



Summary Level Clinical Site Data for Data Integrity Review and Inspection Planning in NDA and BLA Submission

Jingwei Gao, Boehringer Ingelheim (China) Investment Co. Ltd
Nancy Bauer, Boehringer Ingelheim (US)

Objectives

- History
- Draft Guidance
- Summary Level Clinical Site Data
- Data Definition File
- What Else Needs to Be Submitted?
- Site Selection
- Future Developments
- Questions and Answers

Brief History

- 2009 – First initiated by FDA Commissioner Margaret Hamburg
- Sept 2009 – FDA Site Selection tool ready for sponsor pilots
- 2011/2012 – Original version of the FDA draft guidance and specifications released
- The Pharmaceutical Users Software Exchange (PhUSE)/FDA/Clinical Data Interchange Standards Consortium (CDISC) worked together to update the specifications to be CDISC compliant. This version has not been released yet.

FDA Draft Guidance and Specifications –Two Parts

12/2012 Electronic submissions

Guidance for Industry Providing Submissions in Electronic Format — Summary Level Clinical Site Data for CDER's Inspection Planning

Specification Document:

www.fda.gov/downloads/drugs/developmentapprovalprocess/formssubmissionrequirements/ucm332466.pdf

*Specifications for Preparing and Submitting Summary Level Clinical Site Data for
CDER's Inspection Planning*

Revision History

Date	Version	Summary of Changes
11/08/11	1.0	Original Version
10/15/12	1.1	Updated example values in Appendix 2
11/7/12	1.2	Updated Country codes from alpha-2 to alpha-3

Draft Guidance

Applies to:

- **New Drug Applications (NDAs)**
- **Biologics Licensing Applications (BLAs)**
- **NDA and BLA supplemental applications containing new clinical study reports submitted to CDER**

Does not apply to

- **Biopharmaceuticals**
- **Clinical pharmacology**
- **Animal studies**

Draft Guidance

- **THREE parts requested**
 - **Part 1 Site–Level Data Set**
Site level summary data for site selection tool across all pivotal studies (.xpt)
 - **Part 2 Tabular Listings of Site Information**
 - Tabular information for all sites in the study (PDF)
 - Protocols and Annotated CRFs for each study (PDF)
 - **Part 3 Line Listings by Site**
Subject data listings organized by site for each study (PDF)

Draft Guidance

- Assist applicants to include the summary level clinical site data in the submission
- Verify the integrity of data submitted
- Address regulatory issues earlier
- Assure the rights and welfare of human research subjects protected
- Facilitate a risk-based approach to site selection and inspections
- Plan inspections efficiently and effectively
- To meet PDUFA timelines
- Ensure the compliance of US laws and regulations on GCP and GLP

Draft Guidance

FDA purpose for guidance: ‘is intended to facilitate use of a risk-based approach for the timely identification of clinical investigator sites for on-site inspection by CDER during the review of marketing applications.’

Summary Level Clinical Site Data

- eCTD format
- Module 5
- Leaf titles should be “bimo”
- The file name “clinsite.xpt”



- Submit electronically via
FDA Electronic Submission Gateway (ESG) or using
appropriate physical media

Summary Level Clinical Site Data

- Data set contains data from all major (pivotal) studies used to support safety and efficacy
- Data submitted by clinical site and treatment arm for the intent-to-treat (ITT) population
- Clearly identify individual clinical investigator sites within an application or supplement
- Present characteristic and outcomes of the study at the site level

Summary Level Clinical Site Data

Site Administrative Details

- Study number
- Study title
- Number of sponsors
- Other sponsor names
- Investigational new drug (IND) number
- Under IND (y/n)
- NDA number
- BLA number
- Serial number for supplemental application
- Site identification number
- Principal investigator information
 - Name, address, phone and fax number, email address
 - Start and stop dates and duration of responsibility

Summary Level Clinical Site Data

Site Financial Disclosure Information:

- Maximum financial disclosure amount (\$USD) by any single investigator by site
- Total financial disclosure amount (\$USD) by site calculated as the sum of disclosures for the principle investigator and all sub-investigators to include all required parties

