WHY MODERN CDISC DATA VISUALIZATION MATTERS IN CLINICAL TRIALS

VALERIE NEDBAL, PH.D.
VALERIE.NEDBAL@JMP.COM

SR. JMP SYSTEM ENGINEER
SAS INSTITUTE GMBH.

INTRODUCTION

- **Efficient Data Review Practices**
  - Statistical analysis coupled with effective data visualization key to efficient data review
  - Data collection is a large contributor to time-consuming data review process
  - Snapshot comparison tools for tracking data changes enables early and often data review

INTRODUCTION

- Randomized clinical trial remains the gold standard for evaluating efficacy of an new intervention
- Safety profile assessment critical for trial success
- Safety analysis comes with several difficulties
  - Numerous endpoints measured repeatedly
  - Detection of rare events (drug-induced liver toxicity)
  - Limited studied population
  - Limited understanding of biological mechanisms and pathways

SAFETY REVIEW

- **Accelerated safety reviews**
  - Dynamic visualization coupled with tables and reports
  - Leverage data standards (CDISC)
  - Statistically-driven analyses with drill-down and swim-up capabilities
  - Centralized electronic data review coupled with tools that clinicians, data monitors, and biostatisticians can employ

EFFICIENT DATA REVIEW

ADVERSE EVENTS ANALYSIS

- Tables give all the information
- Time-consuming to absorb information and easy to miss signals

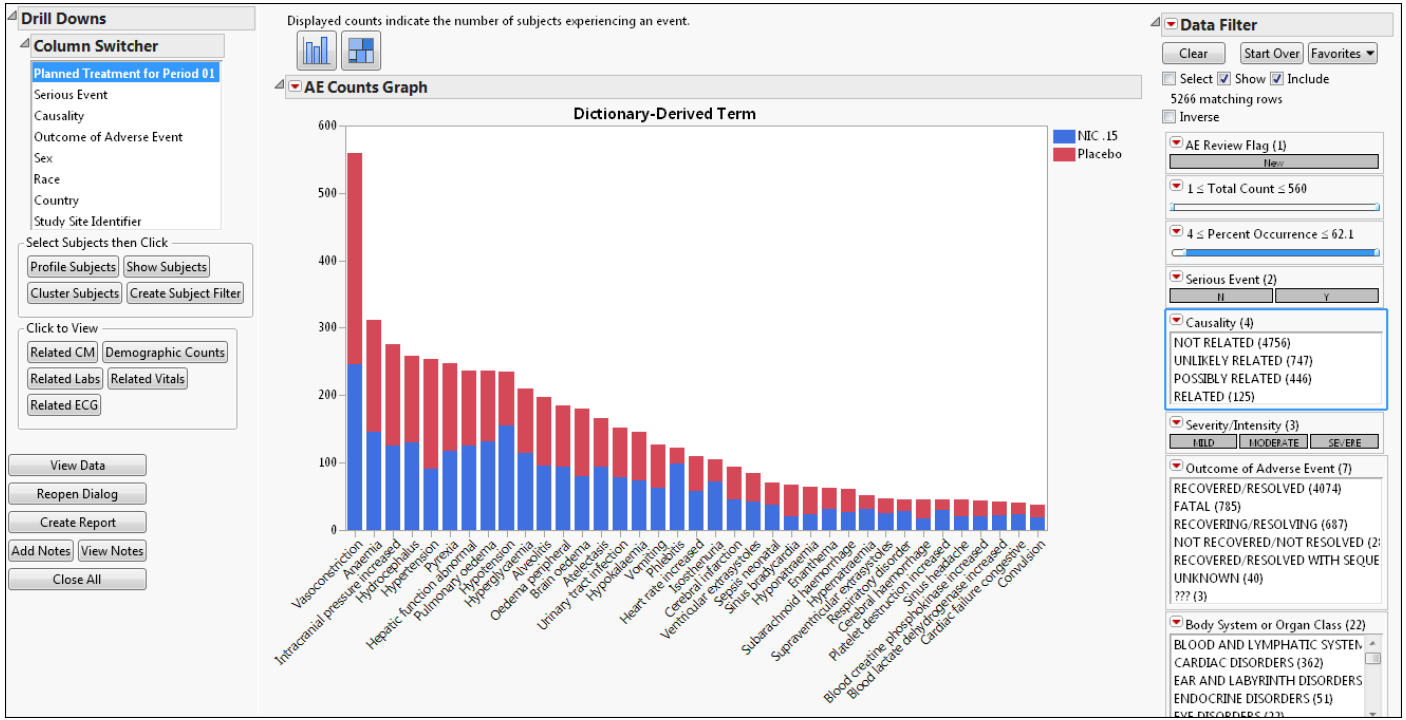
AE Event Occurrence & Proportion Report for Treatment
Reporting Events with at least Overall 4% Occurrence
Nicardipine: All Event Types for All Subjects Chosen

