MINIMIZING VARIABILITY IN TRIALS THAT USE CLINICIAN-ASSESSED OUTCOME MEASURES

Harry Barnett
Managing Partner
Clinical Assistance Programs, LLC
What I Will Discuss Today

1. Introduce and roughly define clinician-assessed outcome measures
2. Explain why variability is a problem
3. Introduce and describe a process for minimizing variability - Outcomes Training
4. Offer a case study with findings
Clinical Trial Outcome Measures

1. Surrogate markers of disease
2. Imaging or histological endpoints
3. Clinician (and subject)-dependent assessments
Outcome Measures:
Surrogate Markers of Disease

- Typically laboratory derived
- Quantitative
- Easily standardized (centralized lab)
- Examples: neutrophil count, PSA, cytokine levels
Outcome Measures:
Imaging or Histological Endpoints

- Based on established measurements of accuracy
- Require standardization of techniques and criteria
- Can centralize reading and interpretation
- Examples: tumor response measured by MRI, cytology, radiologic endpoints, etc.
Outcome Measures:
Clinician-Dependent Assessments

- Meaningful to patient status
- Often based on validated scales
- Applicable to a wide-range of indications
- Subjective and inconsistently applied
- Examples: neurological testing, rash or wound severity, arthritis range of motion, depression, xerostomia, visual acuity, etc.
Outcome Measures: Clinician-Dependent Assessments

Increasing regulatory focus on functional/QOL outcome measures which are clinician-/subject-based

- Biochemical, physiologic and other effects must be accompanied by improvement in function or quality of life
- Examples: emphysema, spinal cord injury, arthritis
Outcome Measure Accuracy

- High
- Moderate
- Low

- Surrogate Endpoints
- Imaging
- Clinician or Subject Based
Factors that Contribute to Lack of Accuracy for Clinician-Dependent Outcome Measures

- Many outcome measures are not designed for clinical trials
- Protocol wiggle-room
- Subjective interpretations (how red is red?)
- Regional variability
- Inconsistencies using multiple scoring systems
- Multiple sites and assessors
- Assessor/site arrogance (“I/we know best”)
- Standard outcome-related source documentation is not designed for clinical trials
Specialized Outcomes Training

What is it?
Why is it necessary?
“You ask for miracles. I give you…. the FDA”
The FDA maintains that comprehensive, consistently applied training is necessary to standardize trial conduct.
The Traditional Standard:
Investigator Meeting
Reliance on Outcomes Training at Investigator Meetings is a Failed Strategy

- Multiple topics discussed in compressed format
- Attendees at the IM are often not the people who will perform the outcome assessments
- Changes in the PI (principal investigator) or other site personnel
- IMs often occur well in advance of site activation and first patient accrual – the learning curve plummets
- Training typically occurs at the end of the meeting when people are less attentive
The Starbucks Approach
The Endpoint: Café Latte
Café Latte Recipes on Google

236,000!!!
Consistency and Uniformity
How Do They Do It?

- Select a clear, measurable, meaningful endpoint
- Define and standardize ingredients, utensils, and appliances
- Rigorous training on how to perform functions
- Constant QA and feedback
The Starbucks Approach for Clinical Trials

- Optimize study design/endpoint selection
- Standardize essential tools and equipment
- Onsite Assessor screening, training, and competency assessment
- Develop clinician-friendly source data capture instruments — source document worksheets
- Real-time data review and analysis — clinical reality check
Ideal Clinical Endpoint

- Accurately reflects severity and course of objective and subjective clinical changes
- Easy to teach and use
- Does not require complex measurements
- Sensitive enough to discriminate treatment efficacy
- Clinically meaningful and easily interpretable endpoints for clinicians, patients, sponsors, and FDA
- Balances regulatory/medical/business interests
Standardizing Essential Tools and Equipment

- Consider everything
- Be proactive
- Regional variability
Onsite Assessor Training - Trainers

- Trainers must be clinically qualified, credible and highly respected
- Trainers must be trained on each study protocol
- Trainers must remember that the training is all about the study data quality
Onsite Assessor Training

