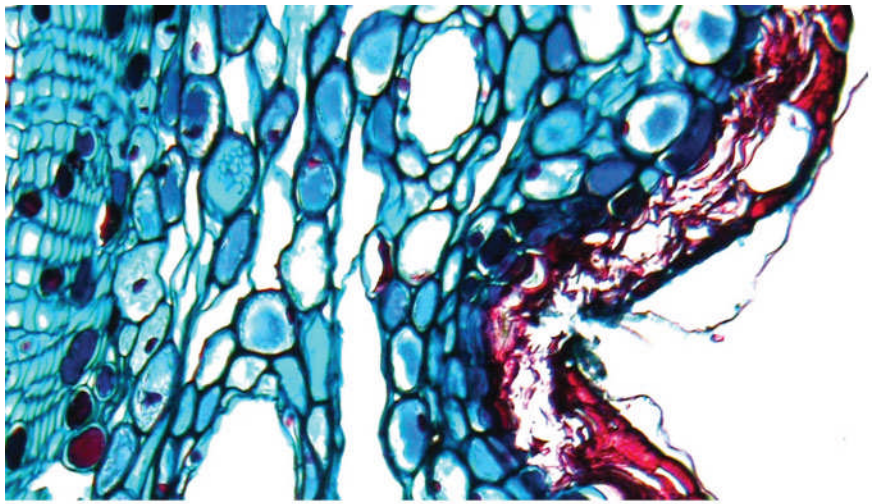
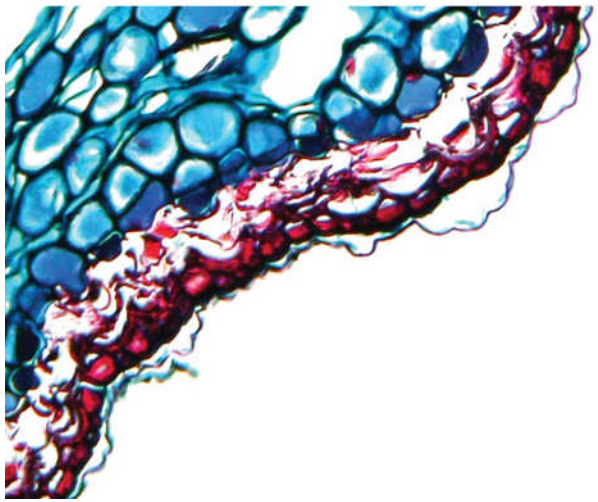


‘From a little SPARC may burst a flame’

-ADaM in the real world

Andrew Tomlinson

11 May 2010



QUINTILES®

Navigating the new health

clinical | commercial | consulting | capital

Contents

- S o – Why am I here?
- P oint of standardisation?
- A DaM – who is he?

  SPARC
- R eal world findings?
- C onclusion

Introduction – Why am I here?

- Motivation to look into

  - New ADaM Implementation guide
  - How this is being applied at Quintiles?
  - Problems encountered?

Motivation: Why Standardise?

- Increased need to standardize / integrate data
- Increased need for efficiency
- FDA has published guidelines recommending the use of SDTM standards
- Increased need for SDTM data for regulatory submissions

– But Also

- Increased quality
- Increased efficiencies -> reduced costs

Who or What is ADaM?

– ADaM (Analysis Data Model)

- "Guides the organization, structure and format of analysis datasets and related metadata."
- Contains key principles that apply to all datasets – not domain specific.
- Supports efficient generation of analysis results ("analysis ready").
- Builds on SDTM

ADaM – Principles

- Analysis datasets and their associated metadata must:
  - facilitate clear and unambiguous communication
  - provide traceability between the analysis data and its source data (ultimately SDTM)
  - be readily useable by commonly available software tools
- Analysis datasets must:
  - be accompanied by metadata
  - be analysis-ready

ADaM – What is does it consist of?