	Planned Treatment for Period 01		
	NIC .15	Placebo	Total
	Count (%)	Count (%)	Count (%)
Total	N=447	N=455	N=902
BLOOD AND LYMPHATIC SYSTEM DISORDERS	159(35.6%)	174(38.2%)	333(36.9%)
Anaemia	145 (32.4)	167 (36.7)	312 (34.6)
Platelet destruction increased	29 (6.5)	16 (3.5)	45 (5)
CARDIAC DISORDERS	91(20.4%)	113(24.8%)	204(22.6%)
Cardiac failure congestive	24 (5.4)	17 (3.7)	41 (4.5)
Sinus bradycardia	20 (4.5)	47 (10.3)	67 (7.4)
Supraventricular extrasystoles	25 (5.6)	22 (4.8)	47 (5.2)
Ventricular extrasystoles	42 (9.4)	43 (9.5)	85 (9.4)
GASTROINTESTINAL DISORDERS	63(14.1%)	64(14.1%)	127(14.1%)
Vomiting	63 (14.1)	64 (14.1)	127 (14.1)
GENERAL DISORDERS AND ADMINISTRATION SITE CONDITIONS	201(45%)	198(43.5%)	399(44.2%)
Enanthema	32 (7.2)	31 (6.8)	63 (7)
Oedema peripheral	94 (21)	91 (20)	185 (20.5)
Pyrexia	117 (26.2)	130 (28.6)	247 (27.4)
HEPATOBILIARY DISORDERS	126(28.2%)	111(24.4%)	237(26.3%)
Hepatic function abnormal	126 (28.2)	111 (24.4)	237 (26.3)

EFFICIENT DATA REVIEW

SUMMARY VISUALIZATIONS

- Summary views with capability to further drill down into the data



EFFICIENT DATA REVIEW

LINK TO PATIENT PROFILES AND NARRATIVES

EFFICIENT DATA REVIEW

LINK TO PATIENT PROFILES AND NARRATIVES

Subject: 101001
Randomized Arm: Placebo
Investigator: 101A

Subject 101001 was a 63-year-old white female. Her medical history included headache associated with sah (1988), hypertension with this sah (1988), vomiting associated with sah (1988) and hypertension prior to sah (1981). She began dosing with 40 mg/h of placebo on 21JAN1988 (Day 1). The subject discontinued the trial on 18FEB1988 (Day 29) due to death.

Other Significant Adverse Event (coded term [reported term]): HYDROCEPHALUS [HYDROCEPHALUS]

On 21JAN1988 (Day 1) the subject experienced a hydrocephalus (mild) which was considered a significant adverse event. At the time of the event, the subject was taking 40 mg/h of placebo and had been at this dose for 1 day. The significant AE occurred on the first day of dosing with any study medication. Trial medication had an action of drug withdrawn as a result of the event. It is not known from the case report form if therapeutic measures were administered to treat the event.

Adverse events that occurred within a ±3-day window of the onset of the significant AE included vasoconstriction (moderate) and vomiting (mild). Concomitant medications taken at the onset of the significant AE included potassium supplements, codeine compound 1/2, docusate sodium and multivitamins.

The investigator considered the AE to be not related to study medication. The final outcome of the event was reported as recovered/resolved on 02FEB1988 (Day 13).

