Workstream 1
Retrospective Data Analysis

May 4, 2011
Overview

- Background
- Data Cleaning
- Results
- Limitations
- Lessons Learned
Background

- Data received from 18 organizations:
  - Academic (2)
  - CRO-ARO (4)
  - Biotechnology (2)
  - Device (2)
  - Government (1)
  - Investigators (2)
  - Pharmaceutical (6)

- Objective was to collect “simple” metrics
## Retrospective Collection - Data Elements

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Suggested Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site ID</td>
<td>List unique site identifier as per your system</td>
</tr>
<tr>
<td>Protocol ID</td>
<td>List a protocol identifier as per your system</td>
</tr>
<tr>
<td>Site Type</td>
<td>Type of site conducting the study (Academic, Hospital based, private practice, independent, VA)</td>
</tr>
<tr>
<td>IRB Type</td>
<td>Central, Local, Regional</td>
</tr>
<tr>
<td>Therapeutic Area</td>
<td>CV/Metabolic, Hematology &amp; Oncology, Infectious Diseases, Neurosciences, Vaccines, Ophthalmology, Pain, Dermatology, Inflammation, Allergy &amp; Respiratory, Urology, Women’s Health, Device and Other</td>
</tr>
<tr>
<td>If TA is &quot;Other&quot; - Specify</td>
<td>If TA is other than one of the options given; enter specific TA</td>
</tr>
<tr>
<td>Sponsor Name</td>
<td>List name of the sponsor for the study</td>
</tr>
<tr>
<td>Field Name</td>
<td>Suggested Definition</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Date protocol sent by sponsor/CRO to site</td>
<td>Date the protocol was sent by the sponsor/CRO to the site. If as a site you do not know this date, leave this field blank.</td>
</tr>
<tr>
<td>Date protocol received by site</td>
<td>Date the protocol was received by the site. If as a sponsor/CRO you do not know this date, leave this field blank.</td>
</tr>
<tr>
<td>Date of protocol submitted by site for IRB approval</td>
<td>Date the protocol is submitted by the site to the IRB of record</td>
</tr>
<tr>
<td>Date of final IRB decision on the protocol</td>
<td>Date the IRB approves the protocol</td>
</tr>
<tr>
<td>Date final signature obtained to fully execute site contract</td>
<td>Date the final signature is obtained on the contract and contract is considered fully executed</td>
</tr>
<tr>
<td>Date site enrolled first patient</td>
<td>Date the site enrolled the first patient into the study (date informed consent signed)</td>
</tr>
</tbody>
</table>
Data Cleaning

Evaluated blank fields to determine if it could be reasonably derived from other fields provided in the sample.

Examples:

- Identification of IRB type from name of IRB
- Identification of site type from name of site
- Reassigning therapeutic area to a broader category
- Filtered a site’s data to include uniform phase data; i.e., all from Phase II, Phase III, or Device studies
- Some data submitted included calculated cycle times without source dates; data still used when we could identify a match with our cycle time definitions
Data Cleaning

- For cycle times, did not include any cycle times that were negative

- Different cutpoints were considered, but all seemed arbitrary

- Cutpoint of zero allowed outlier observations to still be included in the analysis
## Data Cleaning

<table>
<thead>
<tr>
<th>Organization Type</th>
<th># Original Data Lines</th>
<th># Data Lines After Cleaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>4,516</td>
<td>559</td>
</tr>
<tr>
<td>ARO-CRO</td>
<td>703</td>
<td>703</td>
</tr>
<tr>
<td>Biotech</td>
<td>512</td>
<td>512</td>
</tr>
<tr>
<td>Device</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td>Govt</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Investigators</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>Pharma</td>
<td>4,769</td>
<td>3,449</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>10,673</strong></td>
<td><strong>5,396</strong></td>
</tr>
</tbody>
</table>
Results Overview

- Overview of characteristics of data submitted
  - Submitting organization
  - Site type
  - IRB type
  - Therapeutic area

- Missing Data

- Cycle time metrics with various stratifications
  - Protocol sent/received to protocol submitted to IRB
  - Protocol sent/received to IRB final decision
  - Protocol sent/received to contract execution
  - Protocol sent/received to 1\textsuperscript{st} patient enrolled
  - Protocol submission to IRB to IRB final decision
  - IRB final decision to 1\textsuperscript{st} patient enrolled
  - Contract executed to 1\textsuperscript{st} patient enrolled
Who Submitted the Data?
How Much Data Were Received?
Results

