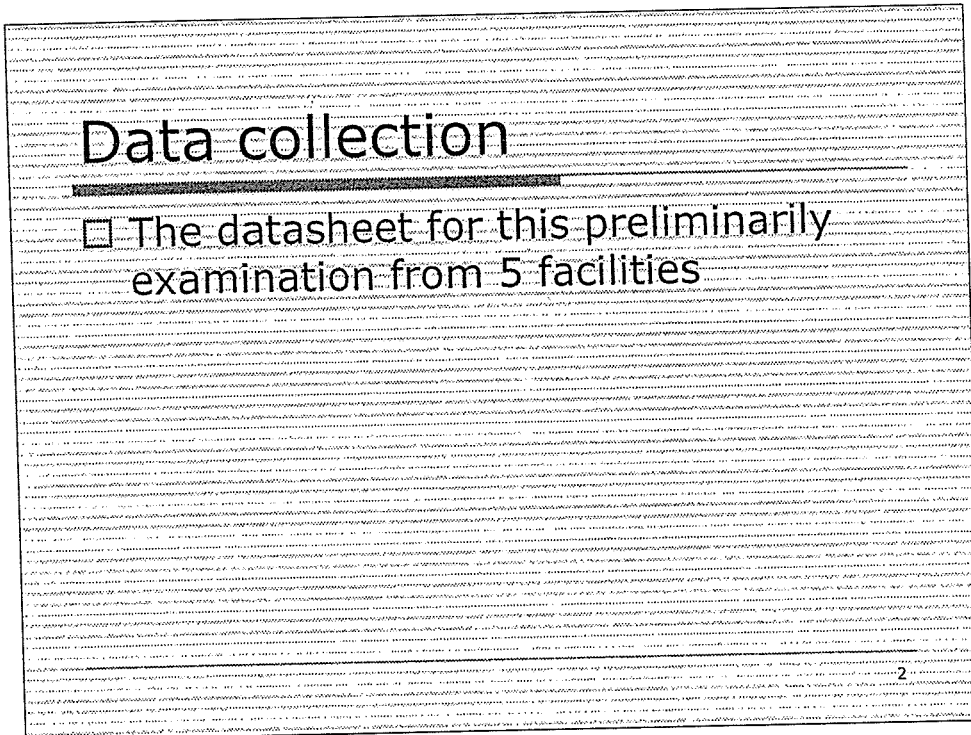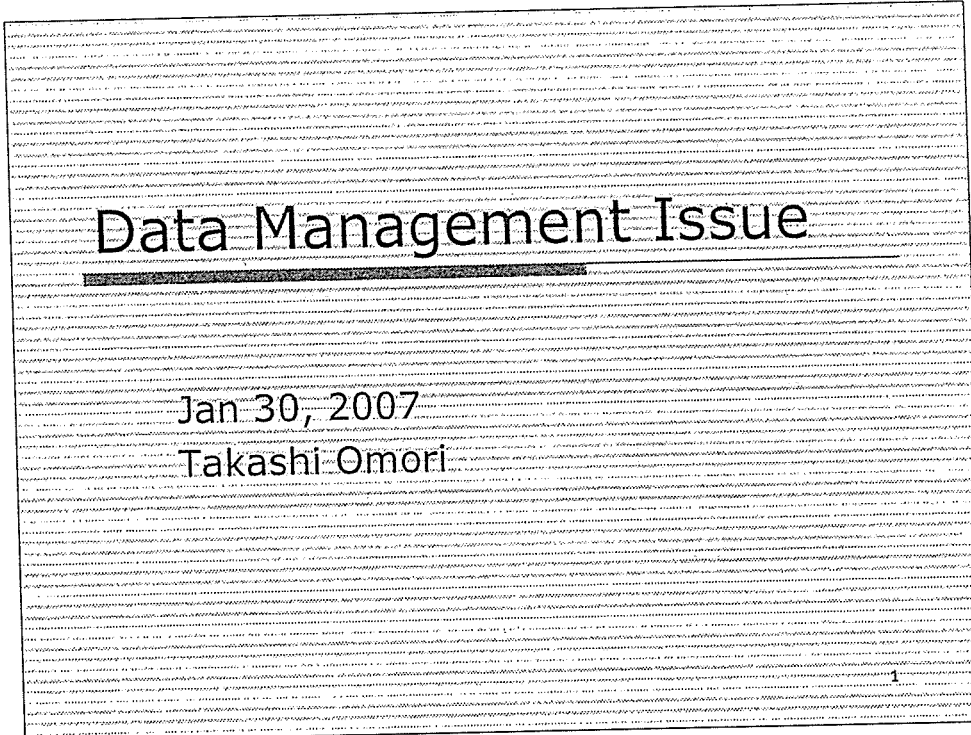|	not randomized	Animals are randomized based on body weights and a computer program that calculates the means and Standard deviation. Animals outside 1.5 SD from the mean are excluded from the study.	Manual Randomisation	Animals are randomly assigned by body weight on the last day of quarantine period to experimental groups using the computer-generated list of permutations.	Yes
Assignment	not randomized	Animals are randomized based on body weights and a computer program that calculates the means and Standard deviation. Animals outside 1.5 SD from the mean are excluded from the study.	Manual Randomisation	Animals are randomly assigned by body weight on the last day of quarantine period to experimental groups using the computer-generated list of permutations.	Yes
Assay solution					
Agarose: bottom	as per protocol	Sigma Lot.	As protocol	Same as protocol	1%agarose
Agarose: cell-containing	as per protocol	Cambrex Lot No. AG5897		Add 10 mM EDTA	0.5% low-melting agarose
Lysing solution	as per protocol	2.5 M NaCl , 100 mM EDTA and 10 mM Trizma	As protocol	Same as protocol	same protocol
Alkaline solution	as per protocol	Electrophoresis buffer, 300 mM NaOH/ 1 mM EDTA, pH 13	As protocol	Same as protocol	same protocol
Neutralization solution	as per protocol	Invitrogen ultrapure tris-HCL solution, pH 7.5	As protocol	Same as protocol	same protocol
Mincing buffer	as per protocol	Hank's balance salt solution, calcium, magnesium and phenol free (HBSS) containing 20mM EDTA and 10% DMSO		Hanks(–) solution (Invitrogen) containing 10 v/v% DMSO (Wako)	same protocol
Staining solution	SybrGold, as per protocol	1X Sybr Gold	x1 Sybr Gold	5xSYBR GOLD (Invitrogen)	×10000 SYBR Gold
All regents used		see workbook			
Identification	see workbook	Detailed in study notes			
Supplier	see workbook	Detailed in study notes			
Lot	see workbook	Detailed in study notes			

