Shared Medical Appointments for Chronic Medical Conditions: A (Traditional) Systematic Review

David Edelman MD, Jennifer R. McDuffie PhD, Eugene Oddone MD, Jennifer Gierisch PhD, John W. Williams Jr. MD.
Shared Medical Appointments (SMAs)

• Subset of Group Visits
• Groups of patients with a defining chronic condition or other health care state meet over time for comprehensive care
• Involve both self-management training and medication management
SMAs-- Structure

• 60-120 minutes, usually 1-3 months apart
• Usually both a prescribing provider and a trained educator/facilitator
• Interactive education, often with techniques such as motivational interviewing; goal is patient activation
• Prescribing provider does med changes, often in one-on-one breakouts
Studies of SMAs

- Some studies in frail elderly
- Most in a single disease, most commonly diabetes
- Wide variability in:
  - Setting and patients
  - Intervention approaches, including staffing
  - Chosen outcomes to measure (ie, behavioral, clinical, cost/utilization)
Objectives

• Summarize the effects of SMA on:
  – Patient outcomes
  – Staff outcomes
  – Economic outcomes

• Evaluate whether these effects vary by clinical condition or specific intervention components.
Outline of Methods

• Topic development
  – Key questions
  – Protocol
• Systematic searches of the literature
• Study selection via eligibility criteria
  – Screening
  – Full text review
• Data abstraction and quality assessment
• Data synthesis and report generation
• Peer review
Key Question 1

• For adults with chronic medical conditions, do shared medical appointments (SMAs), compared with usual care, improve the following:
  – Patient and staff experience?
  – Treatment adherence?
  – Quality process measures?
  – Biophysical markers (laboratory or physiological markers of health status such as HbA1c and blood pressure)?
  – Symptom severity and functional status?
  – Utilization of medical resources or health care costs?
Key Questions 2-3

• **Key Question 2.** For adults with chronic medical conditions, do the effects of SMAs vary by patient characteristics such as specific chronic medical conditions and severity of disease?

• **Key Question 3.** Is the intensity of the intervention or the components used by SMAs associated with intervention effects?
Protocol

• The protocol provides the analytic framework for the report and outlines a priori how each remaining step of the procedure is to be conducted

• The protocol is reviewed and vetted by both internal and external experts as well as the major stakeholders for the report
Literature Search Strategy

• Databases
  – MEDLINE® (via PubMed®), Embase®, CINAHL®, PsychINFO® and Web of Science®

• Search terms
  – Consult master librarian
  – Key words and MeSH Analyzer

• Supplemental searches
  – Bibliographies of exemplary articles
  – ClinicalTrials.gov
Adults with:
- Asthma
- CAD
- CHF
- COPD
- Diabetes
- High lipids
- HTN

**Modifiers**
- Patient
- Social support
- Health care system

**SMA model**
- Group size
- Components
- Team composition
- Rxing professional
- Visit frequency

**Usual care**
- Traditional office visit
- Other systems improvements

**Intermediate outcomes**
- Adherence
- Satisfaction
- QI Measures (A1c, BP, LDL)

**Adverse effects**

**Final outcomes**
- Symptoms
- Functional status
- HR QOL
- Health utilization
Study Selection Criteria

• Exclusion
  – Publication is NOT English language or not peer-reviewed
  – Population selected for substance abuse or is from inpatient setting

• Inclusion:
  – Based on model in previous slide
Inclusion (1)– Study Quality

• Quality of study
  – Study designs recommended by Cochrane EPOC Group
  – Trials, or observational studies with contemporaneous comparator
Inclusion (2)– Patients

• Adult with one or more of 7 chronic medical conditions of *a priori* interest
  – Asthma
  – CAD
  – CHF
  – COPD
  – *Diabetes*
  – High lipids
  – HTN

• Also reviewed– extant literature on older adults without a single unifying disease
Inclusion (3)–
Intervention and its context

• Setting
  – Outpatient primary care or specialty clinic/practice

• SMA Model
  – Intervention defined as ≥2 medical visits where ≥1 healthcare professional (includes prescribing clinician) cares for a patient group

• Comparator
  – Defined as usual care or other quality improvement strategy
Inclusion (4)– Outcomes

