“Table 1”

Slides by
Rachel Richesson,
co-Chair, Phenotype, Data Standards, and Data Quality Core
### Table 1. Participant Characteristics at Baseline

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Nifedipine (n = 201)</th>
<th>Placebo (n = 205)</th>
<th>Nonrandomization (n = 230)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, mean (SD), y</strong></td>
<td>30.2 (5.1)</td>
<td>30.2 (5.1)</td>
<td>28.7 (5.5)</td>
</tr>
<tr>
<td><strong>Body mass index, mean (SD) b</strong></td>
<td>23.3 (4.7)</td>
<td>23.3 (4.3)</td>
<td>23.5 (5.0)</td>
</tr>
<tr>
<td><strong>White race</strong></td>
<td>166 (82.6)</td>
<td>165 (79.2)</td>
<td>166 (73.2)</td>
</tr>
<tr>
<td><strong>Nulliparous</strong></td>
<td>116 (57.7)</td>
<td>115 (56.2)</td>
<td>116 (50.0)</td>
</tr>
<tr>
<td><strong>Prior preterm birth</strong></td>
<td>39 (19.4)</td>
<td>40 (19.5)</td>
<td>39 (19.4)</td>
</tr>
<tr>
<td><strong>Gestational age at study entry, mean (SD), wk</strong></td>
<td>29.2 (1.7)</td>
<td>29.2 (1.7)</td>
<td>29.2 (1.7)</td>
</tr>
<tr>
<td><strong>Multiple gestation</strong></td>
<td>42 (20.9)</td>
<td>42 (20.9)</td>
<td>42 (20.9)</td>
</tr>
<tr>
<td><strong>Twins</strong></td>
<td>40 (19.9)</td>
<td>40 (19.9)</td>
<td>40 (19.9)</td>
</tr>
<tr>
<td><strong>Diagnosis of PROM at study entry</strong></td>
<td>53 (26.4)</td>
<td>53 (26.4)</td>
<td>53 (26.4)</td>
</tr>
<tr>
<td><strong>Vaginal bleeding at study entry</strong></td>
<td>38 (18.9)</td>
<td>38 (18.9)</td>
<td>38 (18.9)</td>
</tr>
<tr>
<td><strong>Additional tocolysis</strong></td>
<td>12 (6.0)</td>
<td>12 (6.0)</td>
<td>12 (6.0)</td>
</tr>
<tr>
<td><strong>Vaginal examination at study entry</strong></td>
<td>(n = 134)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td><strong>Dilatation, median (IQR), cm</strong></td>
<td>0 (0-1)</td>
<td>0 (0-1)</td>
<td>0 (0-1)</td>
</tr>
<tr>
<td><strong>Cervical length, median (IQR), mm</strong></td>
<td>25 (15-35)</td>
<td>25 (15-35)</td>
<td>25 (15-35)</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; NA, not applicable; pPROM, preterm PROM, pregnant with preterm PROM.

**Table 1. Patient Characteristics**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Paracetamol (n = 33)</th>
<th>Morphine (n = 38)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>18 (54.5)</td>
<td>26 (68.4)</td>
<td>.23</td>
</tr>
<tr>
<td>Female</td>
<td>15 (45.5)</td>
<td>12 (31.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Age at surgery, median (IQR), d</strong></td>
<td>5 (1.5-6.5)</td>
<td>20 (1.8-8.7)</td>
<td>.50</td>
</tr>
<tr>
<td><strong>Age, y</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥10</td>
<td>17 (51.5)</td>
<td>16 (47.4)</td>
<td>.73</td>
</tr>
<tr>
<td>&gt;10</td>
<td>16 (48.5)</td>
<td>20 (52.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Weight, mean (SD), kg</strong></td>
<td>3.8 (1.3)</td>
<td>4.4 (2.0)</td>
<td>.17</td>
</tr>
<tr>
<td><strong>Duration of surgery, mean (SD), min</strong></td>
<td>172.1 (83.7)</td>
<td>156.6 (87.9)</td>
<td>.45</td>
</tr>
<tr>
<td><strong>Surgical procedure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thoracic</td>
<td>5 (15.2)</td>
<td>11 (28.9)</td>
<td>.17</td>
</tr>
<tr>
<td>Abdominal</td>
<td>28 (84.8)</td>
<td>27 (71.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Postoperative mechanical ventilation</strong></td>
<td>15 (45.5)</td>
<td>14 (36.8)</td>
<td>.46</td>
</tr>
<tr>
<td><strong>Duration of postoperative ventilation, median (IQR), h</strong></td>
<td>34 (15-45)</td>
<td>23 (16-45)</td>
<td>.43</td>
</tr>
<tr>
<td><strong>Surgical stress score, median (IQR)</strong></td>
<td>10 (9-11)</td>
<td>10 (9-11)</td>
<td>.75</td>
</tr>
<tr>
<td><strong>PRISM2, median (IQR)</strong></td>
<td>2 (0-4.5)</td>
<td>3 (0.0-5.0)</td>
<td>.91</td>
</tr>
<tr>
<td><strong>PIM2, median (IQR), % of mortality</strong></td>
<td>1.3 (0.6-1.9)</td>
<td>1.4 (0.7-2.4)</td>
<td>.34</td>
</tr>
</tbody>
</table>