		33, 100, 200 mg/kg	100 and 200 mg/kg		Also conducted single dose	Also conducted single dose
Dosing procedure	Dose					
	Dosing frequency	twice	Double dose	21h interval	one (3h and 24h treatment) or two	Once and twice
	Dosing interval	24 hours	21h interval	21h Interval	21h interval (2 times)	21h interval
	Route	oral	PO	P.O.	P.O.	Oral
	Dosing volume	10 mL/kg	10 mL/kg	10 mL/kg	10 mL/kg	0.1 mL per 10 g body weight
Animal observation	Body weight	Just before each dose administration and at sacrifice	weighed on the days of dosing and the administered volume will be based on individual body weight	weighed on the days of dosing and the administered volume will be based on individual body weight	Just before administration (3h single), Just before each administration (3+24h tow times), Just before administration and sacrifice (24h single)	
	Clinical sign	none	Animals at 200 mg/kg showed lethargy and piloerection	All normal	All normal	Normal
Tissue sampling	Target	Liver, Glandular stomach	Liver and glandular stomach	Liver and glandular stomach	Liver and glandular stomach	Liver and Stomach
	Sampling time	3h after 2nd dosing	3h after 2nd dosing	3h after 2nd dosing	3 and 24h after single dose; 3h after 2nd dosing	3 and 24h after single dose; 3h after 2nd dosing
	(Exsanguination step)	Yes	No	No	Yes	Yes
Cell preparation	Liver prep. proc.	mince in buffer	Mincing (a couple of seconds)	Mincing	100 times mincing	same protocol
	Stomach prep. proc.	incubate on ice in buffer, scrape with scalpel, triturate 2-3 times	Srapping (a couple of seconds) and filtering	Scraping (5 strokes)	Scrape and 100 times mincing	same protocol
	Use of cell strainer	No	Yes	Yes	Yes	No
Slide preparation	Percentage of agarose	0.50%	0.50%	0.50%	0.47%	0.5%
	No. of slides/organ	3	4	Same as protocol	Same as protocol	5 slides/organ
Lyses	Duration	overnight	in lysis at least overnight	Overnight	O/N	about 70 h
Unwinding	Duration	20 min	20 min	20 min	20 min	20 min