Totals by Organization Type

- Acad: N=559
- ARO-CRO: N=703
- Biotech: N=512
- Device: N=79
- Govt: N=51
- Invest: N=43
- Pharm: N=3449
What is the Distribution of Site Types in the Data?
Results

Totals by Site Type

- Academic: N=1431
- Hospital Based: N=358
- Independent: N=134
- Private Practice: N=1326
- VA: N=33
- Missing: N=2114
Within Each Organization, What is the Distribution of Site Type?
Results

Site Type for ARO-CRO Organizations

- Academic
- Hospital Based
- Independent
- Private Practice
- VA
Results

Site Type for Biotechnology Organizations

- Academic
- Hospital Based
- Independent
- Private Practice
- VA
Results

Site Type for Device Organizations

Academic
Hospital Based
Independent
Private Practice
VA
Results

Site Type for Investigators Organizations

Academic | Hospital Based | Independent | Private Practice | VA

0 | 0 | 0 | 30 | 0
Results

Site Type for Pharma Organizations

- Academic
- Hospital Based
- Independent
- Private Practice
- VA
Which Therapeutic Areas are Presented in the Data?
Results
## Results

<table>
<thead>
<tr>
<th>Other Therapeutic Area</th>
<th>Abbreviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>MED</td>
<td>120</td>
</tr>
<tr>
<td>Endocrinology</td>
<td>ENCRN</td>
<td>110</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>GASTR</td>
<td>109</td>
</tr>
<tr>
<td>Immunoscience</td>
<td>IMSCI</td>
<td>106</td>
</tr>
<tr>
<td>Cystic Fibrosis</td>
<td>CF</td>
<td>73</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>PEDS</td>
<td>61</td>
</tr>
<tr>
<td>Surgery</td>
<td>SURG</td>
<td>43</td>
</tr>
<tr>
<td>Perioperative</td>
<td>PERI</td>
<td>22</td>
</tr>
<tr>
<td>Anesthesiology</td>
<td>ANESTH</td>
<td>19</td>
</tr>
<tr>
<td>ABGEM</td>
<td>ABGEM</td>
<td>14</td>
</tr>
<tr>
<td>Radiology</td>
<td>RAD</td>
<td>14</td>
</tr>
<tr>
<td>GCRC</td>
<td>GCRC</td>
<td>13</td>
</tr>
<tr>
<td>Women’s Health</td>
<td>WH</td>
<td>12</td>
</tr>
<tr>
<td>School of Nursing</td>
<td>NURS</td>
<td>11</td>
</tr>
<tr>
<td>Comm &amp; Fam Med</td>
<td>FAM</td>
<td>10</td>
</tr>
<tr>
<td>Bone</td>
<td>BONE</td>
<td>9</td>
</tr>
<tr>
<td>Urology</td>
<td>UROL</td>
<td>7</td>
</tr>
<tr>
<td>Orthopedic Surgery</td>
<td>ORTHO</td>
<td>6</td>
</tr>
<tr>
<td>Rheumatology</td>
<td>RHEUM</td>
<td>4</td>
</tr>
<tr>
<td>Compassionate Use</td>
<td>COMP</td>
<td>3</td>
</tr>
<tr>
<td>Dermatology</td>
<td>DERM</td>
<td>2</td>
</tr>
<tr>
<td>Ped Migraine</td>
<td>MIGR</td>
<td>2</td>
</tr>
<tr>
<td>Flu</td>
<td>FLU</td>
<td>1</td>
</tr>
<tr>
<td>IDTI</td>
<td>IDTI</td>
<td>1</td>
</tr>
<tr>
<td>Vaccines</td>
<td>VAC</td>
<td>1</td>
</tr>
</tbody>
</table>
What is the Distribution of Use of Central vs. Local IRBs?
Results

All Data: IRB Type

Central: N=2781

Local: N=1829
Results

Type of IRB for Each Site Type

- Private Practice
- Academic
- Hospital Based
- Independent
- VA
Results

Type of IRB for Each Therapeutic Area

- NEURO
- HEM-ONC
- OTH
- CV
- PULM
- PAIN
- INFLAM
- IMMUN
- ID
- OPHTHALM
- DEVICE

Total

Legend:
- Central
- Local
What was the Extent of the Missing Data?
Results

Date Sent by Sponsor: Percent Missing

- All Data
- Academic
- ARO-CRO
- Biotechnology
- Device
- Government
- Investigators
- Pharma
Results