Summary Level Clinical Site Data

Site Characteristics by Counts

- Randomize (ENROLL)
- Screened (SCREEN)
- Discontinue from study (DISCONT)
- Discontinue from treatment (DISCONT2)
- IPV (PROTVIOL)

Summary Level Clinical Site Data

Site-Specific Safety Results by Counts

- Non-Serious Adverse Events (NSAE)
- Serious Adverse Event (SAE)
- Deaths (DEATH)

Summary Level Clinical Site Data

Site-Specific Efficacy Results

- Treatment Arm (ARM)
- Name of the Primary Endpoints (ENDPOINT)
- Type of Endpoint (ENDPTYPE)
- Treatment Efficacy Result (TRTEFFR)
- Treatment Efficacy Result Standard Deviation (TRTEFFV)
- Site-specific Efficacy Effect Size (SITEEFFE)
- Site-specific Efficacy Effect Site Standard Deviation (SITEEFFV)
- The Number of Censored Observations(CENSOR)

Summary Level Clinical Site Data

Data Structure Example

STUDY	STUDYTL	DOMAIN	SPONNO	SPONNAME	IND	UNDERIND	NDA	BLA	SUPPNUM	SITEID	ARM	ENROLL	SCREEN	DISCONT
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	001	Active	26	61	3
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	001	Placebo	25	61	4
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	002	Active	23	54	2
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	002	Placebo	25	54	4
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	003	Active	27	62	3
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	003	Placebo	26	62	5
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	004	Active	26	60	2
ABC-123	Double blind...	DE	1	DrugCo, Inc.	000001	Y	200001	-1	0	004	Placebo	27	60	1

ENDPOINT	ENDTYPE	TRTEFFE	TRTEFFS	SITEEFFE	SITEEFFS	CENSOR	NSAE	SAE	DEATH	PROTVIOL	FINLMAX	FINLISC	LASTNAME	FRSTNAME
Percent Responders	Binary	0.48	0.0980	0.34	0.1405	-1	0	2	0	1	-1	-1	Doe	John
Percent Responders	Binary	0.14	0.0694	0.34	0.1405	-1	2	2	0	1	-1	-1	Doe	John
Percent Responders	Binary	0.48	0.1042	0.33	0.1427	-1	3	2	1	0	45000.00	45000.00	Washington	George
Percent Responders	Binary	0.14	0.0694	0.33	0.1427	-1	0	2	0	3	20000.00	45000.00	Washington	George
Percent Responders	Binary	0.54	0.0959	0.35	0.1448	-1	2	2	0	1	15000.00	25000.00	Jefferson	Thomas
Percent Responders	Binary	0.19	0.0769	0.35	0.1448	-1	3	6	0	0	22000.00	25000.00	Jefferson	Thomas
Percent Responders	Binary	0.46	0.0977	0.34	0.1275	-1	4	1	0	0	0.00	0.00	Lincoln	Abraham
Percent Responders	Binary	0.12	0.0625	0.34	0.1275	-1	1	2	0	1	0.00	0.00	Lincoln	Abraham

MINITIAL	PHONE	FAX	EMAIL	COUNTRY	STATE	CITY	POSTAL	STREET
M	555-123-4567	555-123-4560	John@mail.com	RU	Moscow	Moscow	103009	Kremlin Road 1
M	555-123-4567	555-123-4560	John@mail.com	RU	Moscow	Moscow	103009	Kremlin Road 1
	020-3456-7891	020-3456-7890	george@mail.com	GB	Westminster	London	SW1A 2	10 Downing St
	020-3456-7891	020-3456-7890	george@mail.com	GB	Westminster	London	SW1A 2	10 Downing St
	01-89-12-34-56	01-89-12-34-51	tom@mail.com	FR	N/A	Paris	75002	1, Rue Road
	01-89-12-34-56	01-89-12-34-51	tom@mail.com	FR	N/A	Paris	75002	1, Rue Road
	555-987-6543	555-987-6540	abe@mail.com	US	Maryland	Rockville	20852	1 Rockville Pk.
	555-987-6543	555-987-6540	abe@mail.com	US	Maryland	Rockville	20852	1 Rockville Pk.