Category	Parameter					
	pH	>13	>13	13	<13	>13
	Temperature	ambient ~20C	Room temperature (~21YC) and 2-8YC	4-5C 20-20'C	1.3 - 2.5C (end of unwinding)	1-3C
Electrophoresis	(Re-circulating type)	no	no	No	No	No
	Balanced design	yes	yes	Yes	Yes	Yes
	Voltage	24V (0.7 V/cm)	22 (0.7 V/cm)	22 (0.7 V/cm)	36V (1V/cm) constant voltage	36 V
	Current	300 mA	< 300 mA	290 - 340mA	Adjust to 0.30A at beginning	0.3 A
	Duration	20 min	30 min	20 min	Liver: 40 min., Stomach: 15 min.	15 min
	pH	>13	>13		Not observed	>13
	Temperature	20C	2-8C and 20C	4C and 20C; lost gel problem at 20C	2.5 - 5.1C (end of electrophoresion)	2-5C
Neutralization	Duration	~5 minutes	at least 10 min	10 mins	>20 min	10 min
Slide dehydration	Duration	~5 minutes absolute ethanol	at least 10 min	Not conducted	>10 min	10 min
	Strage condition	dark, ambient temp and humidity	at room temperature	-	R.T.	room temp, closed container
Analysis	Staining procedure	1X SYBR Gold working stock	Stain cells with 1X Sybr Gold and wait for 20 min	Stain cells with 1X Sybr Gold and wait for 15 min	Use 5-fold higher conc. of SYBR Gold for 1 min. or more	
	Use of image analyzer	Comet IV	Comet IV	Comet IV	Comet IV	Komet IV
	Magnification	10X objective	20X	20x	200	x400
	Scoring No. of slides	3 slides	2 slides	2 slides	2 slides	2 slides
	Scoring No. of cells/slide	50	50	50	50	50 cells/slide
	Parameter				-	
	% tail DNA	yes	yes	Yes	Yes	Yes

Tail length	Tail length	(have data)	Tail length ant tail migration is different in the perceptive intruments. I used tail migrations, that goes form the peripheri of the head to the last visible point of the tail.	Yes	Yes	Yes
	Tail moment	(have data)	yes	Yes	Yes	Yes
	Categorical analysis	no	no	No	Not yet	No
Neutral diffusion	Conducted	yes	Yes	Yes	Not yet	Yes
	Duration	1 hour	1 hour in lysis	1 hour in lysis		Yes
	Slide dehydration proc.	~ 5 minutes absolute ethanol	at least 10 min	No		Yes
	Analysis magnification	10X objective, approx. 100X to camera	20X	20x		No
(Histopathology)	Liver	On going	On going	No	No	No
	Stomach	On going	On going	No	No	No
Statistics		f-test of log-transformed animal data (preliminary)	Not yet, bu usually I do a lineal regression analysis, then a ANOVA with dunnets	Not yet	Not yet, maybe Dunnet	Yes(Dunnet)

Data Management Issue

Jan 30, 2007
Takashi Omori

1

Data collection

☐ The datasheet for this preliminarily examination from 5 facilities

2

Construction of DB

☐ Only one datasheet is selected in case of the several datasheets submitted in a facility.

☐ Names of each facility were coded.

☐ A SAS program for reading the imputed values in the datasheet was developed.

3

Construction of DB (Cont.)

☐ Database (DB) include the observed values rather than summary statistics.

☐ One missing observed value is one observation.

☐ Examinations for data analysis based upon the DB.

4

2

Size of DB for this exam.

☐ Facility x Parameter x Region x Dose
  x Animal x Slide x # par slide

  ■ = 5 x 3 x 2 x 3 x 5 x 3 x 50
  ■ = 67500 observed values

☐ 8 Variables

☐ DB is a matrix of 67500 x 8

5

Variables

☐ List of variables

Variable name	Description	Numeric or Character	Length	Note
SHEETNUM	row numbers of observed data entry feiled in the datasheet	C	8	
VALUE	Observed value for the parameter	N	8	
CATCONC	1:Vehicle, 2: Low dose, 3: High dose	N	8	
ANIMALID	Animal id in the dataset	N	8	1-15
SLIDENUM	Slide number	N	8	1-3
LS	1: Liver, 2: Stomach	N	8	
PARAS	1:% tail DNA, 2: Tail length, 3: Olive tail moment	N	8	
LAB_ID	Assained ID for each facility	N	8	

6

Current problems

- ☐ Short of stuffs to deal with DB.
- ☐ We need a few weeks to construct DB.
  - ◼ If we construct more large size DB, we need more weeks.
- ☐ From April in 2007, could JaCVAM fund hire a new stuff for data management?