• One of following outcomes reported at ≥ 3 months:
  – Patient or staff experience
  – Adherence (treatment, medication or self-management)
  – Biophysical marker, (e.g., HbA1c, LDL, BP)
  – Symptom severity or functional status
  – Utilization of medical resources
Data Abstraction

- Extraction of pertinent information from each eligible article into a customized, uniform database in DistillerSR®
- Performed by 1st reviewer and independently over-read by a 2nd reviewer
- Disagreements are resolved by discussion and consensus or referral to a 3rd reviewer
Robustness Score

• Devised to attempt to describe the more potent elements of an SMA intervention
• Seven variables
  – Education session
    • Qualifications of leader
    • Based on theoretical framework?
  – Group composition
    • Closed membership?
    • Stable healthcare team?
  – Intervention Process
    • Individual breakout sessions?
    • Medication changes within visit?
    • Number and length of visits
• Potential score range = 0-9
Quality Assessment

• Elements rated for RCTs
  – Adequacy of randomization
  – Adequacy of allocation concealment
  – Comparability of groups at baseline
  – Blinding of subjects and/or investigators
  – Completeness of and differential loss to followup
  – Management of incomplete data
  – Validity of outcome measures
  – Potential conflicts of interest

• Elements rated for observational studies
  – Selection bias
  – Performance bias
  – Detection bias
  – Reporting bias

• Reference: Agency for Healthcare Research and Quality’s (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews
Data Synthesis

• Summary table of key outcomes
• Quantitative meta-analysis, if feasible
  – Dichotomous outcomes combined using RR; OR
  – Continuous outcomes combined using standardized mean difference and a random effects model
  – Tests for statistical heterogeneity (Q and $i^2$)
• Qualitative synthesis otherwise (e.g., too few studies or subgroup and sensitivity analyses)
• Assessment of publication bias
Strength of Evidence

• Assessment of four domains
  – Risk of bias
  – Consistency
  – Directness
  – Precision

• The strength of the evidence for the proposed answer to each key question is graded – high, moderate, low or insufficient

• Reference: AHRQ’s *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*
Literature Flow

Search results = 1104 references

Excluded = 1009 references
  Excluded at screening level

Retrieved for full-text review = 95 references

Excluded = 71 references
  Not peer-reviewed, or primary data = 34
  Not study population of interest = 6
  Not eligible study design = 17
  Comparator not of interest = 1
  Intervention doesn’t meet definition = 7
  No outcome of interest = 6

Included = 18 unique studies and 6 companion articles

KQ 1: 19 unique studies + 1 companion
KQ 2: 16 unique studies + 4 companions
KQ 3: 13 unique studies + 4 companions
### Study and subject characteristics

<table>
<thead>
<tr>
<th>Study Characteristic</th>
<th>Adults With Diabetes</th>
<th>Older Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>N studies (participants)</td>
<td>16 (3221)</td>
<td>3 (1851)</td>
</tr>
<tr>
<td>Mean age of sample: median (range)</td>
<td>60.8 (27 to 69.8)</td>
<td>74.1 (73.5 to 78.2)</td>
</tr>
<tr>
<td>Randomized controlled trial</td>
<td>13 (2921)</td>
<td>2 (615)</td>
</tr>
<tr>
<td>Observational</td>
<td>3 (300)</td>
<td>1 (1236)</td>
</tr>
<tr>
<td>Study quality: N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>6 (46%)</td>
<td>0</td>
</tr>
<tr>
<td>Fair</td>
<td>6 (46%)</td>
<td>2 (67%)</td>
</tr>
<tr>
<td>Poor</td>
<td>1 (8%)</td>
<td>1 (33%)</td>
</tr>
<tr>
<td>Single-site</td>
<td>14 (2106)</td>
<td>1 (321)</td>
</tr>
<tr>
<td>Multisite</td>
<td>2 (1115)</td>
<td>2 (1530)</td>
</tr>
<tr>
<td>Duration 6 to 12 months</td>
<td>4 (410)</td>
<td>0</td>
</tr>
<tr>
<td>Duration &gt;12 months</td>
<td>12 (2811)</td>
<td>3 (1851)</td>
</tr>
</tbody>
</table>
Effect on A1c: mean = -0.55%