- 2 data structures currently described in ADaM:

- Subject Level Analysis Dataset (ADSL)
  - One record per subject

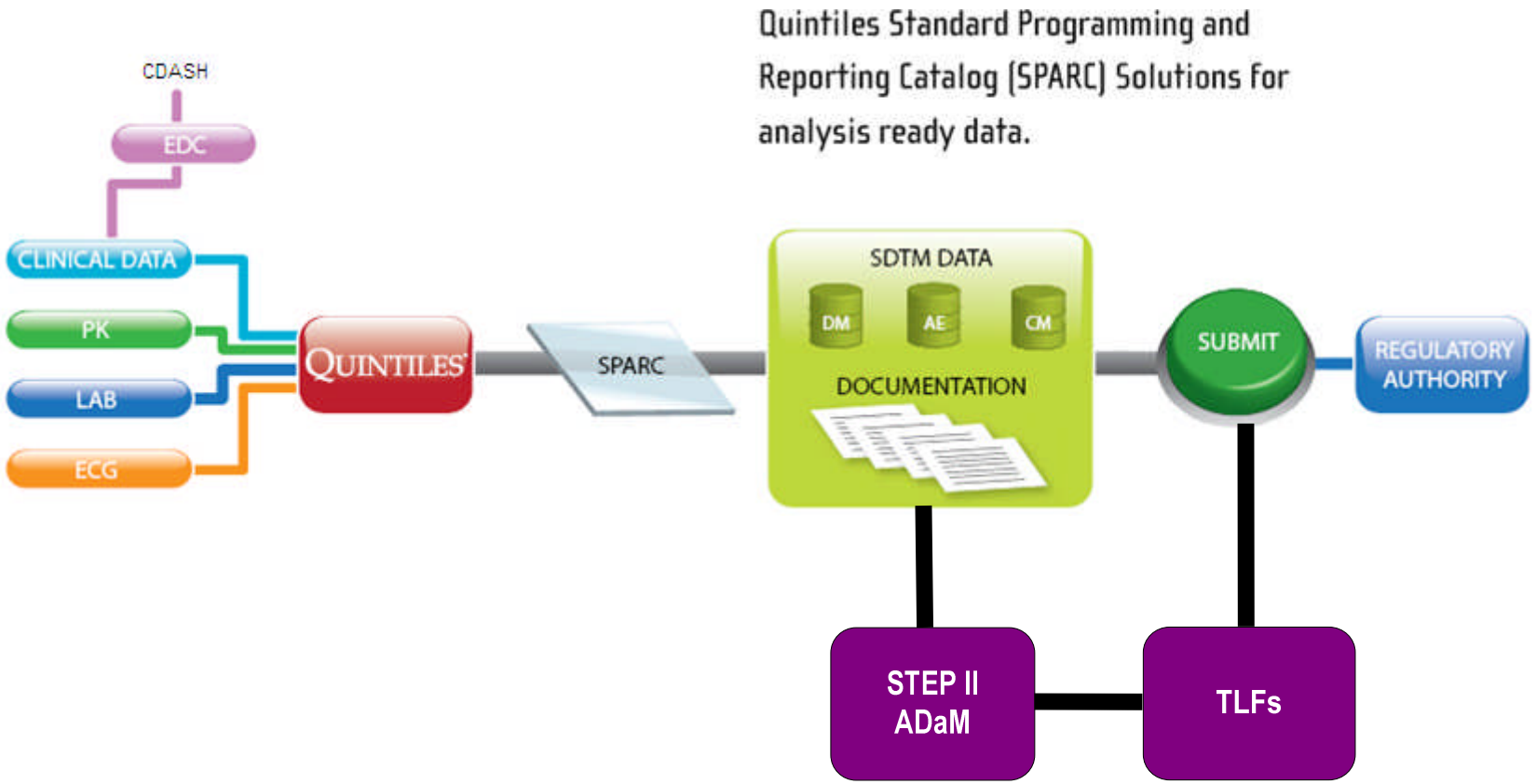
- Basic Data Structure (BDS)
  - One or more record per subject / per analysis parameter / per timepoint

Standard Programming And Reporting Catalogue (SPARC)

- Goal:
  - To develop an innovative solution on an EDC platform to collect, clean, transfer and transform data into an analyzable format using CDISC standards and current processes / department functions
  - SPARC: Standard Programming and Reporting Catalog
  - Step I: CDASH and SDTM (V3.1.2)
  - Step II: ADaM (V1.0)