7

Before examinations
~Endpoint, Estimate and Effect~

Jan 30, 2007

Takashi Omori

Purpose of this presentation

- When considering data analysis of this validation study that we conduct, I found we need to distinct the level of data.

- I would like to propose three different terms for data to make us easier to understand what we do.

2

1

Hierarchical structures for data

- Groups (Concentrations)
- Animals ← unit of randomization
- Slides
- Observed values ← unit of observed data

Note: Ignoring the unit of randomization would lead to serious misunderstanding for variations of data

Three terms

- Endpoint
  - Observed values for a parameter

- Estimate
  - Mean or median for the particular endpoints by individual animal

- Effect
  - Difference or ratio of the averaged estimate between groups

Endpoint

- Endpoint is a particular parameter or observed values of a particular parameter in the assay.
- For example, % tail DNA is one of endpoints.

5

Estimate

- Estimate is a summary statistics for the location on the distribution of a parameter, such as mean or median of an endpoint for each animal.

6

Estimate (Cont.)

- Since the unit of randomization is individual animal rather than the values of an endpoint, the unit should be considered in the data analysis and in the presentation of data.

7

Estimate (Cont.)

- Though the estimate should be used as a basic unit for data analysis when we apply the statistical test, the use of the sophisticated statistical methods, such as the Generalized Estimating Equation (GEE), the mixed effect model or the generalized linear mixed model, can allow the use of the values of an endpoint directly.

8

Effect

- Effect is a measures to evaluate the response on the used a dose group compared to it on the vehicle group in a experiment. Usually, for the aim of it, difference or ratio is used.
- For example, the difference between the average for medians of % tail DNA in the low dose group and it in the vehicle group is one of effects.

9

Effect (Cont.)

- Since the confidence intervals of the effect correspond to statistical test, we can judge whether a tested chemical is positive or negative by using it.
- Different from the use of p-value, the use of effect with its confidence interval can refer to size of amount for difference or ratio directly.

10

5

Effect (Cont.)

- Effect would be one of important measures in the validation study because the calculated values of the effect has more information than the dichotomous judgments of positive or negative.
- In order to evaluate the inter-laboratory variation, for example, we can compare with effects from laboratories.

11

6

Data analysis of the pre-validation study with EMS in five leading laboratories

Date: February 15, 2007

Prepared by: Yoshifumi Uno

Introduction

Dr. Omori analyzed data of the pre-validation study with EMS in five leading laboratories, and he explained the results with attached four PP presentation files in a Japanese domestic Comet meeting on January 30. This document is provided for oversea colleagues to help your understanding of the attached PP files made by Dr. Omori, i.e. this document includes additional explanation of some slides that seem to have some difficulty to understand.

Japanese members can accept Dr. Omori's proposals to investigate the assay validity in our Comet validation studies, and his proposals are highlighted with underlines in the following sentences. If you have any comments or suggestions for Dr. Omori's proposals, please let us know.

PP file #1: Data Management Issue

1) Page 3: Analysis was conducted with data from two times administration of EMS with 21- or 24-hour interval followed by one sampling time at 3 hours (2-4 hours) after the final administration.

PP file #2: Before examinations – Endpoint, Estimate and Effect –

1) Page 3: Data have four hierarchical structures. We should strictly recognize the difference between "Animals are a unit of randomization" and "Observed values are a unit of observed data".

2) Page 4-11: In analysis process, Dr. Omori defines and uses three conceptual key terms described in these slides. Briefly, "Endpoint" is defined as individual observed values for a parameter such as % DNA in tail. "Estimate" is defined as a mean or median calculated with values of a particular "Endpoint" in each animal. "Effect" is defined as difference (or ratio) of an average of "Estimate" between a negative control group and a treatment group. The purpose of this (and further) validation study analysis is to investigate how large variation exists among data from testing facilities, and <u>Dr. Omori proposes "Effect" would be a preferable yardstick (criterion) to understand the variation of Comet parameters among</u>

testing facilities. In this analysis, he summarizes "Endpoint", "Estimate", and "Effect" in the PP file #3.