Note: Information not applicable or not obtainable should be identified with a -1 as the value

Data Definition File (pdf)

Summary Level Clinical Site Data (CLNSITE)					clnsite.xpt
Variable	Label	Type	Display Format	Controlled Terms	Source / Derivation
STUDY	Study Number	text	\$9		From CTMS file.
STUDYTL	Study Title	text	\$200		From study protocol
DOMAIN	Domain Abbreviation	text	\$2		Set to DE for all records
SPONNO	Sponsor Number	integer	8		From CTMS file.
SPONNAME	Sponsor Name	text	\$20		Set to Boehringer Ingelheim
IND	IND Number	integer	6		From CTMS file.
UNDERIND	Under IND	text	\$1		From CTMS file.
NDA	NDA Number	integer	6	NAF	Set to -1 for all records
BLA	BLA Number	integer	6	NAF	Set to -1 for all records
SUPPNUM	Supplement Number	integer	BEST	NAF	Set to -1 for all records
SITEID	Site ID	text	\$5		From CTMS file.

SAE	Number of Serious Adverse Events	integer	\$		Count the number of patients by site and treatment group from POB.AE where AEONDT is within treatment start and stop and AESSERA equal to 1 and ABOUT not equal to 4
DEATH	Number of Deaths	integer	\$		Count the number of patients by site and treatment group from POB.AE where AEONDT is within treatment start and stop and ABOUT equal to 4
PROTOCOL	Number of Protocol Violations	integer	\$		Count the number of patients from ADSP.IPV. Group by study and site for cross-over studies, and by study, site, and arm for parallel studies.
FDNLMAX	Maximum Financial Disclosure Amount	float	BEST		From financial disclosure file. See Data Guide [Data Guide]
FDNLDISC	Financial Disclosure Amount	float	BEST		From financial disclosure file. See Data Guide [Data Guide]
LASTNAME	Investigator Last Name	text	\$18		From CTMS file.
FIRSTNAME	Investigator First Name	text	\$23		From CTMS file.
MENTIAL	Investigator Middle Initial	text	\$2		From CTMS file.
PHONE	Investigator Phone Number	text	\$25		From CTMS file.
FAX	Investigator Fax Number	text	\$25		From CTMS file.
EMAIL	Investigator Email Address	text	\$53		From CTMS file.

ARM	Treatment Arm	text	\$9		Planned treatment arm from each endpoint analysis dataset.
ENROLL	Number of Subjects Enrolled	integer	\$		Count the number of patients by site and treatment group from POB.ENRD where RANDSEL equal to 1.
SCREEN	Number of Subjects Screened	integer	\$		Count the number of patients by site in POB.FATD where CONDT is not missing.
DISCONT	Number of Subject Discontinuations	integer	\$		Count the number of patients by site and treatment group from POB.TTMD where TEROICD greater than 1.
ENDPOINT	Endpoint	text	\$18		PARAM of each endpoint in the analysis dataset. See Data Guide [Data Guide]
ENDPTYPE	Endpoint Type	text	\$10		Set to Continuous for all records
TRTEFFE	Treatment Efficacy Endpoint	float	BEST		Unadjusted treatment means for each endpoint. See Data Guide [Data Guide]
TRTEFFS	Treatment Efficacy Endpoint Std Dev	float	BEST		Standard deviations of the unadjusted treatment means for each endpoint. See Data Guide [Data Guide]
SITEEFFE	Site-Specific Treatment Effect	float	BEST		Unadjusted treatment means differences for each endpoint. See Data Guide [Data Guide]
SITEEFFS	Site-Specific Treatment Effect Std Dev	float	BEST		Standard deviations of the unadjusted treatment means differences for each endpoint. See Data Guide [Data Guide]
CENSOR	Censored Observations	integer	\$	NAF	Set to -1 for all records
NSAE	Number of Non-Serious Adverse Events	integer	\$		Count the number of patients by site and treatment group from POB.AE where AEONDT

What else needs to be submitted

- Reviewer's guide (as needed)
- Table of number of subjects screen, randomized, treated and prematurely discontinued by site (This is general information that may be requested – but not part of the Site level listings).

