	PSUBJID	Previous Subject Identifier for the Study	0030001
STUDY1-0030002	DSSEQ	1	PDSSTDTC	Previous Informed Consent Obtained	2016-06-07



Interprétations des SDTM IG

Normal ranges ($<X$ vs $\leq X$)



- Renseigner Reference Range Lower Limit in Original Units (--ORNRLO) égal à 0 et Reference Range Upper Limit in Original Unit (--ORNRHI) à " $<X$ ".

Par exemple:

Domain LB

LBORNRLO	LBORNRHI	LBSTNRLO	LBSTNRHI	LBSTNRC
0	<100	0		<100
0	100	0	100	

- Renseigner Reference Range Lower Limit in Original Units (--ORNRLO) à vide et Reference Range Upper Limit in Original Unit (--ORNRHI) à " $<X$ ".

Par exemple:

Domain LB

LBORNRLO	LBORNRHI	LBSTNRLO	LBSTNRHI	LBSTNRC
	<100			<100
0	100	0	100	



Interprétations des SDTM IG

Normal ranges ($<X$ vs $\leq X$)



- Renseigner Reference Range Lower Limit in Original Units (--ORNRLO) égal à 0 et Reference Range Upper Limit in Original Unit (--ORNRHI) à " $X-1$ " (définir le nombre de décimales autorisé par paramètre).

Par exemple:

Domain LB

LBORNRLO	LBORNRHI	LBSTNRLO	LBSTNRHI	LBSTNRC
0	99	0	99	
0	100	0	100	



Interprétations des SDTM IG

Randomization number



- Enregistrer dans le domaine SUPPDM
- Enregistrer dans le domaine SUPPDS



Interprétations des SDTM IG

Other specify



- OTHER est ajouté dans la controlled terminology et la description se trouve dans le domaine SUPP--

Par exemple:

Domain EX

EXLOC
OTHER

Domain SUPPEX

QNAM	QLABEL	QVAL
EXLOCOTH	Other Location of Dose Administration	UPPER RIGHT ABDOMEN

- La description est ajoutée dans la controlled terminology

Par exemple:

Domain EX

EXLOC
UPPER RIGHT ABDOMEN



Interprétations des SDTM IG

Inclusion/Exclusion Criteria
IETEST - length >200 characters



- Tronquer à 200 caractères.

Dans le Define.xml colonne Commentaire de l'item IETEST un lien est inséré vers le CRF annoté contenant le texte en entier de chaque critère.

- Reformuler le texte pour ne pas dépasser 200 caractères.

Dans le Define.xml colonne Commentaire de l'item IETEST un lien est inséré vers le CRF annoté contenant le texte en entier de chaque critère.



Interprétations des SDTM IG

Reason Not Performed (--REASND)



- Créer un champ spécifique dans le CRF dès que l'item NOT DONE est coché (--STAT='NOT DONE')

- Collecter les raisons dans le domaine CO puis dériver dans l'item --REASND



Interprétations des SDTM IG

Reference Range Indicator (--NRIND)



- Dérivé à NORMAL quand le résultat est dans les normes
- Reste vide quand le résultat est dans les normes



Interprétations des SDTM IG

Abnormal NCS/CS (--ORRES)



- Renseigner la signification clinique dans SUPP--

Par exemple:

Domain EG

EGTESTCD	EGTEST	EGORRES
INTP	Interpretation	ABNORMAL

Domain SUPPEG

QNAM	QLABEL	QVAL
EGCLSIG	Clinically Significant	N

- Renseigner la signification clinique dans la variable Résultat (--ORRES)

Par exemple:

Domain EG

EGTESTCD	EGTEST	EGORRES
INTP	Interpretation	ABNORMAL NCS



Interprétations des SDTM IG



Entire examination or a group of tests not done

- Soumettre un enregistrement pour chaque paramètre non fait

Par exemple:

Domain LB			
LBCAT	LBTESTCD	LBTEST	LBSTAT
COAGULATION	INR	Prothrombin Intl. Normalized Ratio	NOT DONE
COAGULATION	PT	Prothrombin Time	NOT DONE
COAGULATION	APTT	Activated Partial Thromboplastin Time	NOT DONE

- Soumettre un enregistrement pour le groupe de paramètres non faits (--TESTCD= « --ALL »)

Par exemple:

Domain LB			
LBCAT	LBTESTCD	LBTEST	LBSTAT
COAGULATION	LBALL	Labs Data	NOT DONE



Interprétations des SDTM IG



EPOCH

- Dériver en fonction de la VISITNUM et/ou VISIT
- Dériver quand $SESTDTC \leq --DTC / --STDTC \leq SEENDTC$

Par exemple:

Domain SE		
SESTDTC	SEENDTC	EPOCH
2016-01-01	2016-01-10	SCREENING
2016-01-10	2016-01-16	TREATMENT
2016-01-16	2016-01-30	FOLLOW-UP

Domain EG	
EGDTC	EPOCH
2016-01-15	TREATMENT



Agenda

Différents process dans une CRO

Interprétations des SDTM IG

Interprétations des ADaM IG



Interprétations des ADaM IG

Challenges

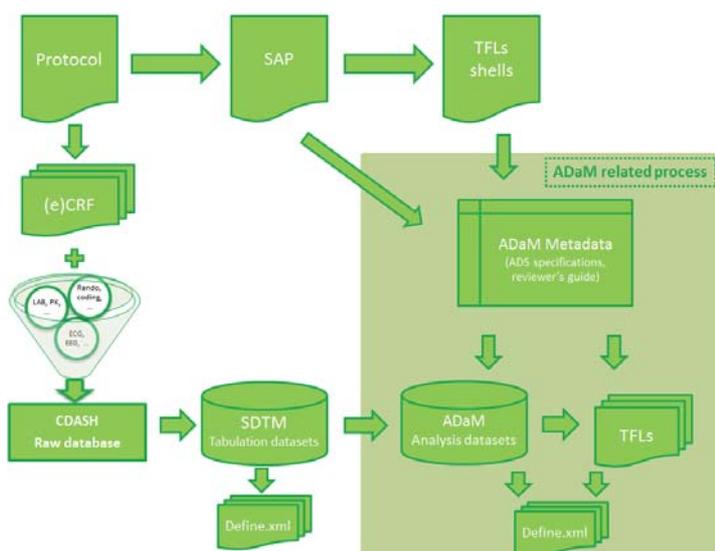
- Traçabilité et report des données sources à déterminer ?
- Suivre les principes ADaM pour être « analysis-ready » ou « one-proc-away » au risque de multiplier les flags et alourdir les datasets ?
- Complexité de la structure ADaM non adaptée pour la production des listings ?
- ADSL
- AVAL / AVALC
- Production des domaines non définis ?



Interprétations des ADaM IG

Pour être ADaM compliant

- Le plus simple et propre est de partir d'une base SDTM compliant.
- La plupart des sponsors ont maintenant adopté le SDTM ou développé un SDTM+.



Interprétations des ADaM IG

Traçabilité

- La FDA veut que toutes les sources utilisées pour les dérivations soient disponibles dans les ADS, mais ce n'est pas requis dans l'ADaM IG.
 - C'est donc un **choix du sponsor**.
- Utilisation des variables SRCDOM, SRCVAR et SRCSEQ notamment dans ADTTE, sinon il est suffisant d'utiliser XXSEQ du SDTM.
- Toutes les variables reprises du SDTM doivent être strictement similaires, pas de changement de valeur, label, codelist, metadata, ...



Interprétations des ADaM IG

« Analysis-ready » « One-proc-away »

- C'est-à-dire qu'il doit y avoir un minimum d'étape pour produire les analyses à partir de l'ADS.
- Cela n'inclut pas les manipulations de présentation quand la sortie de SAS n'est pas au template du TFL requis.