Other Significant Adverse Event (coded term [reported term]): PYREXIA [PYREXIA]

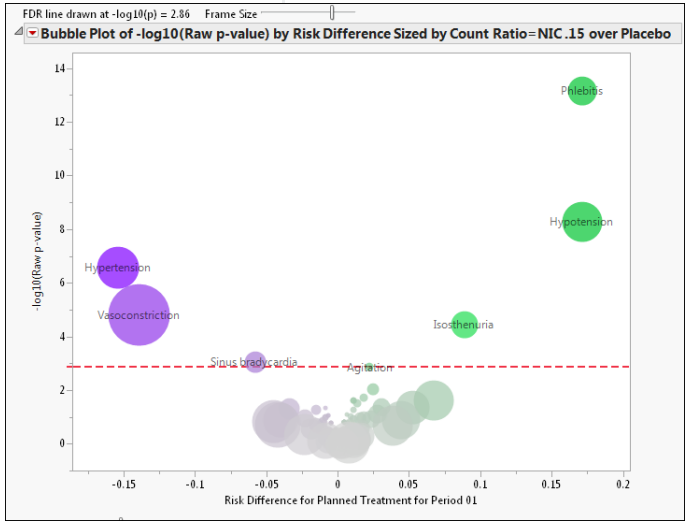
On 25JAN1988 (Day 5) the subject experienced a pyrexia (mild) which was considered a significant adverse event. At the time of the event, the subject was taking 40 mg/h of placebo and had been at this dose for 5 days. The significant AE occured 4 days after the first dose of any study medication. Trial medication had an action of dose not changed as a result of the event. It is not known from the case report form if therapeutic measures were administered to treat the event.

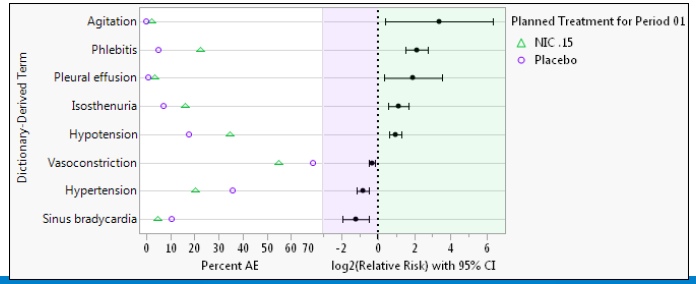
Adverse events that occurred within a ±3-day window of the onset of the significant AE included vomiting (mild). Concomitant medications taken at the onset of the significant AE included potassium supplements, docusate sodium, multivitamins and codeine compound 1/2.

The investigator considered the AE to be not related to study medication. The final outcome of the event was reported as recovered/resolved on 31JAN1988 (Day 11).

EFFICIENT DATA REVIEW

SAFETY SIGNAL DETECTION





- Statistically-driven volcano plots (Jin et al. 2001, Zink et al. 2013)
- Space-constrained view of several hundred AE events
- Difference in observed AE risk vs. statistical significance
- Color illustrates direction of effect
- Bubble size reflects AE frequency
- Traditional relative risk plot (Amit et al. 2008) to display interesting signals

ANALYSIS COMPLEXITIES

- Abundance of endpoints (multiplicity)
  - False discovery rate (FDR) Benjamini & Hochberg (1995)
  - Double FDR (Mehrotra & Heyse 2004, Mehrotra & Adewale, 2012)
  - Bayesian Hierarchical Models
- Repeated/recurrent events
  - Inclusion of time windows across analyses (Zink et al. 2013)
- Trial design complexity
  - Crossover analysis and visualization
- Limited population and understanding of biological underpinnings
  - Cross-domain predictive models
  - Subgroup analysis
  - Pharmacogenomics

DATA QUALITY AND FRAUD IN CLINICAL TRIALS

- Fraud is the “Deliberate attempt to deceive” or the “intention to cheat” (Buyse et al., 1999)

- Fraud thought to be uncommon
  - Investigators committing fraud estimated < 1% (Buyse et al., 1999)

- Difficult to diagnose
  - Lack of tools
  - Variations across subjects, time, sites make comparison difficult
  - Unusual points may indicate quality issues, but determining fraud requires more evidence (Evans, 2001)

CENTRALIZED ELECTRONIC DATA MONITORING

- “Several publications suggest that certain data anomalies (e.g., fraud, including fabrication of data, and other non-random data distributions) may be more readily detected by centralized monitoring techniques than by on-site monitoring” (FDA Guidance 2013)