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Mean</th>
<th>SD</th>
<th>Total</th>
<th>Weight</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.1 A1c Good Quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sadur 1999</td>
<td>8.18</td>
<td>2.1392</td>
<td>82</td>
<td>9.33</td>
<td>2.1392</td>
<td>74</td>
<td>7.4%</td>
<td>-1.15 [-1.82, -0.48]</td>
<td>1999</td>
</tr>
<tr>
<td>Clancy 2003</td>
<td>9.5</td>
<td>2.33</td>
<td>59</td>
<td>9.7</td>
<td>2.33</td>
<td>61</td>
<td>6.8%</td>
<td>-0.20 [-1.03, 0.63]</td>
<td>2003</td>
</tr>
<tr>
<td>Edelman 2010</td>
<td>8.3</td>
<td>1.8417</td>
<td>133</td>
<td>8.63</td>
<td>1.8417</td>
<td>106</td>
<td>8.1%</td>
<td>-0.33 [-0.80, 0.14]</td>
<td>2010</td>
</tr>
<tr>
<td>Trento 2010</td>
<td>7.3</td>
<td>0.9</td>
<td>315</td>
<td>8.8</td>
<td>1.2</td>
<td>266</td>
<td>8.8%</td>
<td>-1.50 [-1.68, -1.32]</td>
<td>2010</td>
</tr>
<tr>
<td>Taveira 2011</td>
<td>7.4</td>
<td>1.2</td>
<td>44</td>
<td>8.4</td>
<td>1.2</td>
<td>44</td>
<td>7.3%</td>
<td>-1.00 [-1.69, -0.31]</td>
<td>2011</td>
</tr>
<tr>
<td>Naik 2011</td>
<td>8.05</td>
<td>1.4</td>
<td>44</td>
<td>8.64</td>
<td>1.39</td>
<td>41</td>
<td>7.7%</td>
<td>-0.59 [-1.18, 0.00]</td>
<td>2011</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>677</strong></td>
<td><strong>592</strong></td>
<td><strong>45.9%</strong></td>
<td><strong>Mean Difference IV, Random, 95% CI</strong></td>
<td><strong>-0.83 [-1.36, -0.30]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.35; Chi² = 33.98, df = 5 (P < 0.00001); I² = 85%
Test for overall effect: Z = 3.06 (P = 0.002)

| **1.1.2 A1c Fair/Poor Quality** |
| Trento 2001       | 7.5  | 1.4  | 43    | 8.3  | 1.8  | 47    | 7.4%     | -0.80 [-1.46, -0.14] | 2001 |
| Wagner 2001       | 7.9  | 1.0001 | 278   | 7.9  | 1.0001 | 429   | 8.8%     | 0.00 [-0.15, 0.15] | 2001 |
| Trento 2005       | -0.38 | 1.2051 | 30    | -0.4 | 1.1605 | 28    | 7.6%     | 0.02 [-0.59, 0.63] | 2005 |
| Clancy 2007       | 9.1  | 2.1947 | 96    | 9    | 2.4666 | 90    | 7.4%     | 0.10 [-0.57, 0.77] | 2007 |
| Taveira 2010      | -0.9 | 1.6   | 58    | 0    | 1.5  | 51    | 7.7%     | -0.90 [-1.48, -0.32] | 2010 |
| Gutierrez 2011    | -1.19 | 1.66  | 50    | -0.67 | 2    | 53    | 7.2%     | -0.52 [-1.23, 0.19] | 2011 |
| Cohen 2011        | -0.41 | 1.1612 | 50    | -0.2 | 1.4274 | 49    | 7.9%     | -0.21 [-0.72, 0.30] | 2011 |
| **Subtotal (95% CI)** | **605** | **747** | **54.1%** | **Mean Difference IV, Random, 95% CI** | **-0.29 [-0.59, 0.01]** |

Heterogeneity: Tau² = 0.09; Chi² = 15.23, df = 6 (P = 0.02); I² = 61%
Test for overall effect: Z = 1.86 (P = 0.06)

**Total (95% CI)**

<table>
<thead>
<tr>
<th>Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
</table>

Heterogeneity: Tau² = 0.57; Chi² = 179.95, df = 12 (P < 0.00001); I² = 93%
Test for overall effect: Z = 2.43 (P = 0.01)
Test for subgroup differences: Chi² = 3.01, df = 1 (P = 0.08), I² = 66.7%

