Adding In Vivo Mammalian Cytogenetics Data to SEND and ToxML

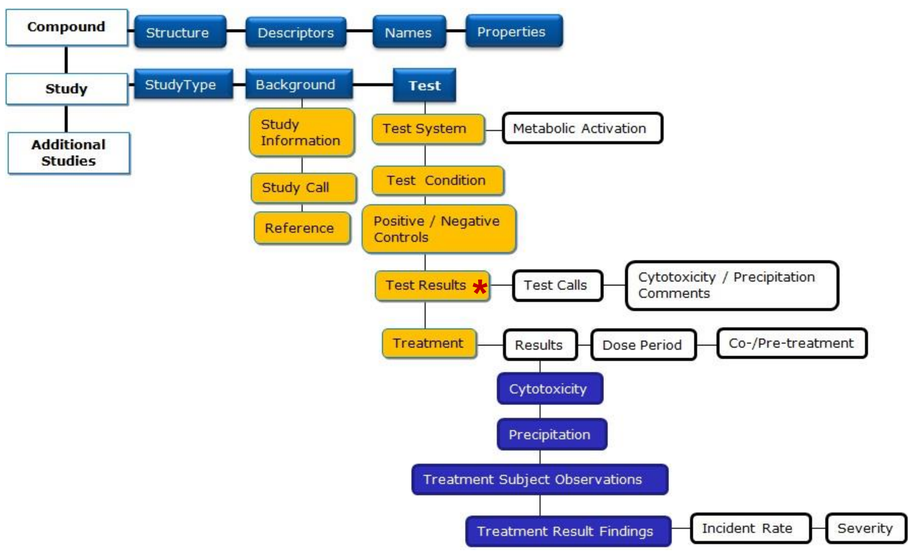
1. Introduction

Both SEND and ToxML are data standards designed to store and transport toxicological and related data in a consistent format. SEND covers study data that is related to animal testing as part of pre-clinical trials and captures detailed subject level observations. ToxML is an open standard based on the Extensible Markup Language (XML) which can cover both clinical and non-clinical data. The standards maintain their own controlled vocabularies that ensure consistency of usage when capturing experimental data. The standards are in the process of being expanded to capture additional endpoints as well as unique/non-standard experimental protocols. This collaboration illustrates the structure of SEND and ToxML and provides a real example of how data is captured and stored in the two formats.

2. Materials & Method

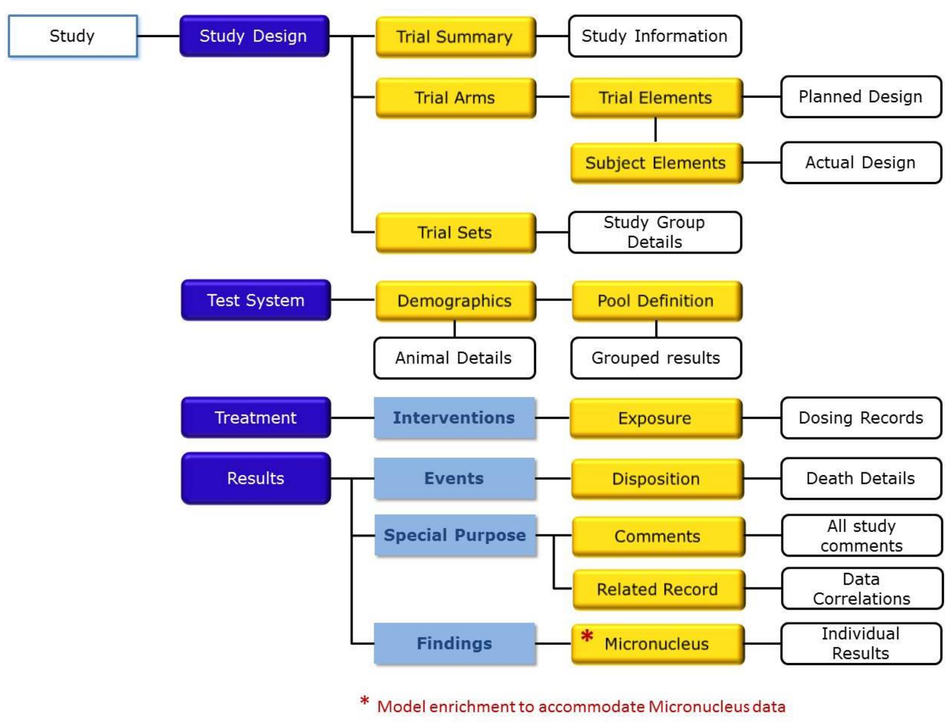
The collaboration focussed on capturing raw experimental data from a micronucleus assay experiment in the two formats. In each case data was captured but the schema had to be extended to capture the individual animal data. In the case of ToxML extra fields were created to support the entry of results for each subject ID. SEND is built up of a set of logically related domains to reflect the structures of data for standard endpoints in general toxicology studies. To accommodate the micronucleus assay data, a new domain had to be created to accurately represent the data. Based on existing concepts in SEND, the micronucleus assay domain was tailored using the "How to build a custom domain" resource from the PhUSE Nonclinical Roadmap team.