- Identify data trends through statistical analysis
  - Missing/inconsistent data, outliers, protocol deviations
  - Unusual distributions of reported data across and among study sites

- Capitalize on:
  - Data standards (CDISC)
  - Graphical display of data through interactive software tools
  - Early and routine data reviews

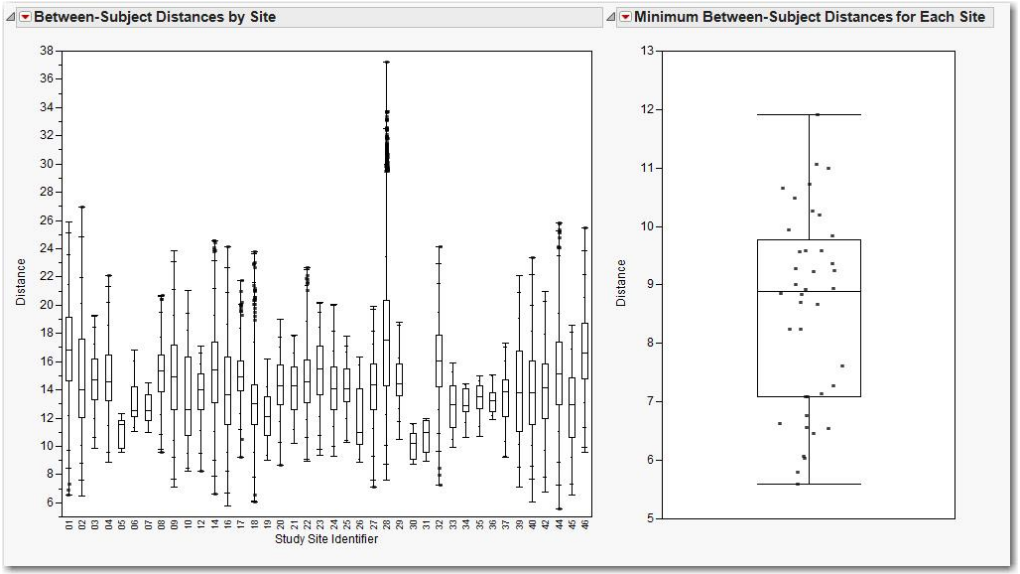
ANALYSIS AND TOOLS

- Types of Fraud
  - Site: modified or fabricated records of participants within study site
  - Patient: subjects with multiple enrollment across sites

- Straight-forward data quality analyses
  - Non-random data distributions
  - Variability and similarity of subjects within and across sites
  - Duplicate detection across measurements
  - Duplicated patient information (Birthdays/Initials)
  - Visit distribution occurrence and attendance

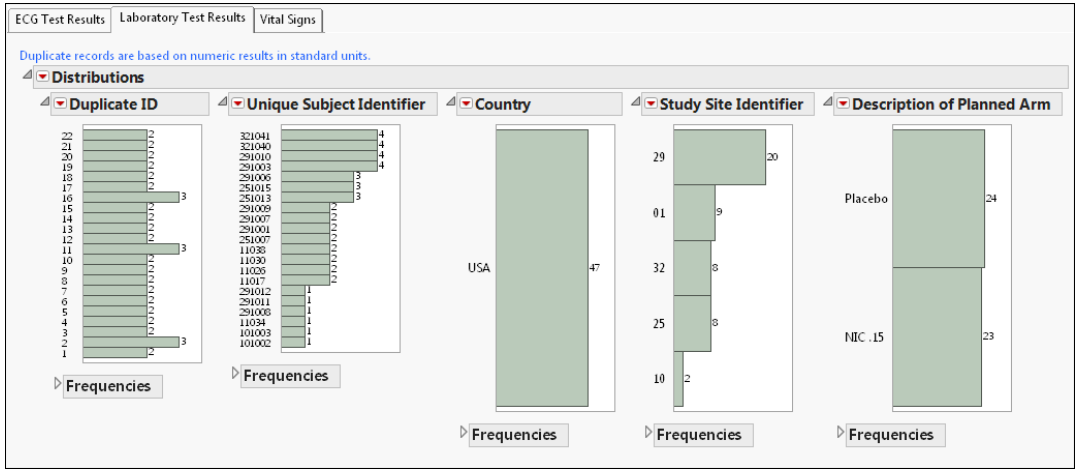
DUPLICATE SUBJECTS WITHIN SITE

- Compare individual data points between all pairs of subjects within a site
- Calculate the Euclidian distance between subjects
  - Investigator is not likely to make an exact copy
- Summarize by site for analysis of subject similarity