Date Received by Site: Percent Missing

<table>
<thead>
<tr>
<th>Category</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data</td>
<td></td>
</tr>
<tr>
<td>Academic</td>
<td></td>
</tr>
<tr>
<td>ARO-CRO</td>
<td></td>
</tr>
<tr>
<td>Biotechnology</td>
<td></td>
</tr>
<tr>
<td>Device</td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td></td>
</tr>
<tr>
<td>Investigators</td>
<td></td>
</tr>
<tr>
<td>Pharma</td>
<td></td>
</tr>
</tbody>
</table>
Results

Date Submitted to IRB: Percent Missing

<table>
<thead>
<tr>
<th>Group</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Data</td>
<td>60</td>
</tr>
<tr>
<td>Academic</td>
<td>90</td>
</tr>
<tr>
<td>ARO-CRO</td>
<td>40</td>
</tr>
<tr>
<td>Biotechnology</td>
<td>30</td>
</tr>
<tr>
<td>Device</td>
<td>70</td>
</tr>
<tr>
<td>Government</td>
<td>100</td>
</tr>
<tr>
<td>Investigators</td>
<td>60</td>
</tr>
<tr>
<td>Pharma</td>
<td>50</td>
</tr>
</tbody>
</table>
Results

Date of IRB Final Decision: Percent Missing

- All Data
- Academic
- ARO-CRO
- Biotechnology
- Device
- Government
- Investigators
- Pharma

Percent
Results

Date of Final Signature: Percent Missing

- All Data
- Academic
- ARO-CRO
- Biotechnology
- Device
- Government
- Investigators
- Pharma
Results

Date First Patient Enrolled: Percent Missing

Percent

0 20 40 60 80 100

All Data Academic ARO-CRO Biotechnology Device Government Investigators Pharma
Results
Results
What are the Cycle Time Results?
Results

All Data: Date Protocol Submitted to IRB to Date of IRB Final Decision

Days

All Data
Results

All Data: Date Protocol Submitted to IRB to Date of IRB Final Decision

median

mean
Date Protocol Sent/Received to Date Protocol Submitted to IRB
Date Protocol Sent/Received to Date Protocol Submitted to IRB

- All Data
  - $p < 2.2 \times 10^{-16}$

- By Site Type
  - $p < 2.2 \times 10^{-16}$

- By IRB Type
  - $p < 2.2 \times 10^{-16}$
Date Protocol Sent/Received to Date of IRB Final Decision
Date Protocol Sent/Received to Date of IRB Final Decision

- p < 2.2 E-16
- By Site Type
- By IRB Type
Date Protocol Sent/Received to Date Contract Executed
Date Protocol Sent/Received to Date Contract Executed

p < 2.2 E-16

By Site Type

p < 2.2 E-16

By IRB Type
Date Protocol Sent/Received to Date First Patient Enrolled
Date Protocol Sent/Received to Date First Patient Enrolled

- All Data: $p = 7.792 \times 10^{-16}$
- By Site Type: $p < 2.2 \times 10^{-16}$
- By IRB Type: $p < 2.2 \times 10^{-16}$
Date Protocol Submitted to IRB to Date of IRB Final Decision
Date Protocol Submitted to IRB to Date of IRB Final Decision

p < 2.2 E-16

By Site Type

p < 2.2 E-16

By IRB Type
Date of IRB Final Decision to Date First Patient Enrolled
Date of IRB Final Decision to Date First Patient Enrolled

- All Data
- By Site Type
  - p = 3.051 E-6
- By IRB Type
  - p = 2.173 E-9
Date Contract Executed to Date First Patient Enrolled
Date Contract Executed
to Date First Patient Enrolled

p < 2.2 E-16

p = 0.4511
Limitations

- Sample size small in many areas
- Limited data submitted by sites themselves
- Lack of standard data elements definitions
  - Data elements definitions provided were not granular enough
- Goal of data analysis - to show variation and themes
- Large percentage of missing data
  - Data element(s) not routinely captured by submitting organization
  - Derived data where possible
Lessons Learned

- Need to develop understanding of a common workflow
- Need to develop enterprise wide standard definitions for data elements
- Need to agree on which cycle time metrics are critical for monitoring study start-up
- Avoid distraction of site-specific processes
- Must provide the “what’s in it for me” explanation to facilitate data collection
- Must be as granular as possible in providing definitions
- Need to provide drop down option lists when possible
- Need more non-date data collected to facilitate identification of processes that could be responsible for cycle time variation
Questions?