



The Missing Link: Results Level Metadata

Jeffrey M. Abolafia
Chapel Hill, NC

Agenda

- Background
- Why Metadata
- New Use for Old Idea
- TFL Library
- Displays
- The Metadata
- Tools and Programs
- Define File
- Management Tracking
- Benefits

1. Background

Background

Metadata systems used effectively for

- Data collection
- Standardized operational (clinical) data
- Standardized Analysis data

Lots of papers and presentations

Not extended to analysis and reported

Not included in Define.xml

Background: Why Metadata

Current *de facto* submission requirements

CDISC -> standards represented as metadata

CDISC Metadata Requirements

Timelines, budgets, resources.....



2. New Use of Old Idea

Metadata Use for Datasets

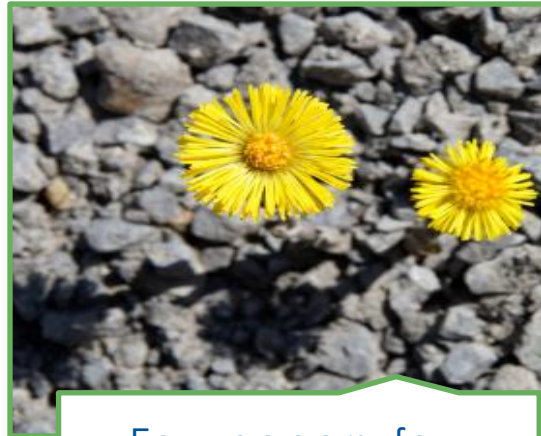
- Specifications stored as machine readable metadata
- Re-used throughout project life cycle
- Programming specifications and documentation
- Input for dataset creation programs
- Extended to produce input for define defile
- Specifications for all
- Traceability

3. Metadata for Displays & Analyses

Metadata for Displays & Analysis



Lags behind
database production



Few papers, few
systems



Rarely a part of
define file

Why Metadata for Displays & Analysis



1 2 3

Functionality of metadata systems for datasets can be easily extended to analysis