- Toutes les dérivations sont alors faites dans l'ADS et cela améliore la qualité et facilite la validation.
- Mais la structure ADaM ne supporte pas l'analyse de critères multiples, comme les analyses de corrélation.



Interprétations des ADaM IG

ADaM non compliant avec la production des listings

- L'ADaM est prévu pour la soumission des données dérivées, pas pour la création des listings de données individuelles.
- Mais avec son adoption dans les process et standards, ADaM est souvent également utilisé comme source pour les listings et peut donc être adapté.
 - Certains sponsors décident d'ajouter des variables pour les listings dans les ADS en nommant les variables additionnelles avec un «_» par exemple.
 - Certains autres choisissent de produire un ADS seulement quand il y a des données dérivées et d'utiliser les SDTM sinon (avec seulement un merge avec l'ADSL dans le programme du listing pour obtenir les informations clés comme le traitement, la population d'analyse ou l'âge et le sexe).



Interprétations des ADaM IG

Analysis Subject Data Level (ADSL) (1/2)

- (Seul domaine) Requis pour toutes les études.
- Pas de contenu global, ni d'ordre des variables imposé, mais doit être **défini par le sponsor** pour une homogénéité des études.
- Les core variables peuvent être adaptées selon les panels pour:
 - Mettre TRTxxA pour les domaines de safety et TRTxxP pour ceux d'efficacité, par exemple;
 - Garder les variables de sous-groupes pour les domaines d'efficacité seulement, par exemple.

A définir dans le ADRG.



Interprétations des ADaM IG

Analysis Subject Data Level (ADSL) (2/2)

- Selon le choix du sponsor, les screening failures peuvent être inclus dans les ADS.

Certains sponsors choisissent de présenter tous les sujets dans l'ADSL (y compris les screening failures) et de ne garder que les sujets analysables dans les autres ADS.

- Plusieurs ADSL possibles pour les pools:

Par exemple: ADSLISE pour ISE (avec seulement les études d'efficacité) & ADSLISS pour ISS (avec toutes les études).



Interprétations des ADaM IG

Analysis Value (AVAL ou AVALC) (1/2)

■ Représente la valeur du paramètre (défini par PARAMCD, PARAM et PARAMN) et est **soit** numérique **soit** caractère, mais **ne** doit **pas** être les 2.

■ C'est seulement toléré quand c'est un mapping exact et doit être rempli pour tout un paramètre.

Par exemple:

- Pour décrire un score d'un questionnaire (AVAL) et ainsi préciser au reviewer la signification avec une correspondance 1:1 dans AVALC.
- Pour trier les valeurs d'un paramètre avec une réponse en caractère (AVALC) avec une correspondance 1:1 dans AVAL pour chaque valeur du paramètre.



Interprétations des ADaM IG

Analysis Value (AVAL ou AVALC) (2/2)

■ AVALC est parfois populé en tant que champ texte de la valeur numérique AVAL, pour être utilisé dans les listings.

Par exemple:

Quand il y a une valeur BLQ et que la valeur d'analyse est substituée par la limite de quantification: AVAL=30 et AVALC='< 30', alors qu'il faut utiliser la variable source dans le listing (par exemple LBSTRESC) et laisser AVALC vide.

■ AVAL ou AVALC doit être créé dans les ADS suivant la BDS même si c'est une copie de la source sans modification.

Par exemple:

AVAL=LBSTRESN ou AVALC=LBSTRESC.



Interprétations des ADaM IG

Domaines non définis

Si un domaine ne rentre pas dans les structures définies (ADSL, BDS, OCCDS, TTE):

- Suivre les principes ADaM pour:
 - Créer de nouvelles observations si nécessaire,
 - Ajouter des flags;
- Garder la structure du SDTM,
Par exemple, pour EX notamment en transposant les données de SUPPEX;
- Copier les variables core de l'ADSL;
- Utiliser DTYPE (plutôt que PARAMTYP qui va être supprimé de la guideline).



