May 20th, 2:30 PM

Data Acquisition, Data Management and Subject Tracking in an RCT: Promoting Breast Cancer Screening in Non-Adherent Women

Roger S. Luckmann
University of Massachusetts Medical School

Follow this and additional works at: https://escholarship.umassmed.edu/cts_retreat

Part of the Databases and Information Systems Commons, Health Information Technology Commons, and the Translational Medical Research Commons

Creative Commons Attribution-Noncommercial-Share Alike 3.0 License

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.
Data Acquisition, Data Management and Subject Tracking in an RCT: *Promoting Breast Cancer Screening in Non-Adherent Women*

**PRINCIPAL INVESTIGATORS**
Mary Costanza MD  
Roger Luckmann MD MPH

**RESEARCH TEAM**
Mary Jo White  MS MPH  
Caroline Cranos MPH  
Chris Foley MS  
Susan Sama  ScD  
Devi Sundersen  
Sally Gore

**CO-INVESTIGATORS**
Bruce Barton PhD  
Robert Yood MD  
Dan Roswig MD
I have no actual or potential conflict of interest in relation to this program/presentation
5-year study funded by National Cancer Institute

Dynamic study population: Women age 51-84 (Later age 40-84)

- Core eligibility: Fallon Community Health Plan (FCHP) member >=18 months with a Fallon (Reliant) Clinic PCP
  - Later added 3 other health plans and reduced time in plan to 13 months
- Baseline n=23,000
- New subjects added as they become eligible (Health plan membership >=18 months with Fallon/Reliant PCP)
- Subjects excluded when no longer eligible, but may return if core eligibility regained
Main objective:

- Compare the effectiveness of 3 different interventions arms in promoting adherence to screening mammography over 4-years.

- When first meeting core eligibility requirements, women are randomized to three study arms:
  
  1. Reminder letter (control, usual care)
  2. Reminder letter followed by reminder/scheduling call to nonresponders
  3. Reminder letter and educational booklet followed by an enhanced tailored telephone counseling call to nonresponders.

The call includes:

- Reminding
- Tailored review of information in educational booklet
- Motivational interviewing as needed
- Scheduling
Establish Study Eligibility

1. Potential study subjects meeting core eligibility criteria
2. Names to PCPs every 3 months
3. Approved by PCP or excluded
4. Introductory letter
5. Passive consent or opt out
6. Eligible for intervention
Establish Study Eligibility

Potential study Subjects Meeting Core Eligibility Criteria

Names to PCPs Every 3 Months

Approved by PCP or Excluded

Introductory Letter

Passive Consent or Opt Out

Eligible for Intervention

Intervention Delivery

>=18 Months from Last Mammogram or Study Entry

Reminder Letter

- Schedules a Mammogram
- No Response

- Reminder Call
- Counseling Call
- No Further Intervention

- If No Mammogram, Recycle in 12 months
- If Mammogram Received, Recycle in 18 Months
Data Flow: Core Eligibility and Mammogram Tracking

- EHR data flows into Clarity data repository
- Automated daily query of Clarity: Data on women ≥40 loads into Staging Database
- Eligibility flag is set (Yes/no) based on core eligibility criteria
- Tracking Database updated:
  - Newly eligible subjects added
  - Eligibility flag, contact info, date of last and next scheduled mammogram updated
Tracking and Application Support Functions of the Tracking Database

• Queries: Identify subjects for introductory or reminder letter, call, or PCP approval
• Letters/lists printed
• Women allocated to call queues
• Data from CATI system and contact logs flow to Tracking Database
• PCP approvals and some “Opt Outs” manually entered
Data Flow to Analytic Datasets

- Monthly snapshots merged as needed for analysis
- Data from fields written to >1 time in a month are lost
- Claims from multiple years extracted and merged with data from Tracking Database
The System looked like this…

Databases and Data Flow
But it FELT like this...
Challenges and Issues

- Inconsistent field names and terminology (e.g. 4 types of “eligibility”)
- Losing/Regaining eligibility (Overwriting fields and loss of history) and specifying eligibility in staging dataset
- 3 data sources for core eligibility, sometimes in conflict (Clarity, FCHP claims, CATI)
- 1 – 4 repeats (waves) of interventions
- Repeated snapshots of tracking data is inefficient and ineffective way to create an analytic dataset
Lessons Learned

- Use consistent field names that indicate data source when possible
- Identify and resolve any potential conflicts in the design phase
- Develop a flow chart of all eligibility processes and waves of intervention when designing the system
- Maintain control of all eligibility rules and of updating of the tracking database
- Do not overwrite values in any variable
- Date and time stamp all entries
- Specify analytic dataset as subset of tracking database in the design phase
Expertise Needed for Building Effective Data Acquisition/Management and Tracking Systems

Expertise in:

- Source data content and organization
- Source data extraction, transfer, and loading (ETL)
- Database design
- System design (Data flow, automated queries, interfaces, hardware)
- Data management (Field names and formats, record structure, analytic dataset construction)
- Data analysis (Biostatistics)
- Software development for custom applications
- Facilitation of process of specifying all system requirements