❑ **ToxML:** structure



- **Design is open standard**
- **Covers single and repeat dose genetic toxicity and other toxicity studies**
- **Data format based on extensible markup language (XML)**
- **Supports a hierarchical data structure**
- **Chemical substance serves as the fundamental base key or root node**
- **Supports both binary & non-binary data (numbers, text, images)**
- **Uses controlled normalized vocabulary to ensure harmonisation**
- **Allows for inclusion of group means, statistical analysis results and study conclusions**

❑ **SEND:** Standard for the Exchange of Non-clinical Data



- **SEND is a study data tabulation model developed by the CDISC SEND team**
- **It is a way of indexing/structuring nonclinical in vivo study data**
- **The study data is tabulated by subject (animal)**
- **Traceability is ensured by representing data both as collected and standardized**
- **Allows for representation of relationships between data records**
- **The planned and actual study conduct is described**

3. Results & Discussion

A published in vivo mammalian cytogenetics study satisfying the OPPTS 870.5395; OECD 474 guideline was used to model endpoints from an erythrocyte micronucleus assay in mice. The full reference can be found here:
www.epa.gov/opp00001/chem_search/cleared_reviews/csr_PC-072506_27-Jul-05_a.pdf

Treatment (24 hr harvest)	Dose (mg/kg)	Animal ID	# MN PCEs/2000 PCEs	Ratio PCE:NCE
Control Vehicle	Corn oil	2602	1	0.47
		2606	0	0.78
		2624	0	0.38
		2628	0	0.48
		2632	2	0.56
Positive Control	80	2600	21	0.63
		2601	39	0.49
		2603	32	0.54
		2610	14	0.81
		2613	17	0.53
Test Compound	31.25	2595	1	0.42
		2596	0	0.43
		2609	1	0.33
		2620	1	0.62
		2633	4	0.61
	62.5	2593	1	0.47
		2597	0	0.58
		2598	4	0.52
		2608	3	0.42
		2615	2	0.47
	125	2614	2	0.62
		2616	1	0.42
		2618	0	0.36
		2619	2	0.36
		2626	2	0.45

Table 1: Raw study results

This table shows the number of micronucleated PCEs/2000 PCEs per animal and treatment doses and the ratios of PCE:NCE for the male mouse micronucleus assay.

Reported data captured in SEND

STUDYID	DOMAIN	TSSEQ	TSGRPID	TSPARMCD	TSPARM	TSVAL
221	TS	1		STITLE	Study Title	Erythrocyte Micronucleus assay in Mice for Silver Chloride
221	TS	1		STCAT	Study Category	TOX
221	TS	1		SSTYP	Study Type	IN VIVO MAMMALIAN CYTOGENETICS
221	TS	1		SDESIGN	Study Design	PARALLEL
221	TS	1		ROUTE	Route of Administration	INTRAPERITONEAL
221	TS	1		DOSDUR	Dosing Duration	P1D
221	TS	1		SPECIES	Species	MOUSE
221	TS	1		STRAIN	Strain/Substrain	Crl:CD-1(ICR)BR
221	TS	1		AGETXT	Age Text	8
221	TS	1		AGEU	Age Unit	WEEKS
221	TS	1		SEXPOP	Sex of Participants	MALE
221	TS	1	1	TRT	Investigational Therapy or Treatment	Silver Chloride
221	TS	1	1	TRTCAS	Primary Treatment CAS Registry Number	7
221	TS	1		TRTV	Treatment Vehicle	C
221	TS	1		GLPTYP	Good Laboratory Practice Type	
221	TS	1		INTSAC	Time to Interim Sacrifice	
221	TS	1		TRMSAC	Time to Terminal Sacrifice	
221	TS	1	2	TSTFNAM	Test Facility Name	Example C
221	TS	1		SSPONSOR	Sponsoring Organization	Examp
221	TS	1	2	STDIR	Study Director	Dr.
221	TS	1		SNDIGVER	SEND Implementation Guide Version	SEND Implemen
221	TS	1		SNDCTVER	SEND Controlled Terminology Version	SEND Term

The Trial Summary dataset (to the left) is an extract showing what the trial summary would look like in SEND. Note: Not all recommended TS parameter codes are included.