What else needs to be submitted

➤ Source data listings by site

II. Request for Subject Level Data Listings by Site

1. For each pivotal trial: Site-specific individual subject data (“line”) listings. For each site provide line listings for:
 - a. Listing for each subject/number screened and reason for subjects who did not meet eligibility requirements
 - b. Subject listing for treatment assignment (randomization)
 - c. Subject listing of drop-outs and subjects that discontinued with date and reason
 - d. Evaluable subjects/ non-evaluable subjects and reason not evaluable
 - e. By subject listing of eligibility determination (i.e., inclusion and exclusion criteria)
 - f. By subject listing, of AEs, SAEs, deaths and dates
 - g. By subject listing of protocol violations and/or deviations reported in the NDA, description of the deviation/violation
 - h. By subject listing of the primary and secondary endpoint efficacy parameters or events. For derived or calculated endpoints, provide the raw data listings used to generate the derived/calculated endpoint.
 - i. By subject listing of concomitant medications (as appropriate to the pivotal clinical trials)
 - j. By subject listing, of laboratory tests performed for safety monitoring

What else needs to be submitted

- Source data listings by site

Listing A: Screened subjects and reason for subjects who did not meet Eligibility requirements

Study: <#> Centre: 01001

Subject	Randomized	Reason for Screen Failure
50000	N	Consent withdrawn
50001	Y	
50002	Y	
50003	N	In/ex criteria
50004	Y	

.....

What else needs to be submitted

- Source data listings by site

Listing B: Treatment assignment (randomization)

Study: <#> Centre: 01001

Subject	Randomized Treatment
5001	<drug a>
5002	<drug b>

.....

What else needs to be submitted

➤ Source data listings by site

Listing C: Drop-outs and subjects that discontinued with date and reasons

Study: <#> Centre: 01001

Subject	Treatment Start Date	DC of <drug a>	Date of DC of <drug a>	Reason for DC of <drug a>	DC of <drug b>	Date of DC of <drug b>	Reason for DC of <drug b>
50018	<DDMMM YYYY>	Y	<DDMMM YYYY>	AE study dis. worse			
50039	<DDMMM YYYY>				Y	<DDMMM YYYY>	AE study dis. worse

.....

What else needs to be submitted

- Source data listings by site

Listing D: Evaluable subjects/non-evaluable subjects and reason not evaluable

Study: <#> Centre: 01001

Subject	Evaluable	Reason not Evaluable
50001	Y	
50002	Y	
50003	Y	
50004	N	Not Treated
50005	Y	
50006	Y	

.....

What else needs to be submitted

➤ Source data listings by site

Listing E: By subject listing of eligibility determination (i.e., inclusion And exclusion criteria)

Study: <#> Centre: 01001

Subject	Randomized	IN01	IN02	IN03	IN04	EX01	EX02	EX03	EX04
50001	Y	Y	Y	Y	Y	N	N	N	N
50002	Y	Y	Y	Y	Y	N	N	N	N
50003	Y	Y	Y	Y	Y	N	N	N	N
50004	Y	Y	Y	Y	Y	N	N	N	N

Y for Inclusion criteria and N for Exclusion criteria indicate patient satisfied criteria for inclusion in the study
IN1 = Inclusion criteria 1, EX1 = Exclusion criteria 1

What else needs to be submitted

➤ Source data listings by site

Listing F, By subject listing of AEs, SAEs and FataIs

Study: <#> Centre: 01001

Subject	Reported Terms	Preferred Terms	System Organ Class	AE Start Date	AE End Date	SAE Flag	Fatal Flag
50002	xxxx	xxxx	xxxx		<DDMMYYYY>	N	N
	xxxx	xxxx	xxxx	<DDMMYYYY>	<DDMMYYYY>	N	N
50005	xxxx	xxxx	xxxx	<DDMMYYYY>	<DDMMYYYY>	N	N

.....

What else needs to be submitted

➤ Source data listings by site

Listing G, By subject listing of protocol violations and/or deviations reported in the NDA, description of diviation/violation

Study: <#> Centre: 01009

Subject	Protocol Deviation Description	Visit
39456	Other PV affecting efficacy and possible safety	20

.....

What else needs to be submitted

➤ Source data listings by site

Listing H, By subject listing of the primary and secondary endpoints

Study: <#> Centre: 01001

Endpoint: <xxxxxxxxxx >

Subject	Visit Date	Visit Day	Time Point	<test 1>	<test2>	Observed record	Used for analysis
50001	<DDMMYYYY>	1	1 hour pre-dose	<value>	<value>	Y	Y

.....

