Strategies for Addressing Data Collection Challenges in a Complex Community-Based Health Evaluation

Gina Cardazone
Landry Fukunaga
Christy Nishita

September 11, 2009
Center on Disability Studies
University of Hawaii at Manoa

Purpose of this Presentation

This presentation will discuss strategies for managing data collection challenges, and implications for analyses, conclusions, and replication.

- Challenges
- Strategies
- Lessons Learned

Project Overview

Hawaii Demonstration to Maintain Independence and Employment (HI-DMIE):

- Funding: $$$ CMS
- Partners: UH, HBHC, HI-DHS
- Objective:
  - Can a program of medical assistance and other supports forestall or prevent the loss of an individual's employment and independence due to potentially disabling complications arising from their diabetes
  - Independent Evaluation

Participants

<table>
<thead>
<tr>
<th>Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>H 1: Intervention group participants will work more hours per week</td>
</tr>
<tr>
<td>H 2: Intervention group participants will significantly increase their health status</td>
</tr>
<tr>
<td>H 3: Intervention group participants will remain independent of SSDI or SSI</td>
</tr>
</tbody>
</table>

Random Assignment:
Treatment : Control ratio of 2 : 1

Project Overview

Hypotheses

Eligibility at baseline:
- Between 18 – 62 years
- Hawaii resident
- Diagnosis of diabetes or HbA1c > 6.5
- Work > 40 hours/month
- Minimum wage

Control Participants
- "Business as Usual"
- No intervention
- Data will be compared with Treatment Group
- $$$ Compensation

Treatment Participants
- Pharmacist
- Life Coach
- Other support services:
  - CDE
  - Nutritionist
  - Fitness Membership
  - Support Groups
  - Diabetes Resources
  - Laptop and internet access
  - Medical copayments related to diabetes

Project Overview

Participants

<table>
<thead>
<tr>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Participants</td>
</tr>
<tr>
<td>Treatment Participants</td>
</tr>
</tbody>
</table>

Activity
- Health Assessment
- Monthly Work Calendar
- Work Productivity Survey
- 6-month Assessment Demographics
- Satisfaction Survey

Timeline
- Months: 0, 6, 12, 18
- Every month
- Months: 0, 3, 6, 9, 12, 15, 18
- Months: 0, 6, 12, 18
- Months: 6, 12

Description
- Physician reported HbA1c, Cholesterol, triglycerides, BMI, blood pressure
- Self-reported daily hours worked
- Work Productivity and Activity Impairment Diabetes Specific: Self-reported ability to work and perform regular activities
- Self reported: Participation in government assistance, Quality of Life, Health and Well Being, Diabetes Empowerment
- Self reported: Satisfaction, Effectiveness, Treatment participants only

Participant Data Requirements
Factors Affecting Data Collection

Participant data collection may be affected by:

- Lack of time and logistical barriers
- Expenses related to participation
- Accessibility of locations for implementation and/or data collection
- Residence instability
- Participant psychosocial issues such as self-efficacy, distress, or readiness to change
- Presence or absence of timely incentive payments
- Participant attitudes toward the scientific and medical community
- Complexity and stringency of protocol
- Dislike of randomization
- Physician attitudes toward intervention

- Yancey, Ortega & Kumanyika, 2006
- Mills et al., 2006

Participant Data Collection Challenges

<table>
<thead>
<tr>
<th>Issue</th>
<th>Influences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disenrollment</td>
<td>• Participant motivation</td>
</tr>
<tr>
<td></td>
<td>• External life events</td>
</tr>
<tr>
<td>Missing data</td>
<td>• Participant motivation</td>
</tr>
<tr>
<td></td>
<td>• Number of required items</td>
</tr>
<tr>
<td></td>
<td>• Overall demands of project on participant</td>
</tr>
<tr>
<td></td>
<td>• Physician data refusal</td>
</tr>
<tr>
<td>Completeness</td>
<td>• Participant motivation</td>
</tr>
<tr>
<td></td>
<td>• Attention to detail</td>
</tr>
<tr>
<td>Timeliness</td>
<td>• Awareness of deadlines</td>
</tr>
<tr>
<td></td>
<td>• Participant and physician diligence</td>
</tr>
<tr>
<td></td>
<td>• Health assessment deadlines incongruent with</td>
</tr>
<tr>
<td></td>
<td>typical physician appointments</td>
</tr>
<tr>
<td>Accuracy</td>
<td>• Social desirability bias</td>
</tr>
<tr>
<td></td>
<td>• Recall bias and reliability</td>
</tr>
</tbody>
</table>

Efforts to Improve Data Collection

- Contingent on submission of monthly calendars, quarterly surveys & semi-annual health assessments
- Reminders prior to deadlines
- Follow-up re: missing items
- Follow-up re: incomplete items

Current Status of Data Collection

<table>
<thead>
<tr>
<th>Activity</th>
<th>Timeline</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session logs</td>
<td>Each</td>
<td>Pharmacist and Life Coach self reported:</td>
</tr>
<tr>
<td></td>
<td>session</td>
<td>• Session duration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Focus of meeting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Referrals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reflections</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Meeting format (life coach only)</td>
</tr>
<tr>
<td>Session recordings</td>
<td>Each</td>
<td>Digital audio recording of entire session</td>
</tr>
<tr>
<td></td>
<td>session</td>
<td></td>
</tr>
<tr>
<td>Project specific</td>
<td>Each</td>
<td>Pharmacist:</td>
</tr>
<tr>
<td>requirements</td>
<td>session</td>
<td>• Quality Assessment Prescription (QARX) or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Digital Outcomes Communication System (DOCS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Life Coaches:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Online Coaching tool</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Case notes/session summaries/participant messages</td>
</tr>
</tbody>
</table>