Abbreviations: IQR, interquartile range; PIM2, Pediatric Index of Mortality 2; PRISM2, Pediatric Risk of Mortality 2.
Project Aim

• To identify important person characteristics and clinical features, along with explicit definitions and representations, for the reporting of baseline characteristics of research populations in interventional and observational studies.
Table 1 supports:

- Submission of data sets and data results from NIH-funded studies for archival and secondary use and for analyses and comparisons across trials
- Standardizing reporting of results from NIH-funded studies to ClinicalTrials.gov
- Better practices for describing populations in submissions to medical journals
- Conduct of multisite pragmatic trials
- Conduct of multisite observational studies
- Others?
Question 1: What characteristics should be on Table 1?

Question 2: How should they be defined? represented?

Question 3: Do we want to standardize reporting conventions? Categories?

- e.g., continuous variables: mean and SD
- median, 25th, 75th %-ile
- categories (e.g., deciles of age)
Different Definitions Yield Different Cohorts

N=24,520

- Presence of ICD-9 Diagnosis Codes Indicative of Diabetes (n = 18,980)
- Presence of Diabetes-related Medications (n = 11,800)
- Presence of Abnormal Lab Results (n = 18,833)

A comparison of phenotype definitions for diabetes mellitus
Challenges

- Multiple approaches to defining important characteristics and “standard” presentation:
  - Top-down (expert opinion)
  - Bottom-up (data-driven)

- Condition-specific and study specific components
- Timing of observations/measurements vs. inception of trial
- Observations and fragmentation of care

- Need clearly defined objective
- Need to engage potential users and stakeholders to ensure uptake / future endorsement or adoption

- Connection to pragmatic trials implies EHR as source data
  (→ more stakeholders)
Opportunities for the Collaboratory

• Researchers can define best practices in data collection and use
• High-visibility effort – ideal to build (and endorse) a case for standards
• Our members can be a conduit to health care organizations
• We are uniquely focused on getting data from EHRs (in contrast to de novo data collection)
Question: What characteristics should be on Table 1?

- Demographics
  - age, sex, race, ethnicity (federal standards)

- Relevant clinical or behavior/exposure data
  - pan-disease
  - disease-specific
Pan-Disease Features (top half)

- Age
- Gender
- Race
- Ethnicity
- SES
- Height
- Weight
- Blood Pressure
- Insurance Status (to infer access to care)
Condition-Specific (Bottom half)

- MANDATORY:
  - Study specific relevant co-morbidities
  - Study specific medications
  - Study specific labs
  - Study specific non-medication interventions
  - Cohort ID variables (baseline)

- CAD
- HTN
- Diabetes
- Hyperlipidemia
- CKD
- Anemia
- CHF

- COPD
- Asthma
- PVD
- PUD
- CVA
- Tumor, Leukemia, Lymphoma
- AIDS
- Atrial Arrythmeia
- Dementia
- Connective Tissue Disease
- Cohort identification variables (baseline)

- Co-morbidities - options:
  - Charlson Index
  - Top 10 comorbidities by frequency