An Approach to Data Management and Evaluation for Evidence-Based Practice Projects

Martha Sylvia, PhD, MBA, RN
Mary Terhaar, DNSc, RN

STTI 23rd International Nursing Research Congress, Brisbane Australia, August 2012
Objectives

Within the context of the Doctorate of Nursing Practice:

- To describe effective approaches for managing and analyzing data for evidence-based projects
- To improve the quality of clinical data management for evidence-based projects.
“The focus for (DNP) faculty and students should be on the translation of evidence to improve the quality of care and patient outcomes”

(AACN, 2006)

Demonstrating “improvement” implies an understanding of:

- How to **define** improvement
- How to **measure** improvement
- How to **analyze** data for improvement
- How to **demonstrate** improvement
Clinical Data Management (CDM)

Our Definition:

“The process of planning, designing, collecting, cleansing, manipulating, analyzing, and reporting data generated in the assessment, development, delivery, and evaluation of health-related interventions, products, and services.”
Phases of CDM

1. Planning
2. Data Collection
3. Data Cleansing
4. Data Manipulation
5. Exploratory Analysis
6. Outcomes Analysis
7. Reporting
Practice Example of Process: Guided Care Pilot (GCP)

- Nurse-led, patient centered, comprehensive evidence-based project incorporating 7 successful innovations in chronic care:
  - Disease management
  - Case management
  - Self management
  - Geriatric evaluation and management
  - Transitional care
  - Lifestyle modification
  - Caregiver education and support

Planning

- What is your evidence-based project question?
- What is the design of your analysis?
  - Unit of analysis:
    - Groups/events
  - Eligibility criteria
  - Description of “intervention” and assignment of intervention group
  - Aims/outcomes/measures
  - Descriptive variables
  - Independent (IV) and dependent variables (DV)
  - Statistical tests/models
  - Power
Capstone Project Purpose
To Determine if there is a difference in the trend of costs between guided care and usual care enrollees

Define and Describe Population
2 groups of patients >=65, highly morbid, community dwelling, of 4 primary care providers (randomized) in same office

Aim: The GC group will have less costs than the comparison group

Outcomes: All costs of health care services

Measures: Mean costs over 6 months by group

Calculation: Sum of costs by group/total number of patients by group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Data Source</th>
<th>Possible Range of Values</th>
<th>Level of Measurement</th>
<th>Statistical Test</th>
</tr>
</thead>
</table>

Data Collection

◦ New data sources, some considerations:
  • Procedures for survey administration or other newly created data fields
  • Validation of conditions under which data was collected
  • Quality checks during collection

◦ Existing data sources, some considerations:
  • Conditions under which the data was collected
  • Definitions of fields
  • Request for data from other sources: details, details…
  • Information that can be calculated/derived from existing data

◦ Planning for final data structure:
  • Systems for data entry/import
  • Unique identifiers for unit of measurement
  • Rows and columns
  • Longitudinal vs. cross sectional
Data Cleansing

◦ Importing data into SPSS or other similar software:
  • Variable and value definitions, labels, etc.

◦ Running descriptive statistics on each variable looking for:
  • Missing values: setting rules
  • Text in number fields and vice versa
  • Erroneous values
  • Values outside of set range of expected
  • Looking for duplicate cases
  • Combinations of values that should not occur
Data Manipulation

- Creating final analysis data set
  - Merging/aggregating files
  - File restructuring
    - Transposing, cases to variables and vice versa
  - Transforming values
    - Calculating new values using functions
    - Recoding values
    - Banding values
    - Manipulating character values
    - Date and time values
  - Creating a data dictionary
GCP: Data Collection/Cleansing/Manipulation

**Administrative Enrollment Files**
- Gender
- Age
- Ethnicity
- Time enrolled in plan

**Administrative Claims Files**
- Medical Payments
- Pharmacy Payments
- Diagnoses/Services for morbidity determination

**Clinical Management Files**
- Time in program
- Dose of program
- Clinical indicators

**Analysis File**

Exploratory Data Analysis

- Explore and describe distribution of independent and dependent variables
- Describe groups/events
- Explore differences in “other factors” between groups/events
- Determine confounding of relationship between independent and dependent variables
<table>
<thead>
<tr>
<th>Demographics</th>
<th>GC (n=63)</th>
<th>UC (n=65)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>76.1 (6.15)</td>
<td>75.8 (6.53)</td>
</tr>
<tr>
<td>% Female</td>
<td>60.3%</td>
<td>47.7%</td>
</tr>
<tr>
<td>ACG-PM*</td>
<td>0.34 (0.22)</td>
<td>0.20 (0.14)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Health Status**</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># Chronic Conditions (max = 9)</td>
<td>2.95 (1.54)</td>
<td>2.85 (1.31)</td>
</tr>
<tr>
<td>Ischemic Heart Disease</td>
<td>52.2%</td>
<td>49.2%</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>31.7%</td>
<td>21.5%</td>
</tr>
<tr>
<td>Hypertension</td>
<td>88.9%</td>
<td>86.1%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>30.2%</td>
<td>20.0%</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>49.2%</td>
<td>46.1%</td>
</tr>
<tr>
<td>Parkinson’s Disease</td>
<td>1.6%</td>
<td>7.7%</td>
</tr>
<tr>
<td>Dementia</td>
<td>7.9%</td>
<td>13.8%</td>
</tr>
<tr>
<td>Depression</td>
<td>12.7%</td>
<td>18.5%</td>
</tr>
<tr>
<td>COPD</td>
<td>20.6%</td>
<td>21.5%</td>
</tr>
</tbody>
</table>

*Statistically Significant p<0.05
**Expanded Diagnostic Categories (EDCs) from ACG methodology used to define disease categories

GCP: Exploratory Data Analysis

ACG—PM Distribution by GC Assigned and UC Assigned Groups

Percent

GC Assigned

Percent

UC Assigned

ACG—PM
GCP: Exploratory Data Analysis

Costs per Member for 6-month Period
Unadjusted

Outcomes Data Analysis

- Final determination of statistical tests/models
- Statistical testing of outcomes measures
- Statistical testing of outcomes measures adjusting for confounding
Reporting and Presentation

- Relevance and importance to stakeholders
  - Common methods used to display certain representations of data
    - Tables, graph types, flow charts, etc.
  - Summarizing:
    - Tables, graphs, diagrams
    - Written and oral presentation of findings
GCP Outcomes Data Analysis and Reporting

GC vs UC: Costs at ACG-PM Cut Points
Adjusted for Age, Gender

Insurance Expenditures

$0
$2,000
$4,000
$6,000
$8,000
$10,000
$12,000
$14,000
$16,000
$18,000
$20,000

ACG-PM score

Usual Care (n=65)
Guided Care (n=62)

Summary

- Scholarly DNP projects using evidence-based practice frameworks require strong data management skills for management and evaluation.
- This clinical data management process provides a methodical and rigorous approach to meet this challenge.
Questions?
Contact Information

Martha Sylvia, PhD, MBA, RN
Assistant Professor
Johns Hopkins University School of Nursing
525 N. Wolfe St. Room 456
Baltimore, MD 21205-2110
443-287-8529
msylvia@jhmi.edu