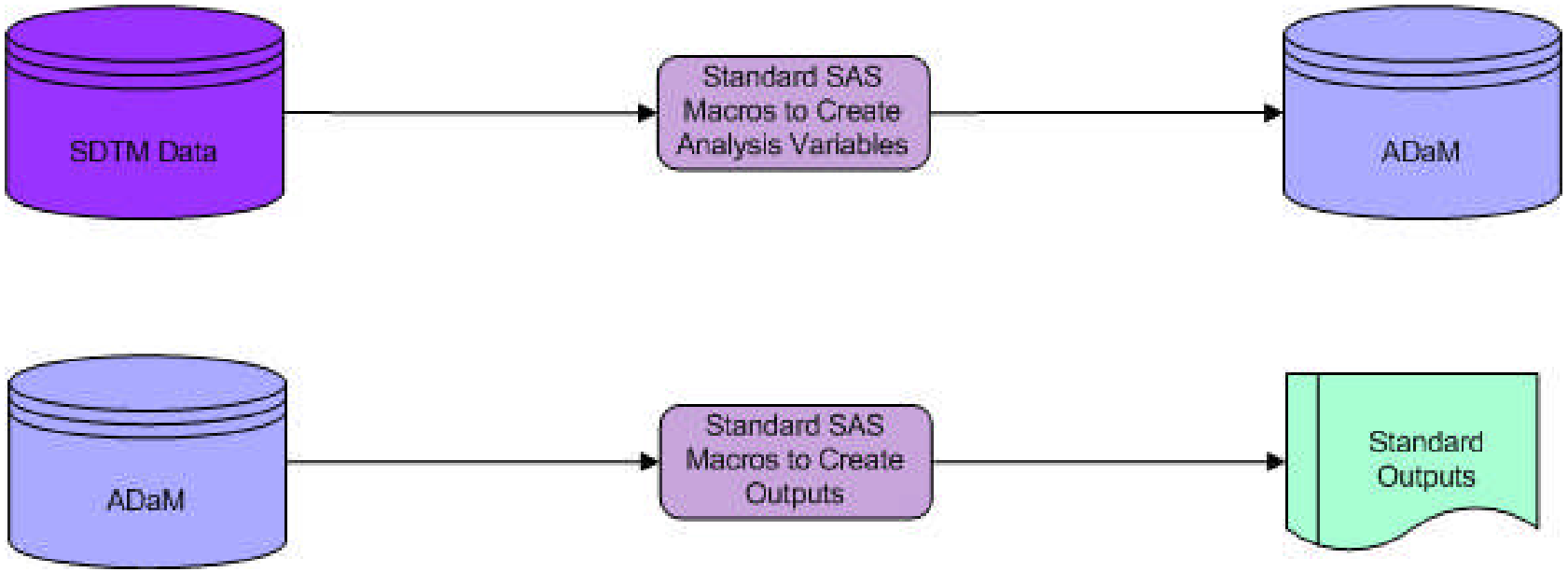
- Joint effort from various departments
  - Data Management
  - Biostatistics
  - Marketing / Business Development

SPARC – process



Data Flow – Biostatistics (ADaM STEP)

– Full SDTM to ADaM



SPARC – Subject level data (ADSL)

– What variables do we keep from ADSL when merging into other datasets?

- Control variable to be kept by:
  – A study specific list of core ADSL variables and additional variables can be added to specific datasets.
  – Use of the metadata specifications to control variables and attributes

- Ability to drop period dates from ADSL when merging the BDS and create a period value by the use of a standard macro.

SPARC – Subject level data (ADSL)

– Can use SDTM to drive part of ADSL creation

- E.g. Trial design datasets such as SDTM.SE (Subject Elements) can be used in SPARC to automate creation of period information in ADSL.

SPARC – Subject level data (ADSL)

- APxx variables
  - Period start and stop variables
- TRxx variables
  - Treatment start and stop variables

  - Generally these might appear to be identical in most parallel group studies but these have to be created as part of a standard approach as the two might differ e.g. crossover studies.

  - Populations will generally be created as part of the SDTM.SUPPDM dataset and these will not be recreated in ADSL. However some populations may have to be created within ADaM e.g. Per protocol populations based on exclusion from a certain time point.