DUPLICATE RECORDS WITHIN SITE

- Falsify new records from existing data
- Example with Findings domain tests
  - Triplicate matching values of Systolic, Diastolic BP and Heart Rate vitals measurements
- Data quality based on trends of missing records



FRAUD AND DATA QUALITY

PATIENTS WITH MULTIPLE ENROLLMENT

- For access to drug or other reimbursement, patients enroll at two or more sites
- Screen failure, independence and sample size concerns
- Match on birth date or initials, can summarize demographics, height, weight
- More complex analyses could cluster across sites using baseline information

Matching Variable Value	Unique Subject Identifier	Study Site Identifier	Date/Time of Birth	Sex	Race	N
15Apr1915	141001	14	1915-04-15	M	WHITE	1
	442001	44	1915-04-15	F	WHITE	1
01Jan1917	392012	39	1917-01-01	F	OTHER	1
	41018	04	1917-01-01	M	WHITE	1
11Aug1925	221020	22	1925-08-11	M	BLACK OR AFRICAN AMERICAN	1
	282003	28	1925-08-11	F	WHITE	1
28Jul1928	241003	24	1928-07-28	F	WHITE	1
	81011	08	1928-07-28	F	WHITE	1
13Nov1935	141021	14	1935-11-13	F	OTHER	1
	231027	23	1935-11-13	F	WHITE	1
25Aug1937	31006	03	1937-08-25	M	WHITE	1
	31012	03	1937-08-25	F	WHITE	1
23Mar1938	171001	17	1938-03-23	M	WHITE	1
	271014	27	1938-03-23	F	WHITE	1
29May1939	11034	01	1939-05-29	F	WHITE	1
	161003	16	1939-05-29	F	WHITE	1
21Jan1941	282008	28	1941-01-21	F	BLACK OR AFRICAN AMERICAN	1
	351006	35	1941-01-21	M	WHITE	1
09Jul1941	11033	01	1941-07-09	F	WHITE	1
	392017	39	1941-07-09	F	WHITE	1
27Aug1943	11048	01	1943-08-27	F	WHITE	1
	291002	29	1943-08-27	F	ASIAN	1
22Feb1947	201018	20	1947-02-22	M	WHITE	1
	81023	08	1947-02-22	M	WHITE	1
23Jan1949	141013	14	1949-01-23	F	ASIAN	1
	461023	46	1949-01-23	M	WHITE	1
09Feb1954	161016	16	1954-02-09	F	WHITE	1
	241011	24	1954-02-09	F	WHITE	1
15Oct1955	141054	14	1955-10-15	F	WHITE	1
	231020	23	1955-10-15	F	WHITE	1
18Oct1956	91038	09	1956-10-18	F	BLACK OR AFRICAN AMERICAN	1
	91040	09	1956-10-18	F	BLACK OR AFRICAN AMERICAN	1

RISK-BASED MONITORING

- ICH Guideline E6 (1996)
  - “sponsor should ensure trials are adequately monitored”
  - “sponsor should determine the appropriate extent and nature of monitoring”
  - “statistically controlled sampling may be an acceptable method for selecting the data to be verified.”
- 100% source data verification
  - Error-prone (Tantsyura et al., 2010)
  - Expensive, as much as 25-30% of trial cost (Eisenstein et al, 2005; Funning et al., 2009; Tantsyura et al., 2010)
  - Time-consuming on fields of little importance
  - Limited-in-scope (comparisons across pages, subjects and sites)

RISK-BASED MONITORING

FDA Guidance (August 2013) recommends risk-based monitoring, including centralized monitoring, where appropriate

- Apply statistical sampling to CRF pages for review
- Minimize the number of site visits
- Limit amount of manual work
- Sampling rates can vary based on data importance
  - Primary endpoint and SAEs (100%)
  - Medical history or physical exams (0%)
- Targeted on-site visits based on risk evaluation