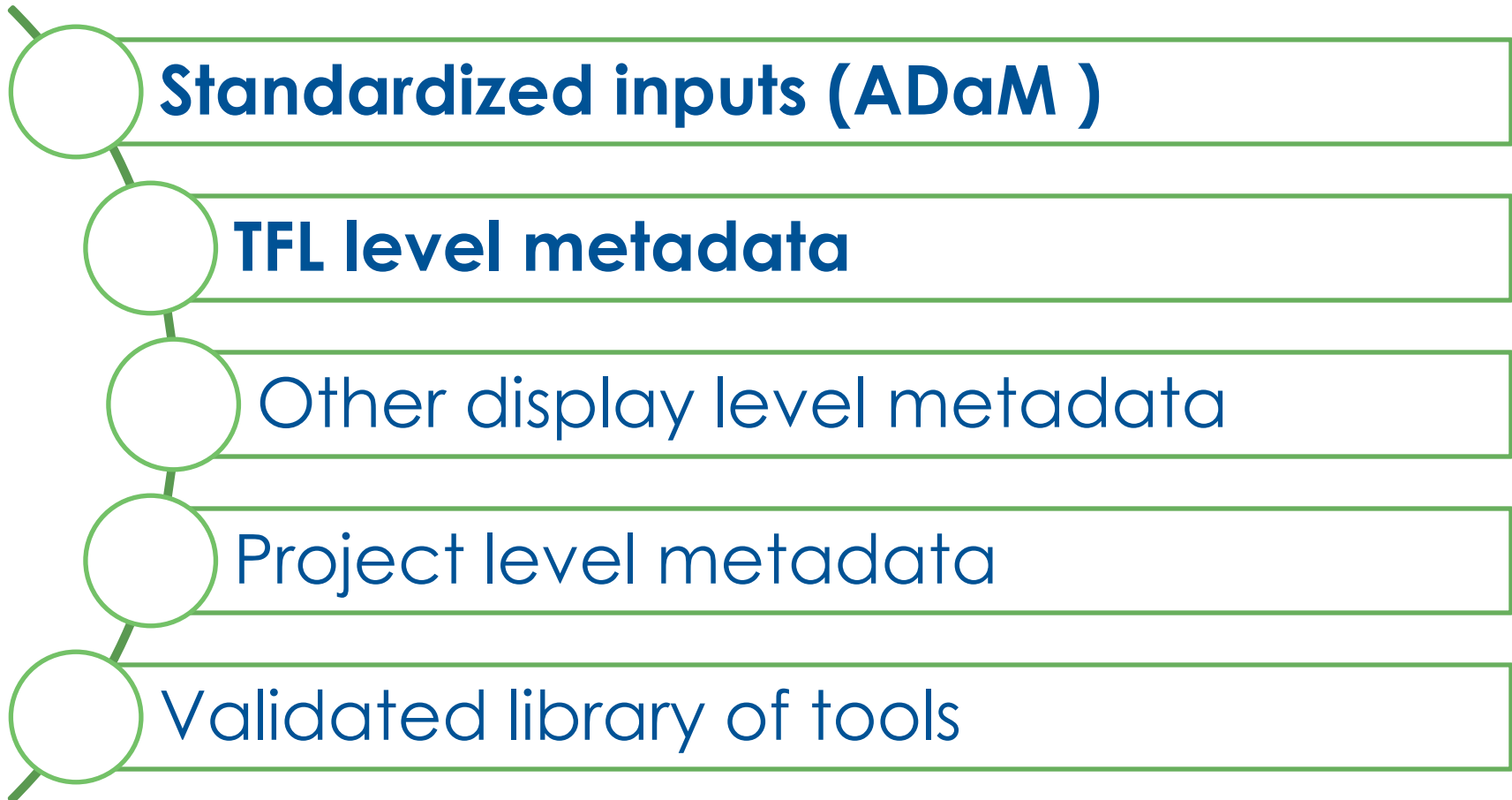


Traceability should start with *results* not *analysis* datasets



Adds Value to a submission

Components for Efficient Display Production



Note: TFL → Tables, Figures, and Listings

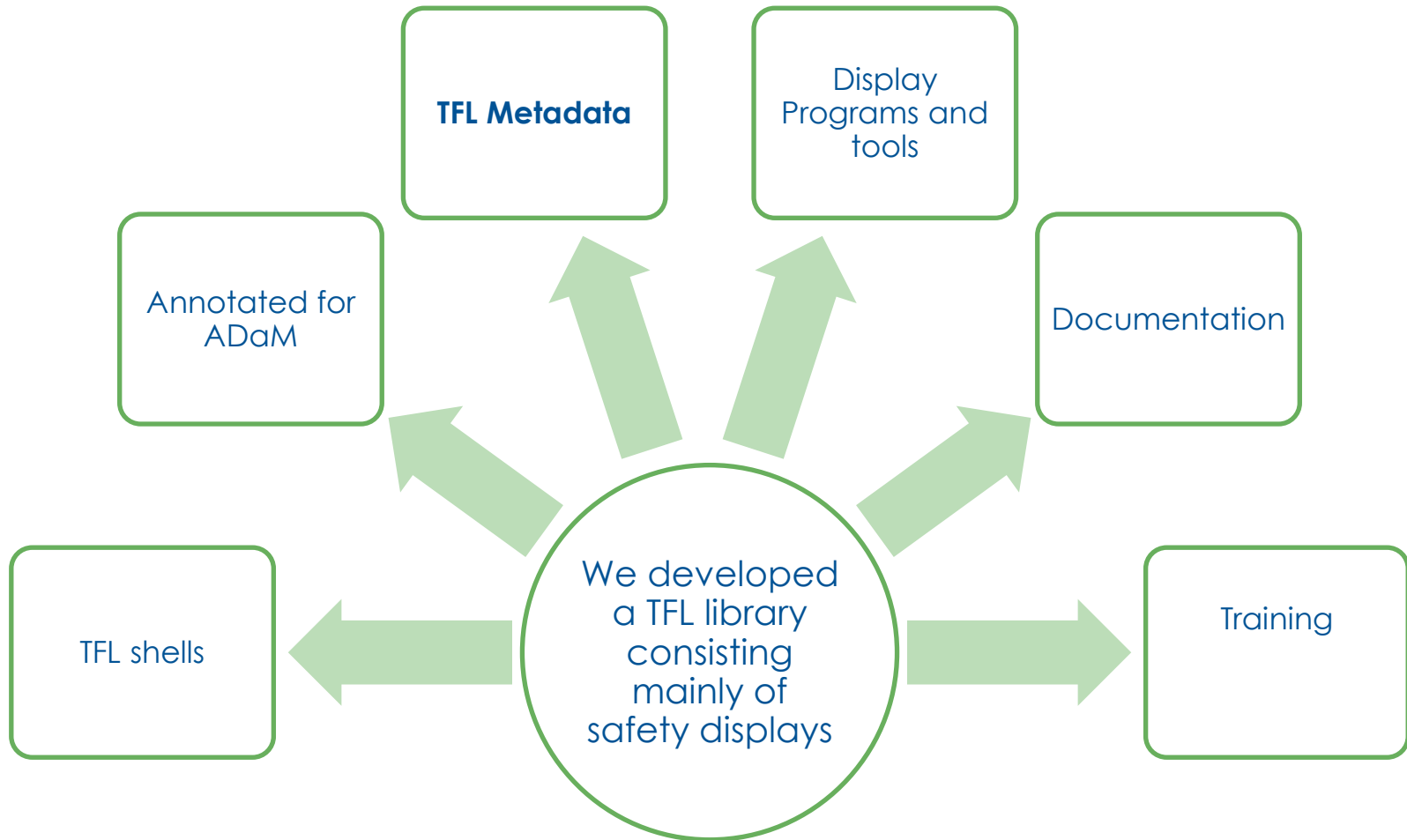
Standard Displays

We evaluated displays from over 15 sponsors across a wide spectrum of therapeutic areas

Result

- There are many displays that are common across sponsor and therapeutic area
- Most common displays were safety related
- With well designed metadata, the same display shell can be re-used many times throughout a project and across domains and projects
- Same domains as CDASH (i.e. AE PE VS MH ECG)
- **Most of what changes from project to project can be extracted from display and project level metadata**

Standard Displays: TFL Library



Standard Displays: Sample Shell

XXXX Pharmaceuticals, Inc.
NDA xx-xxx
Integrated Summary of Efficacy



Table 8.7.3.9
Demographics and Background Characteristics
Studies xx-xxx, xxx-xx, and xx-xxx
Intent to Treat Population

Metadata

Characteristic	Placebo	0.5% XXXX	1.0% XXXX	All Treatment Groups Combined	P-Value
Artificial Tear Use					
No	00 (00%)	00 (00%)	00 (00%)	00 (00%)	0.000
Yes	00 (00%)	00 (00%)	00 (00%)	00 (00%)	
Qualifying Eye					
Both	000 (00%)	000 (00%)	000 (00%)	000 (00%)	0.000
Neither	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Right	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Left	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Worst Symptom					
Burning	000 (00%)	000 (00%)	000 (00%)	000 (00%)	0.000