STUDYID	DOMAIN	SETCD	SET	TXSEQ	TXPARMCD	TXPARM	TXVAL
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	1	ARMCD	Arm Code	1
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	2	SPGRPCD	Sponsor-Defined Group Code	1
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	3	GRPLBL	Group Label	Group 1, Control Vehicle
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	4	TCNTRL	Control Type	Vehicle Control
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	5	TRTDOS	Dose Level	0
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	6	TRTDOSU	Dose Units	mg/kg
221	TX	1	Control Vehicle (Corn oil), 24 hour harvest	7	SPLANSUB	Planned Number of Subjects	5
221	TX	2	Positive Control (Cyclophosphamide), 24 hour harvest	8	ARMCD	Arm Code	2
221	TX	2	Positive Control (Cyclophosphamide), 24 hour harvest	9	SPGRPCD	Sponsor-Defined Group Code	2
221	TX	2	Positive Control (Cyclophosphamide), 24 hour harvest	10	GRPLBL	Group Label	Group 2, Positive Control
221	TX	2	Positive Control (Cyclophosphamide), 24 hour harvest	11	TRTDOS	Dose Level	80
221	TX	2	Positive Control (Cyclophosphamide), 24 hour harvest	12	TRTDOSU	Dose Units	mg/kg
221	TX	2	Positive Control (Cyclophosphamide), 24 hour harvest	13	SPLANSUB	Planned Number of Subjects	5
221	TX	3	Test Compound (31.25 mg/kg), 24 hour harvest	14	ARMCD	Arm Code	3
221	TX	3	Test Compound (31.25 mg/kg), 24 hour harvest	15	SPGRPCD	Sponsor-Defined Group Code	3
221	TX	3	Test Compound (31.25 mg/kg), 24 hour harvest	16	GRPLBL	Group Label	Group 3, Test Compound low dose
221	TX	3	Test Compound (31.25 mg/kg), 24 hour harvest	17	TRTDOS	Dose Level	31.25
221	TX	3	Test Compound (31.25 mg/kg), 24 hour harvest	18	TRTDOSU	Dose Units	mg/kg
221	TX	3	Test Compound (31.25 mg/kg), 24 hour harvest	19	SPLANSUB	Planned Number of Subjects	5

The Trial Sets dataset (to the right) provides a high level overview of the treatment groups in the study.
Only the first three treatment groups are shown the table.

The dataset below shows data for animal # 2600, 2601, 2602 and 2603.
The micronucleus data were collected per 2000 PCEs per animal but also standardised and reported per 1000 PCEs per animal in the SEND file.
Actual date/time information would be likely to be included in data collected in today's data capture systems, but it was not available in this legacy data.

STUDYID	DOMAIN	USUBJID	MNSEQ	MNTESTCD	MNTEST	MNORRES	MNORRESU	MNSTRESC	MNSTRESN	MNSTRESU	MNSPEC	MNANTREG	MNBLFL	VISITDY	MNDTC	MNDY	MNTPT	MNTPTNUM	MNELTM	MNTPTREF	MNRFTDTC
221	MN	221-2600	1	MN_PCE	Micronucleated Polychromatic Erythrocyte	21	/2000 PCEs /animal	10.5	10.5	/1000 PCEs /animal	BONE MARROW			2		2	24 hours after dose	1	P24H	Day 1 dose	
221	MN	221-2600	2	PCE/NCE	Polychromatic Monochromatic Ratio	0.63	Ratio	0.63	0.63	Ratio	BONE MARROW			2		2	24 hours after dose	1	P24H	Day 1 dose	
221	MN	221-2601	1	MN_PCE	Micronucleated Polychromatic Erythrocyte	39	/2000 PCEs /animal	19.5	19.5	/1000 PCEs /animal	BONE MARROW			3		3	48 hours after dose	2	P48H	Day 1 dose	
221	MN	221-2601	2	PCE/NCE	Polychromatic Monochromatic Ratio	0.49	Ratio	0.49	0.49	Ratio	BONE MARROW			3		3	48 hours after dose	2	P48H	Day 1 dose	
221	MN	221-2602	1	MN_PCE	Micronucleated Polychromatic Erythrocyte	1	/2000 PCEs /animal	0.5	0.5	/1000 PCEs /animal	BONE MARROW			2		2	24 hours after dose	1	P24H	Day 1 dose	
221	MN	221-2602	2	PCE/NCE	Polychromatic Monochromatic Ratio	0.47	Ratio	0.47	0.47	Ratio	BONE MARROW			2		2	24 hours after dose	1	P24H	Day 1 dose	
221	MN	221-2603	1	MN_PCE	Micronucleated Polychromatic Erythrocyte	32	/2000 PCEs /animal	16	16	/1000 PCEs /animal	BONE MARROW			2		2	24 hours after dose	1	P24H	Day 1 dose	
221	MN	221-2603	2	PCE/NCE	Polychromatic Monochromatic Ratio	0.54	Ratio	0.54	0.54	Ratio	BONE MARROW			2		2	24 hours after dose	1	P24H	Day 1 dose	