- Important to standardize the assessment methodology and grading criteria
- Assessors should be trained to assess using the same technique, same standards, same equipment, same order, same time frame, same source documents, etc.
- Training without competency evaluation and continuous feedback is of reduced value
Collecting the Endpoint: Source Worksheets

- Protocol-specific
- Provide sites with source documentation for the endpoint assessments
- Enable the tracking of subjects throughout the study
- Assist in ensuring that the assessments occur in the proper order (e.g., patient-reported then examination)
- Minimizes data collection and calculation errors
**Neurological Level**

R  
Sensory  
Motor  
L  

**Single Neurological Level**

**SENSORY CLINICAL EXAM**

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>NT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Impaired</td>
<td>Normal</td>
<td>Not Testable</td>
</tr>
</tbody>
</table>

**Please note that the Right and Left sides are grouped! Record your light touch and pin prick scores carefully**

**Right Side**

- Light Touch
- Pin Prick

**Left Side**

- Light Touch
- Pin Prick
Data Quality Review: The Traditional Standard

- Collect data on traditional or electronic CRFs during entire study
- Transfer data to sponsor’s data management group
- Lock database at the end of study – often many years after study commencement
- Analyze data
- Discover problems
The New Paradigm: Real-Time Data Quality Review

Site Endpoint Data (24-48 hours)

Endpoint Data Quality Management (DQM) Team
Real Time Data Quality Review: Feedback Loop

Site Endpoint Data

DQM Team

Evaluated

Scale Concordance

Data Patterns

Compliance

Sponsor

Database

DQM Team
Benefits of Real-Time Data Quality Review

- Confirms findings during onsite training and identifies deficiencies
- Provides valuable, independent, expert analysis of outcomes data during study
- Identifies systemic problems, toxicities, dosing compliance, formulation tolerability, etc.
- Permits data modeling and trend assessment
- Facilitate the collection of quality data specifically related to the efficacy endpoints
- Provide continuous data review and communication with the sponsor and study sites
Real-time Data Quality Review: The Family Oral Care Analogy
Case Study

Amgen’s Phase 3 Study of Kepivance® for Oral Mucositis
Oral Mucositis: Worst Complication of Ablative Chemotherapy

Close to 450,000 Patients Per Year Suffer from Mucositis During Cancer Therapy

- Stem cell transplant and radiation +/- chemotherapy for solid tumor (head and neck cancer, non-small-cell lung cancer) patients have the highest risk for severe mucositis
  - Mild, moderate, and severe mucositis can have serious clinical and economic consequences

Source: Mattson, Jack. Database 2003; NCI; Note: 400,000 patients in the US; CRC = colorectal cancer; NHL = Non-Hodgkin’s lymphoma
For Every 55 Patients with Grade 3-4 Mucositis and Myelosuppression...

41 will develop infection ... and 5 will die.

Elting, et al; Cancer 2003
Kepivance (KGF) was being tested in cancer patients receiving autologous BMT for ability to treat or prevent OM.

Previous studies had been confounded by inter-observer and inter-site variability.

Previous studies had operational issues that went uncorrected until nearly the end of the study.

No successful Phase III in the indication (many failures).
Case Study – Kepivance®

Actions

- Standardize method of examinations and scoring including use of source worksheets
- On-site training required at time of site initiation
- Continuous data review for abnormal trends and consistency
- Early recognition of, and intervention to address, site/investigator issues
- Immediate feedback provided to the sites
- Real-time inquiries fielded regarding study assessments
- Refresher training provided throughout the study
Accuracy Comparison

- Overall accuracy with IM training alone: 62.87%
- Overall accuracy with on-site training: 87.95%
Accuracy: Trained and Untrained

Accuracy

Study quarter

Accuracy

Trained

Untrained
Recap: Learning Objectives

- Selecting outcome measures in the design of clinical studies that will help get drugs to market faster
- Minimizing variability in clinical research involving the assessment of subjective clinical outcomes
- Improving the accuracy and consistency of outcomes data during a clinical study
THANKS AND QUESTIONS

THANKS TO:

- Steven T. Sonis, DMD, DMSc
- Charlotte Harrison