Foreign Body	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Itching	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Photophobia	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Vision Blurred	000 (00%)	000 (00%)	000 (00%)	000 (00%)	

Formatting

Note: The Intent to Treat (modified) population consists of all subjects randomized and receiving at least one dose of study drug who also completed at least one post-baseline efficacy assessment.

Standard Displays: Sample Shell

xxxx Pharmaceuticals, Inc.
 NDA xx-xxx
 Integrated Summary of Efficacy

**ADaM
 Annotation**

**PARAM /
 AVALC**

TRTP

Table 8.7.3.9
 Demographics and Background Characteristics
 Studies xx-xxx, xxx-xx, and xx-xxx
 Intent to Treat Population

Characteristic	Placebo	0.5% xxxx	1.0% xxxx	All Treatment Groups Combined	P-Value
Artificial Tear Use					
No	00 (00%)	00 (00%)	00 (00%)	00 (00%)	0.000
Yes	00 (00%)	00 (00%)	00 (00%)	00 (00%)	
Qualifying Eye					
Both	000 (00%)	000 (00%)	000 (00%)	000 (00%)	0.000
Neither	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Right	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Left	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Worst Symptom					
Burning	000 (00%)	000 (00%)	000 (00%)	000 (00%)	0.000
Foreign Body	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Itching	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Photophobia	000 (00%)	000 (00%)	000 (00%)	000 (00%)	
Vision Blurred	000 (00%)	000 (00%)	000 (00%)	000 (00%)	

Note: The Intent to Treat (modified) population consists of all subjects randomized and receiving at least one dose of study drug who also completed at least one post-baseline efficacy assessment.