- Favors SMA
- Favors Usual Care
Effect on SBP: mean = -5.2 mmHg

### 1.4.1 Systolic Blood Pressure

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>SMA Mean [mmHg]</th>
<th>SMA SD [mmHg]</th>
<th>Usual Care Mean [mmHg]</th>
<th>Usual Care SD [mmHg]</th>
<th>Total</th>
<th>Weight</th>
<th>Mean Difference IV, Random, 95% CI [mmHg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen 2011 (1)</td>
<td>-9.19</td>
<td>20.2676</td>
<td>-0.8</td>
<td>16.746</td>
<td>50</td>
<td>8.8%</td>
<td>-8.39 [-15.71, -1.07]</td>
</tr>
<tr>
<td>Edelman 2010</td>
<td>139.2</td>
<td>21.5523</td>
<td>146.5</td>
<td>21.5523</td>
<td>133</td>
<td>15.6%</td>
<td>-7.30 [-12.80, -1.80]</td>
</tr>
<tr>
<td>Taveira 2010</td>
<td>-7.3</td>
<td>20.3</td>
<td>-1.7</td>
<td>19.6</td>
<td>58</td>
<td>8.4%</td>
<td>-5.60 [-13.10, 1.90]</td>
</tr>
<tr>
<td>Taveira 2011</td>
<td>123.4</td>
<td>12.3</td>
<td>127</td>
<td>17.3</td>
<td>44</td>
<td>12.0%</td>
<td>-3.60 [-9.87, 2.67]</td>
</tr>
<tr>
<td>Trento 2010</td>
<td>138.01</td>
<td>16.1</td>
<td>142.43</td>
<td>18.9929</td>
<td>295</td>
<td>55.1%</td>
<td>-4.42 [-7.35, -1.49]</td>
</tr>
</tbody>
</table>

Total (95% CI)

- **Mean Difference**: -5.22 [-7.40, -3.05]

Heterogeneity: Tau² = 0.00; Chi² = 1.82, df = 4 (P = 0.77); I² = 0%

Test for overall effect: Z = 4.71 (P < 0.00001)

Test for subgroup differences: Not applicable

(1) Cohen 2011 and Taveira 2010 is mean change
Effect on LDL-C: mean = -6.6 mg%

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>SMA Mean [mg/dl]</th>
<th>SMA SD [mg/dl]</th>
<th>Total</th>
<th>Usual Care Mean [mg/dl]</th>
<th>Usual Care SD [mg/dl]</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI [mg/dl]</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clancy 2003</td>
<td>107.6</td>
<td>31.2</td>
<td>59</td>
<td>116.2</td>
<td>31.2</td>
<td>61</td>
<td>19.0%</td>
<td>-8.60 [-19.77, 2.57]</td>
<td>2003</td>
</tr>
<tr>
<td>Taveira 2010</td>
<td>82.8</td>
<td>24.1</td>
<td>58</td>
<td>85.2</td>
<td>26.7</td>
<td>51</td>
<td>20.4%</td>
<td>-2.40 [-12.00, 7.20]</td>
<td>2010</td>
</tr>
<tr>
<td>Trento 2010</td>
<td>107.9</td>
<td>36.4</td>
<td>315</td>
<td>127.9</td>
<td>37.5</td>
<td>266</td>
<td>23.3%</td>
<td>-20.00 [-26.04, -13.96]</td>
<td>2010</td>
</tr>
<tr>
<td>Taveira 2011</td>
<td>92.5</td>
<td>24.3</td>
<td>44</td>
<td>93.9</td>
<td>30.6</td>
<td>44</td>
<td>18.6%</td>
<td>-1.40 [-12.95, 10.15]</td>
<td>2011</td>
</tr>
</tbody>
</table>

Total (95% CI) 526 471 100.0% -6.64 [-16.11, 2.82]

Heterogeneity: Tau² = 90.57; Chi² = 19.46, df = 4 (P = 0.0006); I² = 79%
Test for overall effect: Z = 1.38 (P = 0.17)
Test for subgroup differences: Not applicable
(1) Mean change
Other outcomes—patient-level

- Patient experience?
  - No effect— but only 2 studies measured
- Staff experience?
  - Completely unmeasured
- Treatment Adherence?
  - No effect— but only 3 studies measured, and no single behavior was assessed in more than 2 studies
- HRQoL?
  - 3 studies used a disease-specific measure, found a positive effect, not highest quality trials
  - 2 used a general measure, no effect
Other Outcomes—costs and utilization

• Utilization?
  – 5 studies, 4 of 5 with reduced admissions
  – Same 5 studies, variable results on ER visits

• Costs
  – 4 studies, mixed results on overall costs
Older Adults?