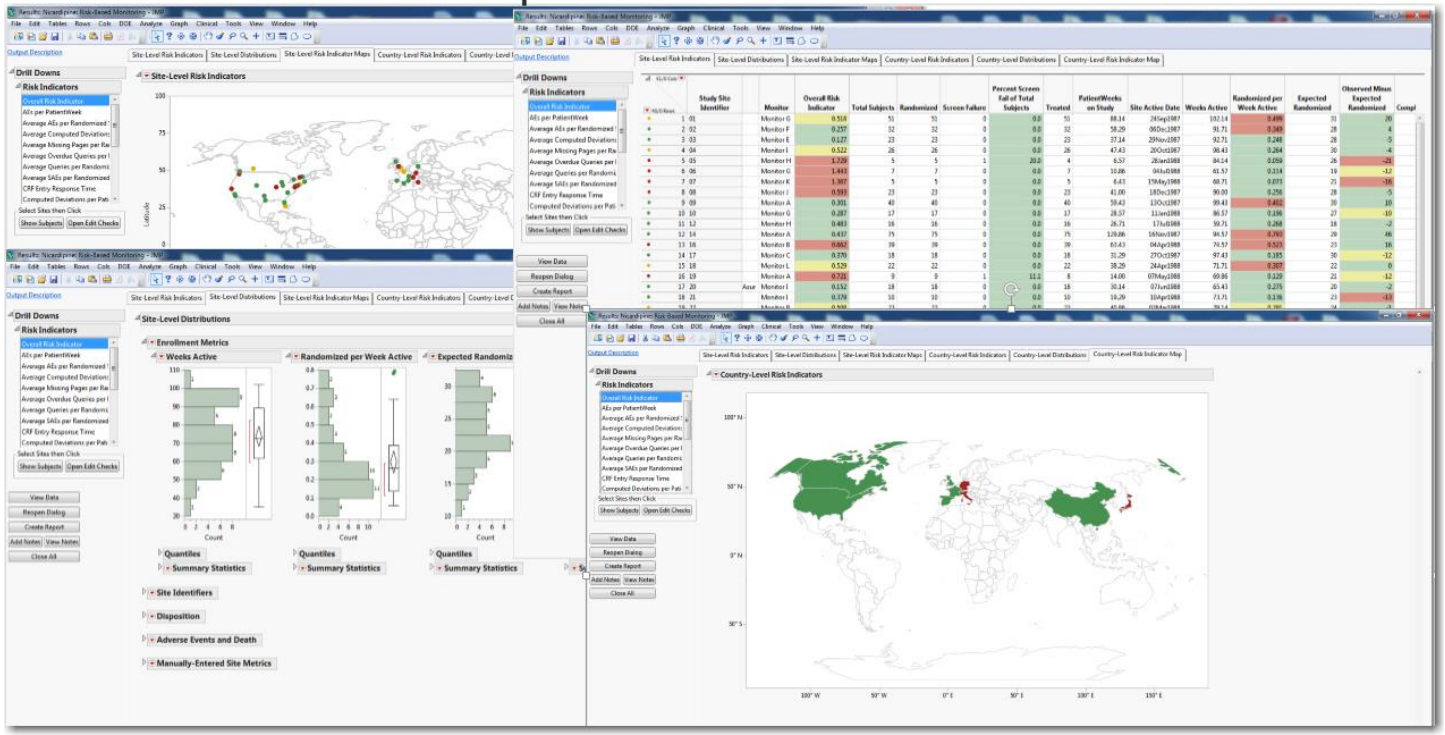
- Centralized monitoring capable of detecting over 90% of findings identified by on-site monitoring (Bakobaki et al., 2012)

EXAMPLE RISK-BASED MONITORING METHODS

- Create indicators of excessive:
  - Screen Failures
  - Discontinuations
  - Serious Adverse Events
  - Deaths
  - Dose interruptions
  - Queries
  - Protocol Deviations
  - Missing CRF Pages
  - Poor Query Response
- Indicators can be examined individually or a weighted combination can give an overall score
- High scores on items of particular importance will necessitate onsite review