DRAFT

4. Metadata: Display Level



The Metadata: Display Level

Contents – specific information about a single display

- Display number
- Title lines
- Footnote codes
- Datasets used by the table
- Display type (Table, Figure, Listing)
- Population criteria
- Filtering information Treatment and time point information

Structure – one record per display

Display Level Metadata

Number	TableLetter	Title	Population	Population_Label	Selection_Criteria	Footnotes	Dataset
8.7.3.1	DM_TAA	Demographics Characteristics	ITTFL = 'Y'	Intent to Treat Population		ITT5, FT30, FT31	ADSL
8.7.3.2	AE_TAB	Adverse Events by Treatment Group	SAFFL = 'Y'	Safety Population		PPP5, FT30	ADAE
8.7.4.1	VS_TAB	Vital Sign Values by Treatment Group	SAFFL = 'Y'	Safety Population		FT1, FT2, FT3, SRC5	ADVS

Outcome	Treatment	TreatmentValues	TimeVariab	TimePoints	BaselineV	Analysis Variables
	TRT01P	1,2				AGE, WEIGHTB, HEIGHTB
	TRT01P	1,2				ANYAE, AEBODSYS, AEDECOD
AVAL	TRT01P	1,2	AVISIT	1,2,3,4,5	BASE	AVAL

5. Metadata: Footnote Level

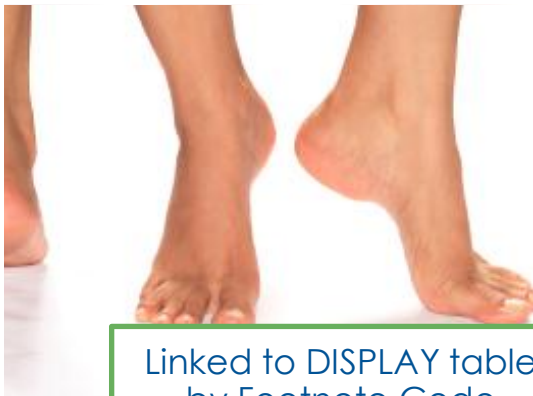
The Metadata: Footnote Level



Contents – actual footnotes used in displays



Structure – one record unique footnote



Linked to DISPLAY table by Footnote Code



Can also be used for Titles

Footnote Level Metadata

Values in the “Tables” sheet column “Footnotes” correspond to row values of “Footnote Code” in the “Footnotes” sheet

N
Footnotes
ITT3, FT30, FT31, FT31b, SR C3

Footnote codes are entered in the order in which the formatted footnote will appear in the table.

	A	B	C	D
1	Footnote Code	Population Label	Study ID Label	Footnote Text
2	ITT3	Intent to Treat Population	Study 03-103	Note: The Intent to Treat (modified) population consists of all subjects randomized and receiving at least one dose of study drug who also completed the baseline and at least one post-baseline assessment of corneal staining, conjunctival staining, and ocular symptoms.
13	SRC3		Study 03-103	Source: Study Report 03-103, Appendix 16.4
14	FT30			Note: Worst symptom is defined as the symptom with the highest (most severe) rating in the study eye at Visit 2, averaging across diary entries for days during the placebo run-in period.
15	FT31			Note: P-values correspond to the test of an overall treatment effect adjusting for site. Categorical variables were assessed via the Cochran-Mantel-Haenszel chi-square test and continuous variables were assessed via analysis of variance.
16	FT31b			Note: For treatment comparisons for race, categories are collapsed to Caucasian and Non-Caucasian.

A change to the footnote text will be reflected in all tables using the footnote. There are no changes to individual programs.



6. Metadata: Project Level

Project Level Metadata

- ✓ Name and description of study
- ✓ Sponsor information
- ✓ Location on network of study
- ✓ Location of study database, format, and macro libraries
- ✓ Other directory locations
- ✓ SAS Options
- ✓ Treatment group information
- ✓ Visit information
- ✓ Client specific preferences for displays

7. Define.xml: Results

Define.xml: Results Level Metadata

Table_14-3.12

Display	Table 14-3.12 Mean NPI-X total score from Weeks 4 through Week 24 - Windowed (Efficacy Population)
AnalysisResult	Descriptive statistics (n, mean, standard deviation, median, min, and max)
Analysis Parameter(s)	NPTOTMN=Mean NPI-X (9) Total (Week 4 to 24)
Analysis Variable(s)	AVAL BASE
Reason	pre-specified in SAP
Data References (incl. Selection Criteria)	ADQSNPIX [EFFFL="Y" and PARAMCD="NPTOTMN" and AVISIT="Weeks 4-24" and ANL01FL="Y"]
Documentation	SAP Section 10.2
Programming Statements	PROC UNIVARIATE data= ADQSNPIX (where =(EFFFL = 'Y' and PARAMCD = 'NPTOTMN' and AVISIT='Weeks 4-24' and ANL01FL= 'Y')); Var BASE AVAL; By TRTPN; Run;