Interprétations des ADaM IG

Pharmacocinétique (PK)

- Pas de guideline disponible pour l'analyse de la PK
 - Mais l'équipe ADaM y travaille
- ADPC & ADPP à partir de SDTM.PC, PP et EX en suivant la BDS (suffisamment flexible pour décrire ces données).
 - Incluant la population de PK:
 - qui doit être définie dans l'ADSL,
 - quand données PK disponibles après gel, tolérance pour que la population spécifique soit seulement présente dans l'ADS correspondant.
 - Gestion des BLQ, données manquantes ou aberrantes (AVAL).
 - Dérivation de flags:
 - spécifiques aux analyses de PK (CRITyFL),
 - données analysables (ANLzzFL),
 - et déviations entre temps théorique et actuel (ARELTM).
- Dérivations de paramètres additionnels.



Interprétations des ADaM IG

Taille des fichiers

Quand la taille de l'ADS est trop importante, on peut diviser le domaine en plusieurs datasets, mais il n'y a pas de convention de découpe et de nom:

- Découpe par catégorie le plus souvent:
Par exemple: ADLBH ou ADHEAM pour l'hématologie, et ADLBC ou ADCHEM pour la biochimie: **à définir par le sponsor.**
- Attention à prévoir cela pendant la programmation pour que les programmes de TFLs pointent bien vers ces domaines, à moins que la découpe soit juste faite pour la soumission: **à définir par le sponsor.**
- Cela doit être précisé dans l'ADRG.



Interprétations des ADaM IG

ADS spécifications

- Elles ont 2 objectifs:
 - Préciser les dérivations à faire par le programmeur
 - Source de metadata pour le define.xml
- Souvent produites dans Excel:
 - Avantages:
 - Largement disponible
 - Lisible informatiquement pour l'intégration dans le define.xml
 - Pratique pour définir les ADS par onglet et avec les variables prédéfinies
 - Inconvénients:
 - Pas de suivi de modifications (audit trail), il faut identifier les changements manuellement
 - Facile de faire des erreurs/fautes de frappe
 - Pas de contrôle de cohérence



Interprétations des ADaM IG

Analysis Results Metadata (ARM)

- Attendu dans le define-xml 2.0.
- Pour le moment, peu requis par les sponsors.
- Seulement l'analyse principale et/ou les tables avec des modèles statistiques quand demandé.
- Pas d'outils adaptés disponibles pour le moment.
 - SAS y travaille



Interprétations des ADaM IG

Analysis Data Reviewer's Guide (ADRG)

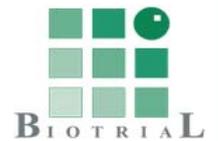
- Optionnel, mais fortement recommandé.
- Template du Phuse disponible.
 - Mieux de suivre les sections et mettre non applicable si besoin
 - Pas obligatoire de répéter les versions des dictionnaires et terminologies contrôlées, mais c'est plus complet.
- Doit préciser les conventions du sponsor.



Interprétations des ADaM IG

Remarque: versions de l'ADaM IG à utiliser

- La FDA demande un Study Data Standardisation Plan (SDSP) définissant notamment les versions des guidelines utilisées afin d'homogénéiser les données d'une soumission.
- Peu transmis aux CRO par les sponsors pour les ADaM.



SDTM / ADaM Implementation FAQ

■ Le Phuse a créé un groupe de travail pour compiler les questions sur l'implémentation des IG et ils les transmettront à la FDA:

Issues captured led to six primary topic areas:

- I. Data Submission (SDTM and ADaM)
- II. Validation/Conformance Rules
- III. IG Nuances
- IV. Therapeutic Area Specific
- V. Legacy Mapping to SDTM
- VI. Trial Design Domains



BIOTRIAL



***Merci de votre
attention !***

Karine PROVOST - karine.provost@biotrial.com
Hélène TRIBODET – helene.tribodet@biotrial.com