- 3 studies, 2 trials, one observational
- Lower study quality than diabetes studies
- All studies measured patient satisfaction but with different, non-validated measures
  - All showed satisfaction improvement
- No change in global health or function
- Both trials showed lower ER use and admissions, ER statistically significant in both
- Costs lower, but not significantly so, in both trials
Key Questions 2 & 3

• No study reported specific patient characteristics that led to better response to SMAs (Key Question 2)
  – We evaluated whether baseline A1c was associated with response; it was not

• No study reports specific intervention components, or intensity, associated with effects of SMAs (Key Question 3)
  – We evaluated whether robustness was associated with effect size; it was not
Evidence Synthesis found no data to assess:

- Cost-effectiveness
- Non-patient benefits, such as improved access or staff satisfaction
- Key elements to successful implementation, especially outside academic or vertically integrated systems
Lessons learned

• Precise, highly scientifically valid estimates of SMA efficacy
  – SMAs are pretty efficacious; effects on A1c and SBP close to those seen in drug trials
  – Effect sizes 0.5, 0.33, and 0.25 for SBP, A1c, and LDL-c respectively
  – Taken together, a marked improvement in risk of complication; would still be important risk reduction if half the efficacy were lost in translation
Are these findings generalizable?

- Use **PICOTS** framework
- **Population**—likely generalizable, populations were well demographically balanced
- **Intervention**—components VERY heterogeneous, remains quite possible that not all group strategies are effective.
- **Comparator**—Also heterogeneous, “usual care” poorly described in most studies
- **Outcome**—likely generalizable, biophysical outcomes are the generally agreed upon set
- **Timing**—likely generalizable, general agreement that 6 months improvement is important
  - But no studies of maintenance of improvement
- **Setting**—All studies in highly academic settings, not any “real-world” studies.
Where do we go from here?

• Other chronic illnesses
• Study designs that allow evaluation of particular components
• Multi-site implementation studies with good measurement of patient and staff impacts
  – Strongly consider mixed-methods studies
  – Carefully measure unintended consequences on the system
• Cost and cost-effectiveness analyses
Theory Driven, Context Dependent Studies of Shared Medical Appointments: A Realist Work in Progress

Susan Kirsh, MD, MPH and David Aron, MD, MS

with Kim Johnson, RN PhD, Katherine Jones, PhD, Brian Mittman, PhD, John Øvretveit, PhD, Laura Santurri, PhD, MPH, CPH, Lauren Stevenson, PhD

Louis Stokes Cleveland DVAMC and Case Western Reserve University
Cleveland, OH 44106

The views are those of the presenters and do not reflect those of VHA or any other agency/institution.
Objectives

• Discuss theory-driven context-dependent review (realist) and why needed
• Describe our experience
• Q&A
Main results: benefits HbA1c; some evidence for SBP; none for LDL

Conclusions: should be considered by clinicians as an effective, non-pharmacologic intervention that can have a positive impact on biologic markers such as HbA1c and SBP.

Implications for Practice: most powerful model includes clinician prescriber

Implications for Research: RCTs needed

What managers want to know is what works when and for whom.

In other words, context matters.
45 CFR 46.102(d) *Research* means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.
Scientists, including social scientists, are often dismissive of philosophy. Philosophy, it is said, is too abstract, too fussy and too taken up with its own problems to matter to real practice. With the issues discussed here (efficacy, evidence, RCTs and policy) I think just the opposite is the case. Bad practice, I maintain, is being recommended without intention and without sufficient notice in part because prissy issues that philosophy fusses about are being ignored, issues like what counts as a proper definition and whether an argument has been laid out with all the necessary premises.