- Used in updated Pilot Project
- If already using to produce displays, only several additional fields needed for Define file
- Adds significant value to submission



Define.xml: Results Level Metadata

Variable	Display Production	Define.xml
Display Identifier	Yes	Yes
Display Name	Yes	Yes
Analysis Result	No	Yes
Analysis Parameters	Yes	Yes
Analysis Variables	Yes	Yes
Reason	No	Yes
Dataset	Yes	Yes
Selection Criteria	Yes	Yes
Documentation	No	Yes
Programming Statements	Maybe	Yes



8. Tools / Metadata Access

Metadata tools

Metadata great, not so great without tools

Tools are necessary to effectively access and use metadata

Standardized Study Setup macro

Gettables macro

RhoTables®

Auto-validation macro

Create “Define” files for FDA submission

Validate define file

Metadata Tools: Display Production

Complete Calling Program (ISE_TAI.SAS)

```
%inc 'S:\RHO\sponsor\project\prog\setup.sas' / nosource2 ;  
%setup(cad=yes);
```

Read Project level metadata to get project level information

```
%GetTable(tbl=TAI);
```

Read display metadata for rows related to table TAI. Write SAS macro variables, then run the shell associated with the table.

Shell Program Excerpts Illustrating Macro Variable Usage

```
data randsel; *Randomized selected for study and subpopulation*;  
  set demo;  
  &WhereStatements.  
run;
```

Created by GetTable, using data from Display metadata.

```
ods pdf file="&PDFdest." style=&style. notoc;  
ods escapechar='`';  
ods listing close;
```

```
&TitleLines.  
&FootnoteLines.
```

Created by Setup macro. It identifies a style sheet that controls the appearance of the PDF output.

Created by GetTable, using data from Titles/Footnotes metadata. The values have the appropriate indentation, page numbers, etc.

Metadata Tools: Define.xml

```
%inc 's:\submissions\macros\util\setup.sas' / nosource2 ;
```

```
%setup(project=S:\RHO\sponsor\project)
```

```
%let crf = &path\Doc\CRF\blankcrf\blankcrf.pdf;
```

```
%defineXMLprep4(type=ADAM, order=meta)
```

```
%defineXML(pdf=no, type=ADAM)
```



9. Management Tracking

Results Level Metadata: Management/tracking

Name	Number	Task	Status	Owner	Created	Started	Completed	Next Due Date	Instructions
TAA_DS_01	14.1.1	Create Program	Completed	kreece	08-Oct-2012 09:37:36 AM	11-Oct-2012 11:12:46 AM	11-Oct-2012 11:12:47 AM		
TAA_DS_01	14.1.1	Create Validation Program	Completed	pnguyen	08-Oct-2012 09:37:36 AM	16-Oct-2012 04:24:23 PM	16-Oct-2012 04:24:23 PM		
TAA_DS_01	14.1.1	Validate Cosmetics	Completed (Pass)	rwoolson	08-Oct-2012 09:37:36 AM	06-Dec-2012 09:50:05 AM	06-Dec-2012 09:50:18 AM		
TAA_DS_01	14.1.1	Validate Statistics	Completed (Pass)	pnguyen	08-Oct-2012 09:37:36 AM	19-Oct-2012 10:53:05 AM	19-Oct-2012 03:52:05 PM		
TAA_DS_01	14.1.1	QC	Superseded		08-Oct-2012 09:37:36 AM				
TAA_DS_01	14.1.1	Revise Program(1)	Completed	kreece	21-Nov-2012 12:39:13 PM	04-Dec-2012 09:42:05 AM	04-Dec-2012 09:42:05 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	Verify Revision(1)	Completed (Pass)	pnguyen	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:43:05 AM	06-Dec-2012 09:43:05 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	Validate Cosmetics(1)	Completed (Pass)	rwoolson	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:50:05 AM	06-Dec-2012 09:50:18 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	Validate Statistics(1)	Completed (Pass)	pnguyen	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:43:08 AM	06-Dec-2012 09:43:08 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator
TAA_DS_01	14.1.1	QC(1)	Completed (Pass)	rwoolson	21-Nov-2012 12:39:13 PM	06-Dec-2012 09:50:28 AM	06-Dec-2012 09:50:36 AM		Update so that all subjects in the Not Randomized column show as discontinued. For all columns, calculate reason for discon % using discontinued as denominator