3) Cf. another idea for analysis would be to compare the variation of assay results, i.e. positive or negative, among testing facilities. In this analysis, an expected assay results for each test compound should be clarified (decided?) in advance.

PP file #3: Data on the protocol ver. 10

1) Pages 2, 9, 16, 20, 24, 28, 32, 36, 40, and 44: These slides show how to see the following six or three slide figures. For example, in page 2, the slide title means the following (six) slides show "Distribution of observed values by individual animals", the term "PARAS 1" means "% DNA in tail", the term "Level 1: Region" means data were obtained from the "Liver" and the "Stomach", the term "Level 2: Facility code" means five leading laboratory names were coded and simply described as "1", "2", "3", "4" and "5", and the term "Level 3: Animal" means the animal number used ("1"-"5": vehicle control, "6"-"10": low dose of EMS, and "11"-"15": high dose of EMS).

2) Page 3: This type of slides shows distributions of "Endpoint". This slide is for "PARAS 1", i.e. "% DNA in tail", in the "Liver". Each row shows the coded facility name (from lower to upper, "1", "2", "3", "4" and "5") and each column shows the animal number (from left to right, columns 1-5: vehicle control, columns 6-10: low dose of EMS, and columns 11-15: high dose of EMS).

3) Page 4: This slide shows "PARAS 1", i.e. "% DNA in tail", in the "Stomach". Here, I guess you must understand how to see the same type of slide figure, and I skip the explanation how to see the following slides.

4) Page 17: This type of slides shows "Estimate". This slide is for "PARAS 1", i.e. "% DNA in tail". Upper five figures show the "Stomach", and lower five show the "Liver". Each column shows the coded facility name (from left to right, "1", "2", "3", "4" and "5"). In each figure, a open circle shows a mean of observed values, and a group of open circles shows each treatment group, i.e. from left to right, vehicle control, low dose of EMS, and high dose of EMS. Here, I guess you must understand how to see the same type of slide figure, and I skip the explanation how to see the following slides.

5) Page 33: This type of slides show "Effect". Each open circle shows a value of difference between the mean of vehicle control and the mean of treatment group: the left circle shows the low dose of EMS and the right circle shows the high dose of EMS.

214

6) Page 48-50: Dunnett's one side test is applied to data (also see another data analysis sheets entitled "#5 DunnettU table 1"). Here, I explain briefly how to see the table in page 50. "L-V" means "the difference between low dose and vehicle control", and "H-V" means "the difference between high dose and vehicle control", i.e. these means "Effect". "Lower" means a value of the lower limit for the confidence interval obtained from Dunnett's one side test, and the value is considered significant when it is zero or more. "Diff." means a value of "Effect", and <u>Dr. Omori proposes comparison of these values would be a preferable yardstick (criterion) to understand the variation of Comet parameters among testing facilities</u>. "Sig." means the statistical significance between the vehicle control group and the treatment group.
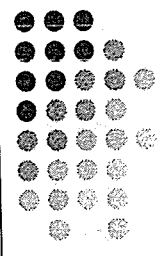
PDF file #4: Slide-to-slide variation

1) Page 3: This slide shows boxplots for "PARAS 1", i.e. "% DNA in tail", in the "Liver". Each row shows the coded facility name (from lower to upper, "1", "2", "3", "4" and "5") and each column shows the animal number (from left to right, columns 1-5: vehicle control, columns 6-10: low dose of EMS, and columns 11-15: high dose of EMS).

2) Page 4: This slide shows "PARAS 1", i.e. "% DNA in tail", in the "Stomach". Here, I guess you must understand how to see the same type of slide figure, and I skip the explanation how to see the following slides.
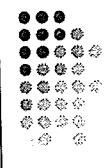
3) Slide-to-slide variation seems negligible.

The end of this document

Data on the protocol ver.10

Jan 30, 2007

Takashi Omori

Distributions of observed values by individual animals

- PARAS
  - 1: % DNA in tail
  - 2: tail length
  - 3: Olive tail moment

- Level 1: Region
  - Liver
  - Stomach

- Level 2: Facility code
  - 1-5

- Level 3: Animal
  - 1-15

2

1

PARAS: 1

Percent of Total

VALUE

3

PARAS: 1

Percent of Total

VALUE

4

PARAS: 2

Percent of Total

VALUE

5

PARAS: 2

Percent of Total

VALUE

6

3

218