SPARC – Dosing data (ADEX)

- Are all variables relevant to ADSL ?
  - The SPARC team plans to create a standalone ADEX dataset.
    - 1 row per dose level per subject
    - Allows the storage of all dosage information (e.g number of capsules, titration information etc.)
    - Allows information on any study interruptions etc.
    - Could record by visit to allow easier traceability in ADDA or record by date (e.g. titration log) to allow easier merge with data such as AEs.

SPARC – Basic Data Structure

- General structure presented in the implementation guide.
  - Successful in dealing with majority of safety domains in findings data
  - Not so good with incidence or occurrence data (events/interventions in SDTM)

    As the implementation guide says “The BDS was not designed to support analysis of incidence of Adverse event or other occurrence data”. This applies to:

    - AEs, Conmeds, Compliance, Medhist, Exposure

SPARC – Variable control

– SPARC tries to reduce the number of variables to those required by the output.
  - E.g. Lab start and stop dates from SDTM could generate 12 related variables on ADLB

SPARC – Standard outputs

– No listings produced from ADaM datasets by SPARC

– Listings expected to be produced from SDTM data but this would be client specific.

– Currently safety tables are the only standard outputs in the catalogue.

SPARC – Metadata

– Version 2.1 of the Analysis data model gives guidance on metadata.

- Analysis dataset metadata fields
- Analysis variable metadata

– Currently SPARC specifications containing this metadata are created in Excel as specifications. This allows specs to be read electronically.

– This allows programmatic checks between the specification and the ADaM dataset to ensure consistency in terms of variable ordering and format.

SPARC – Traceability

- Traceability i.e. the ability to trace a variable in the ADaM dataset back to its source.

- One of the fundamentals of the ADaM standard.
  - Metadata traceability
    - Through specifications
  - Data point traceability
    - Through a combination of dataset content and specifications

- Section 4.3 of the IG indicates : ‘The ADaM methodology is to include all observed and derived rows for a given analysis parameter’

SPARC – Traceability

- Keep the source data in the ADaM BDS datasets
  - i.e. Keep
    - a parameter identifier describing the original datapoint (PARAM)
    - A code for this description (PARAMCD)
    - The original value (AVAL)
  - Optional:
    - The SDTM domain the data originated from (SRCDOM)*
    - The variable any data originated from (SRCVAR)*
    - The sequence number the data originated from (SRCSEQ)*
    - The Visit from the original dataset
    - The Visitnum from the original dataset

*When combining variables from several SDTM datasets

Traceability – example

– SDTM Data (Vital signs)

Row	USUBJID	VISITNUM	VSSEQ	VSDTC	VSDY	VSTESTCD	VSSTRESN	VSSTRESU
1	10001	1	11	2009-08-25	1	WEIGHT	80	Kg
2	10001	1	18	2009-08-25	1	Height	190	cm

– ADaM Data with derived BMI

USUBJID	VISITNUM	VSDTC	ADY	PARAMCD	AVAL	VSSTRESU	SRCDOM	SRCVAR	SRCSEQ	PARAMTYP
10001	1	2009-08-25	1	WEIGHT	80	Kg	VS	VSSTRESN	11	
10001	1	2009-08-25	1	HEIGHT	190	Cm	VS	VSSTRESN	18	
10001	1	2009-08-25		BMI	22.2	Kg/m^2				DERIVED

Conclusion

- ADaM is sufficiently developed to make a start on some standard datasets
- Further information on incidence data required but concepts can be applied with a view to modification later
- A subset of datasets can be produced initially with associated ‘standard’ outputs
- Cost and time savings will ultimately be recognised when applying these standard models.

Questions

- Has anyone else started work on implementing ADaM guidelines into a standard offering?
- How have people handled exposure / compliance ?
- Who agrees with the assertion that ADaM V1.0 Final is sufficient to act as a foundation for a standard approach to analysis data and associated outputs?
- Have people applied the permissible traceability variables as a standard?
- Has anyone tried to analyse a cross-over trial using ADaM V1.0 final – how was it?
- Does anyone know when a standardised ADaM approach to AEs will be available?