Results Level Metadata: Management/tracking

								Hide/Show Columns	CSV	PDF	Print
								Search: <input type="text"/>			
Name	Links	Status	Title	Population	Group	Number	Validation Method				
TAD_DM_01 [Tasks]		Started	Demographic and Baseline Characteristics	Intent-to-Treat	Demographics	14.1.3	Double Programming				
TAD_DM_02 [Tasks]		Started	Demographic and Baseline Characteristics	Clinically Evaluable	Demographics	14.1.4	Double Programming				
TAD_DM_03 [Tasks]		Started	Demographic and Baseline Characteristics	Enrolled Safety	Demographics	14.1.5	Double Programming				
TBI_EX_01 [Tasks]		Started	Exposure to Final Dose by Final Dose Group	Enrolled Safety	Exposure and Compliance	14.3.5.5	Double Programming				
TBI_EX_02 [Tasks]		Started	Exposure to Final Dose by Final Dose Group	Randomized Safety	Exposure and Compliance	14.3.5.6	Double Programming				
TAG_EX_01 [Tasks]		Started	Exposure to Treatment	Enrolled Safety	Exposure and Compliance	14.3.5.1	Double Programming				
TAG_EX_02 [Tasks]		Started	Exposure to Treatment	Randomized Safety	Exposure and Compliance	14.3.5.2	Double Programming				
TAH_FD_01 [Tasks]		Started	Final Dose by Subgroups	Enrolled Safety	Exposure and Compliance	14.3.5.3	Double Programming				
TAH_FD_02 [Tasks]		Started	Final Dose by Subgroups	Randomized Safety	Exposure and Compliance	14.3.5.4	Double Programming				
TAX_AEO_01 [Tasks]		Started	Overview of Adverse Events During Double-Blind Phase	Randomized Safety	Adverse Events	14.3.1.3	Double Programming				
TAW_AEO_03 [Tasks]		Started	Overview of Adverse Events During Entire Study	Enrolled Safety	Adverse Events	14.3.1.4	Double Programming				
TAW_AEO_01 [Tasks]		Started	Overview of Adverse Events During Open-Label Phase	Enrolled Safety	Adverse Events	14.3.1.1	Double Programming				
TAW_AEO_02 [Tasks]		Started	Overview of Adverse Events During Open-Label Phase	Randomized Safety	Adverse Events	14.3.1.2	Double Programming				
TBE_VS_02 [Tasks]		Started	Potentially Clinically Significant (PCS) Vital Sign Values During Double-Blind Phase	Randomized Safety	Vital Signs	14.3.5.10	Double Programming				
TBE_VS_01 [Tasks]		Started	Potentially Clinically Significant (PCS) Vital Sign Values During Entire Study	Enrolled Safety	Vital Signs	14.3.5.9	Double Programming				



10. The Pay-Off

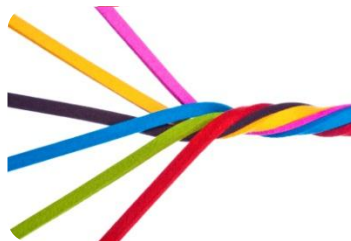
Pay-off



Traditional approach



Table creation program is *greatly* simplified



Handling of display creation is in one place (the metadata) rather than “n” places (separate programs)

Pay-off



Exploits similarities across displays/projects



Reduces manual labor



Increased use of cheaper labor



Increase in automated validation



Improved communication – changes are made in one place: the metadata



Faster, cheaper, and higher quality



Happier Programmers!

The Future

- Better use of SDTM trial design datasets
- ADaM Evolution
- Extend TFL library and metadata
- Develop therapeutic specific displays
- CDISC (or other) Standard set of displays
- Re-define validation requirements

The End

Thank you for attending our presentation.

Send questions or comments to:

Jeff Abolafia (jeff_abolafia@rhoworld.com)