Reported data captured in ToxML

```
<Tests>
 <Test>
  <TestSystem>
   <Species>Mouse</Species>
   <Strain>Crl:CD-1 (ICR) BR</Strain>
   <TargetCell>Bone marrow</TargetCell>
   <Sex>Male</Sex>
   <Supplier>Charels River Laboratories, Raleigh, NC</Supplier>
   <InitialAge>8 Weeks</InitialAge>
   <InitialWeight>29.9-36.3 g</InitialWeight>
  </TestSystem>
  <Treatments>
   <TreatmentGroup>
    <DosageRegimen>
     <Dose>
      <Value>31.25</Value>
      <Units>mg/kg</Units>
     </Dose>
     <HarvestTime>
      <Value>24.0</Value>
      <Units>hour</Units>
     </HarvestTime>
    </DosageRegimen>
    <IndividualSubjectCount>6</IndividualSubjectCount>
*   <Results>
     <PCENCERatio>0.48</PCENCERatio>
     <MNPCEPer1000PCE>0.035</MNPCEPer1000PCE>
     <MNPCESignificance>Not significant</MNPCESignificance>
    </Results>
   </TreatmentGroup>
  </Treatments>
 </Test>
</Tests>
```

```
<Results>
 <PCENCERatio>0.48</PCENCERatio>
 <MNPCEPer1000PCE>0.035</MNPCEPer1000PCE>
 <MNPCESignificance>Not significant</MNPCESignificance>
 <IndividualResults>
  <IndividualResult>
   <SubjectId>2595</SubjectId>
   <MNPCE>
    <MNPCECount>1</MNPCECount>
    <TotalPCECount>2000</TotalPCECount>
   </MNPCE>
   <PercentPCE>
    <PCENCERatio>.42</PCENCERatio>
   </PercentPCE>
  </IndividualResult>
  <IndividualResult>
   <SubjectId>2596</SubjectId>
   <MNPCE>
    <MNPCECount>0</MNPCECount>
    <TotalPCECount>2000</TotalPCECount>
   </MNPCE>
   <PercentPCE>
    <PCENCERatio>.43</PCENCERatio>
   </PercentPCE>
  </IndividualResult>
 </Results>
```

Treatment Group Summary Result Data

This version of the ToxML specification is already available as the 'working draft' from https://toxml.org/working-draft/

Extended Individual Result Level Data

* ToxML was enriched to incorporate the capture and association of individual subject identifiers with their respective result findings. The web-based Specification Editor created by the ToxML Organisation was used to add the new field (SubjectId) to the working draft version of the specification. The SubjectId field is depicted in the figure above.

4. Conclusion

The collaboration focused on capturing raw experimental data from a micronucleus assay experiment. In each case the existing models accommodated many of the necessary concepts and the schema was successfully extended to allow capture of all the relevant data. While SEND required enrichment of the data structure, ToxML required definition of subject level data and the addition of extra fields for subject level observations.

We found overlap between the structures of the models as well as areas where each model could be enriched by taking concepts from the other.

Acknowledgements: MA Ali, Lhasa Limited, UK; PN Judson, Judson Consulting Service, UK; W Houser, Bristol-Myers Squibb, USA; G Frausing, Data Standards Decisions Aps, Denmark; D Bower and K Cross, Leadscope Inc, USA. **Note:** The opinions expressed in this poster are those of the authors and do not necessarily represent the opinions of their respective companies or organisations.