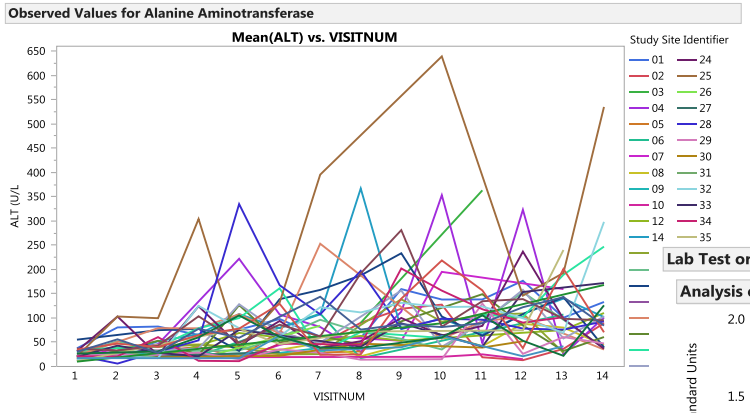
EXAMPLE RISK-BASED MONITORING METHODS

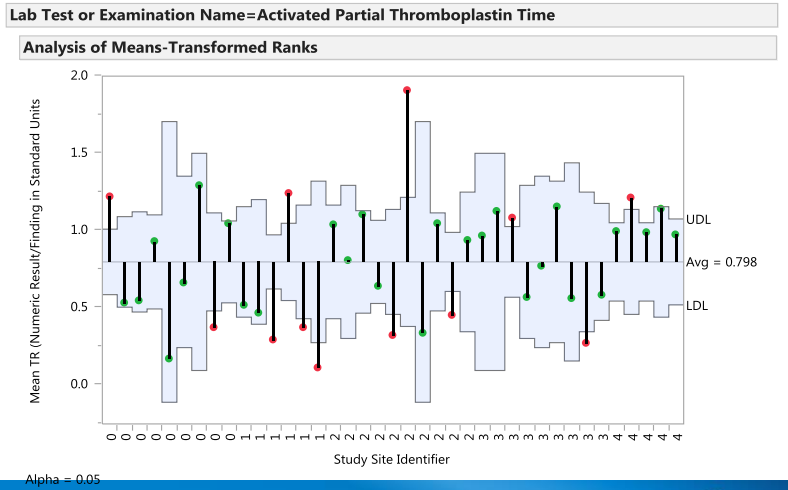
- Employ risk indicators to rate potentially problematic sites for further follow-up



RISK-BASED MONITORING WITH FINDINGS TRENDS

- Detect unusual trends in findings measurements across sites





CONCLUSIONS

- Statistically-driven, dynamic data visualization necessary for efficient clinical safety review
- Data standards allow for automated analyses that enable clinicians, data monitors, data managers, and statisticians
- Tools for snapshot comparison accelerate reviews
- Centralized-monitoring enhances the accessibility and transparency of data
- Several straight-forward analyses to interrogate data quality and potentially fraudulent activities

CONCLUSIONS

- Statistically-driven, dynamic data visualization necessary for efficient clinical safety review
- Data standards allow for automated analyses that enable clinicians, data monitors, data managers, and statisticians

REFERENCES

- Amit O, Heiberger RM and Lane PW. Graphical approaches to the analysis of safety data from clinical trials. *Pharmaceutical Statistics* 2008; 7: 20-35.
- Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* B 1995; 57: 289–300.
- Jin W, Riley RM, Wolfinger RD, White KP, Passador-Gurgel G and Gibson G. The contributions of sex, genotype and age to transcriptional variance in Drosophila melanogaster. *Nature Genetics* 2001; 29: 389-395.
- Mehrotra DV and Adewale AJ. Flagging clinical adverse experiences: reducing false discoveries without materially compromising power for detecting true signals. *Statistics in Medicine* 2012 (in press).
- Mehrotra DV and Heyse JF. Use of the false discovery rate for evaluating clinical safety data. *Statistical Methods in Medical Research* 2004; 13: 227-238.
- Zink RC, Wolfinger RD, Mann G (2012). Summarizing the incidence of adverse events using volcano plots and time intervals. *Clinical Trials*. **10(3)**:398-406.

FOR MORE INFORMATION;
Contact valerie.nedbal@jmp.com