What else needs to be submitted

➤ Source data listings by site

Listing I: By subject listing of concomitant medications

Study: <#> Centre: 01001

Subject	Reported Name	Indication	Start Date	End Date	<Med 1 name>	<Med 2 name>	<Med 3 name>
50001	xxxx	xxxx	<DDMMMYYYY>	<DDMMMYYYY>	Y		
	xxxx	xxxx	<DDMMMYYYY>	<DDMMMYYYY>			Y
50002	xxxx	xxxx	<DDMMMYYYY>	<DDMMMYYYY>		Y	

.....

What else needs to be submitted

➤ Source data listings by site

Listing J: By subject listing of laboratory tests performed for safety monitoring

Study: <#> Centre: 01001

Lab Group: Biochemistry

Subject	Actual Treatment	Visit	Visit Date	Visit Day	Alkaline Phosphatase U/L	Calcium mmol/L	Chloride mmol/L
50001	<treatment>	SCREENING	<DDMMYYYY>	-14	<value>	<value>	<value>
		BASELINE	<DDMMYYYY>	-1	<value>	<value>	<value>
		WEEK 12	<DDMMYYYY>	82	<value>	<value>	<value>

.....

.....

What else needs to be submitted?

Table of number of subjects screen, randomized, treated and prematurely discontinued by site

Study	Centre	Screened	Randomized	Treated	Early Discontinuation
3809.1	100001	4	0	0	0
	100002	4	0	0	0
	100003	4	2	2	0
	100004	12	9	9	0

Site Selection

What is the FDA doing with the information?

Site Selection

Site Selection:

Site selection should focus on:

- Impact to review of the study
 - **Largest site**
 - **Outliers**, e.g., site with large proportion of responders/trt effect
 - **Drop-outs; adverse events; protocol violations**
 - **Smaller than average variation in patient response**
- Impact to the clinical trial process
 - Volume of FDA-regulated research performed by the Clinical Investigator
 - Past inspectional history
 - Open door policy to reports of scientific misconduct

Site Selection

Site Selection: Additional Considerations

- Suspicion of false or fraudulent data
- Data that appear unrealistic
- Evidence that a sponsor has rejected data from an investigator
- Evidence of under-reporting or delay in submitting adverse events
- Evidence of inadequately monitored clinical investigations
- Evidence of inadequate or inappropriate informed consent
- Evidence of delayed or inappropriate IRB approval
- Evidence that an investigator has a significant financial interest in the product