Treating Staff Data Requirements

Activity: Session logs
Timeline: Each session
Description: Pharmacist and Life Coach self reported:
- Session duration
- Focus of meeting
- Referrals
- Reflections
- Meeting format (life coach only)

Activity: Session recordings
Timeline: Each session
Description: Digital audio recording of entire session

Activity: Project specific requirements
Timeline: Each session
Description: Pharmacist:
- Database tools
- Quality Assessment Prescription (QARX) or Digital Outcomes Communication System (DOCS)
Life Coaches:
- Online Coaching tool
- Case notes/session summaries/participant messages
Factors Affecting Data Collection

Data collection by treatment staff in community settings:
May be adversely affected by the following:
- Research needs are secondary to treatment/clinical needs
- Insufficient training, monitoring, or administrative support
- Lack of motivation, incentives, or recognition
- Ethical issues (confidentiality, consent)
- Clinical features (research skills, beliefs about merit of research)
- Political factors (collaboration, competing research agendas)
- Systemic barriers (caseload, tasks, $)

Boyd et al., 2007
- Butler, Little, & Grimard, 2009

May be reluctant to engage in data collection due to the following:
- Ambivalence
- Competing work demands
- Questionable evidence supporting outcome measures
- Fear about how data will be used

Meehan et al., 2006

Treatment Staff Data Challenges

<table>
<thead>
<tr>
<th>Issues</th>
<th>Possible influences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Refusal</td>
<td>Political/ethical/clinical issues</td>
</tr>
<tr>
<td></td>
<td>Fear of how data will be used</td>
</tr>
<tr>
<td>Missing Data</td>
<td>Other job responsibilities/double reporting</td>
</tr>
<tr>
<td></td>
<td>Recording errors/downloads</td>
</tr>
<tr>
<td></td>
<td>Change in database tool</td>
</tr>
<tr>
<td>Incomplete Data</td>
<td>Motivation/incentive</td>
</tr>
<tr>
<td></td>
<td>Training/monitoring</td>
</tr>
<tr>
<td></td>
<td>Time constraints</td>
</tr>
<tr>
<td></td>
<td>Recording/hardware errors</td>
</tr>
<tr>
<td>Inaccurate Data</td>
<td>Low motivation/ambivalence</td>
</tr>
<tr>
<td></td>
<td>Recall bias</td>
</tr>
<tr>
<td></td>
<td>3rd party data entry</td>
</tr>
</tbody>
</table>

Efforts to Improve Data Collection

- P.I. letter on use and intent of session recordings
- Supervisor mediation
- Project IT data collection
- Pharmacists – Month 3
- Life Coaches – Month 6
- Follow-up re: missing items
- Follow-up re: incomplete items
- Payment contingent on submission of data
- Only first page of session logs
- No session recordings
- 3rd party data entry (supervisor)

Current Status of Data Collection

Pharmacist Missing Data:

<table>
<thead>
<tr>
<th>No log</th>
<th>Before</th>
<th>163 (44%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No log</td>
<td>125 (34%)</td>
<td></td>
</tr>
<tr>
<td>After</td>
<td>16 (4%)</td>
<td>163 (44%)</td>
</tr>
</tbody>
</table>

Missing Data (Specific Items, Complete Subjects)

1. A Potential Source of Bias
2. Impacts the Ability to Interpret Results
3. Influences the Certainty With Which Conclusions are Drawn

Implications for Analyses

Analysis Strategy:
“As Randomized... So Analyzed”

Use All Randomized Participants -Regardless of Withdrawal, Non-Compliance, etc.
- Preserve Sample Size, Power
- Deviation can Contaminate Treatment Comparison
- Compliant Participants Have Better Outcomes, Regardless of Group Assignment

- LaValley (2003)
Randomized (n = 190)

Life Coaching (n = 128)

Complete Data (65.6%)
Missing Data (34.4%)
Withdraw
• Moved Out of State (n = 2)
• Refused (n = 9)
• Unable to Contact (n = 6)

Business As Usual (n = 62)

Complete Data (72.6%)
Missing Data (27.4%)
Withdraw
• Moved Out of State (n = 1)
• Refused (n = 1)
• Unable to Contact (n = 1)

Analysis Strategies

Multiple Imputation (MI)

• A Range of Plausible Values are Generated that Approximate Each Missing Value
• Allows for the Inclusion of All Subjects in the Analysis
• Less Bias, More Statistical Efficiency, Straightforward

Multiple Imputation Process

Original Dataset

Imputed Datasets

Multiple imputation to eliminate missing values

Analysis

Analysis

Analysis

Analysis

Analysis

Merging of Results from Individual Datasets

Final Results

Missing Data in Health Interventions

• Goal to Minimize Bias
• Method for Handling Data Should Provide Conservative Estimate of the Treatment Effect
• Pragmatic Approach- Determine Utility of Intervention for Practice and Dissemination
• Need to Make Appropriate Conclusions

Lessons Learned

1. Pre-specify in protocol strategies to minimize amount of missing data and how missing data will be handled
2. Pre-service training with community partners on merits of research and data collection
3. Continued data monitoring, feedback, and retraining
4. Clearly defined contractual agreements regarding data collection, payment of incentives
5. Staff resources to follow-up on missing data

References