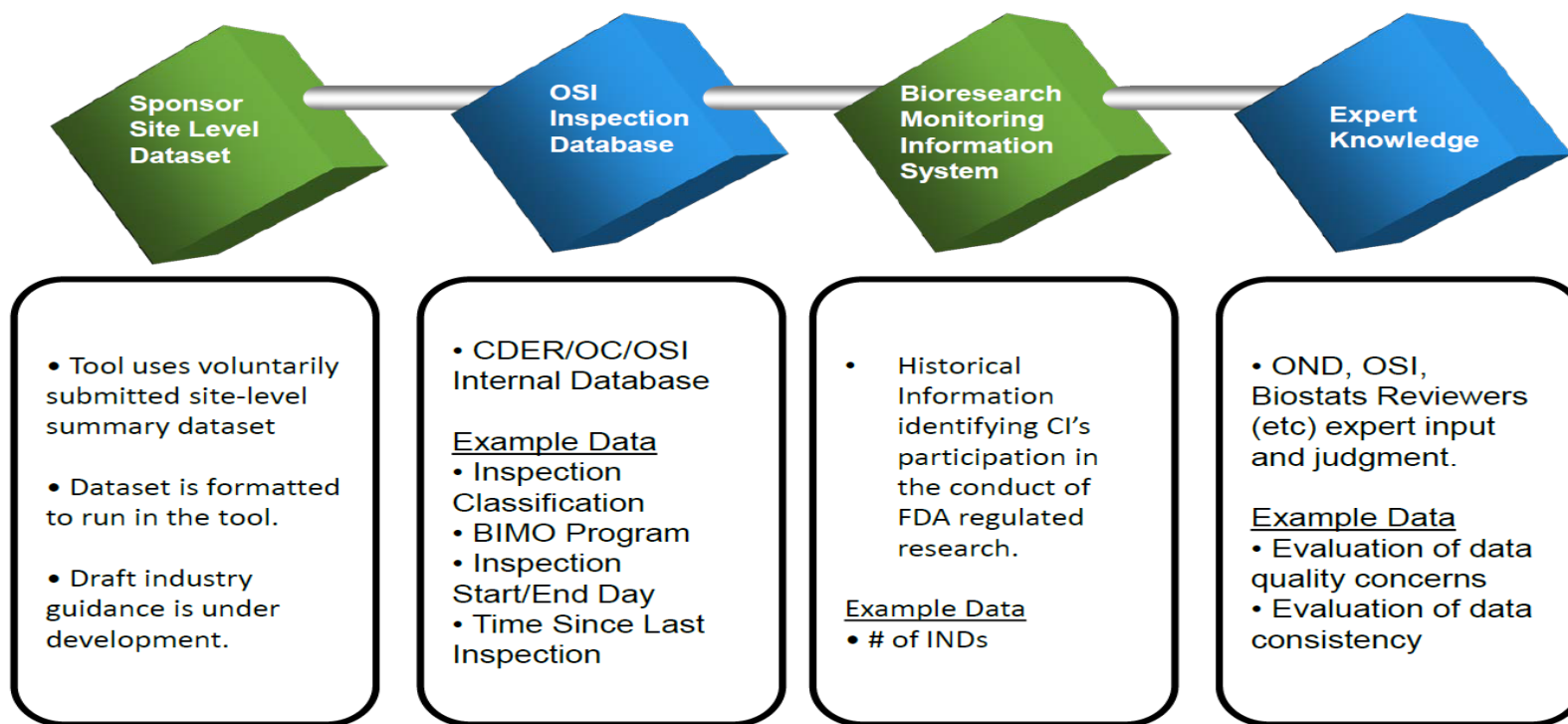
Site Selection



U.S. Food and Drug Administration
Protecting and Promoting Public Health

www.fda.gov

CDER Inspection Site Selection Tool Data Sources



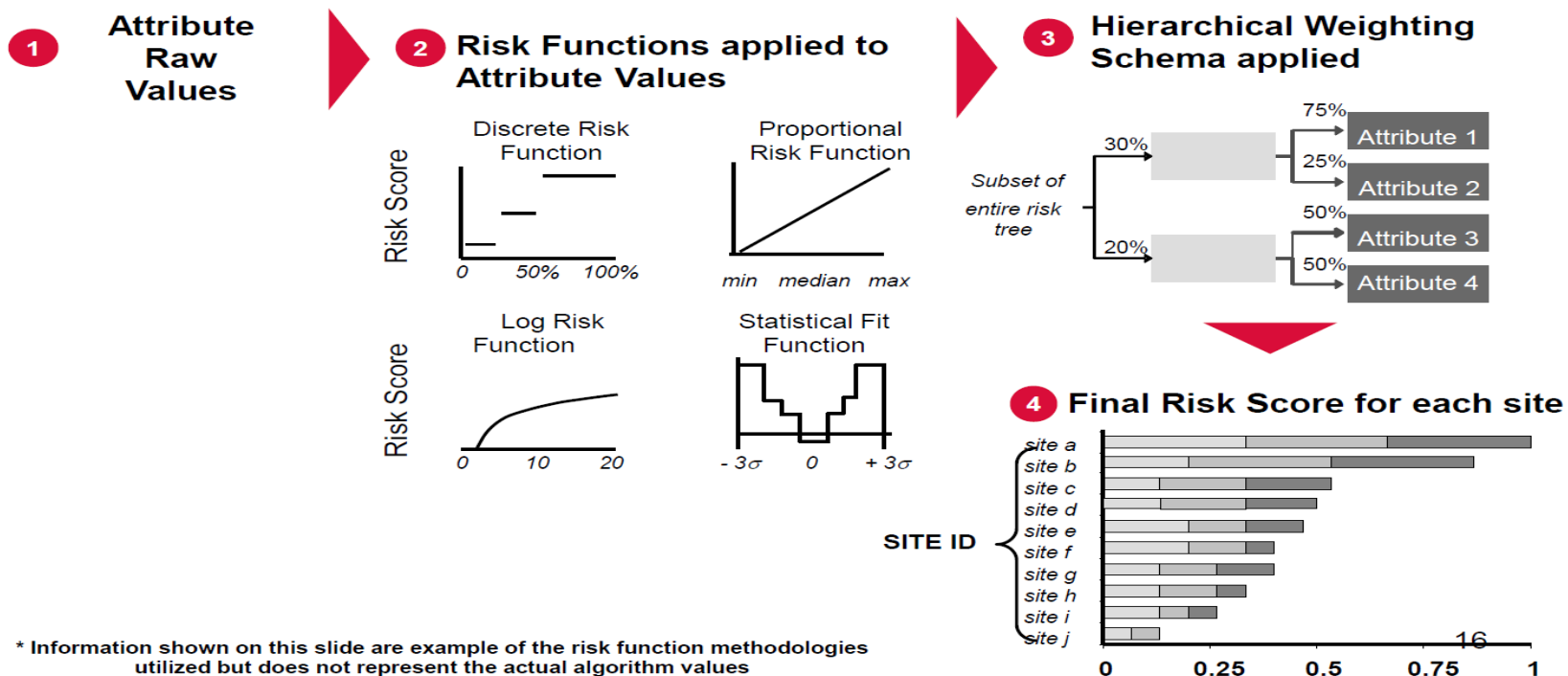
Site Selection



U.S. Food and Drug Administration
Protecting and Promoting Public Health

www.fda.gov

Input data is processed with a decision analysis algorithm



Site Selection Output

Risk Ranking of Sites

Site Identification		Risk Scores			
Site Number	PI Name	Rank	Total Score	Site Score	Trial Score
111		1	.5464	.1464	.2500
146		2	.5355	.1355	.2500
128		3	.5329	.1329	.2500
105		4	.5302	.1302	.2500
265		5	.5262	.1262	.2500
70		6	.5201	.1201	.2500
356		7	.5132	.1132	.2500
350		8	.5102	.1102	.2500
205		9	.5100	.1100	.2500
91		10	.5096	.1096	.2500
85		11	.5095	.1095	.2500
354		12	.5093	.1093	.2500
313		13	.5014	.1014	.2500
208		14	.5011	.1011	.2500
296		15	.5011	.1011	.2500

Future Goals and Plans



U.S. Food and Drug Administration
Protecting and Promoting Public Health

www.fda.gov

Future Goals

- Request updates to CDISC to include variables not currently present in SDTM (PhUSE working group gap analysis/recommendations nearing completion)
- HLA7 Pilot – Potential in future for direct pull of CI address and contact information
- Begin to develop methods to pull site data for those variables available at patient level in SDTM format
- Enhanced statistical methods to detect data irregularities and outliers at the site-level
- Ongoing development of learning algorithm so risk attributes and weights will evolve over time with a greater understanding of risk based on actual inspection results

26

Future Goals and Plans

- FDA continues to refine ‘The Request for Site Level Data Set’ document. You should always check on the FDA website to get the most current version.
- FDA may want this information (or most of it) during conduct for the purpose of 'real time inspections’.
- FDA to develop an inspection data warehouse and learning algorithm to select problem sites with a greater accuracy.
- IRB and Bioequivalence inspection models are under development.

Jingwei Gao (郜竞薇)

Regional Head of Statistical Programming, Pan Asia/META

Boehringer Ingelheim (China)Investment Co., Ltd.

29/F, Park Place, 1601 Nanjing Road(West), Jingan District, Shanghai, 200040

Work Phone: +86-21-52883958

Fax: +86-21-58780088

E-mail: jingwei.gao@boehringer-ingelheim.com

Web: www.boehringer-ingelheim.cn



Nancy Bauer

Regional Head of Statistical Programming, United States

Boehringer Ingelheim Pharmaceuticals, Inc.

900 Ridgebury Road, Ridgefield, CT, US 06877-0368

Work Phone: +01 203-798-5045

E-mail: nancy.bauer@boehringer-ingelheim.com

Web: www.boehringer-ingelheim.com