What is This Thing Called ‘Efficacy’?
http://personal.lse.ac.uk/cartwrig/PapersOnEvidence/What%20is%20that%20thing%20called%20efficacy.%202018%20June%20edited%20for%20web%20page.pdf
## Philosophical differences: Positivism, Realism and Constructivism

<table>
<thead>
<tr>
<th></th>
<th>Positivism</th>
<th>Realism</th>
<th>Constructivism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ontology</strong></td>
<td>There is an objective reality, which exists independent of us.</td>
<td>Material &amp; social reality – we interact with reality.</td>
<td>Subjective reality – we ‘create’ reality</td>
</tr>
<tr>
<td><strong>Epistemology</strong></td>
<td>Truth and final knowledge exists.</td>
<td>No final truth or knowledge, but improvement in knowledge is possible.</td>
<td>No way to choose between interpretations. What we jointly believe is true.</td>
</tr>
<tr>
<td><strong>Causation</strong></td>
<td>Constant conjunction, linear causation. Programs cause outcomes.</td>
<td>Mechanisms operating differently in different contexts generate patterns of outcomes.</td>
<td>Co-constructed interpretations lead to actions and outcomes.</td>
</tr>
<tr>
<td><strong>Implications for evaluation</strong></td>
<td>Evaluators ‘tell facts’. Context factors should be eliminated: Randomised Control Trials/ Quasi-experimental methods.</td>
<td>Evaluators explain how and where programs generate outcomes. Mixed methods (qualitative and/or quantitative).</td>
<td>Evaluators describe stakeholder interpretations. Qualitative methods.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RCT</th>
<th>Theory Driven/Context Dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Internal validity-independent from other external changes.</td>
<td>External validity-external changes part of interventions and must be reported together with the results.</td>
</tr>
<tr>
<td>Intervention group is compared with similar control or an internal control is used.</td>
<td>Based on identification of relevant context and mechanisms, the intervention group can be compared with similar external or internal control. Differences between controls and intervention are reported.</td>
</tr>
<tr>
<td>Isolation of confounding factors</td>
<td>Isolation of confounding factors is not possible, outcome cannot be seen isolated from context and mechanisms</td>
</tr>
</tbody>
</table>

Realist Evaluation and Realist Synthesis

- **Real** – deals with the real world
- **Realist** – grounded in ‘scientific realism’
- **Realistic** – “The whole point is that it is a form of applied research, ... pursued to inform the thinking of policy makers, practitioners, program participants and public.”

Patricia Rogers, RMIT University

“What causes something to happen has nothing to do with the number of times we observe it happening” (Sayer, 2000 p. 14).

Not: “does it work or not?”

But rather, “what works, for whom, and in what circumstances?”

Justin Jagosh, Ph.D, Participatory Research at McGill (PRAM)
Methods, but...

1. identifying the review question
2. searching for primary studies (A search to track program theories and a search for primary studies)
3. quality appraisal (Assessment of relevance and rigour)
4. extracting the data (Annotation, Collation, Reportage)
5. synthesis
There is more than one way to conduct a “realist synthesis.”

Thus, on the continent the disease was pulled to pieces, bit by bit without the benefit of chemical determinations: in America it was put together by a group of men, almost no one of whom had ever looked a parathyroid cell face to face in a microscope. All of which proves that there are two ways of killing a cat (see Fig. 8).

Fig. 8. Two ways of killing a cat.
At the Massachusetts General Hospital we prefer “A.”

• Biographical Memoirs V.48 (1976) National Academy of Sciences (NAS)
What Realist Review Does

• Identifying mechanisms, the contexts in which they are (or are not) activated, and the outcomes to which they lead

• Categorizing and building these Context-Mechanism-Outcome clusters into Demi-Regularities
  – Not laws, but things that tend to happen

• Bringing to bear mid-range theories to help understand the patterns of these demi-regularities

• Ultimately building or testing a theoretical model of how a program works


Justin Jagosh, Ph.D , Participatory Research at McGill (PRAM)
Department of Family Medicine, McGill University, Montréal, Canada.
Understanding Mechanisms:

• Mechanism may be defined as:
  – “…underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest.”*

• Mechanism:
  – Are usually hidden
  – Sensitive to variations in context
  – Generate outcomes

• For social interventions, mechanism typically refer to a cognitive process or what ‘turns on’ in the mind of program participants to make them want to participate in the program

How do you do that?

• By identifying the basic logic (theory) behind programs under review;
• By configuring the contextual features and mechanisms which determine outcomes (C-M-O configuring) and comparing cases;
• By refining the theory that was originally identified, based on the CMO synthesis.


Modified from Justin Jagosh, Ph.D, Participatory Research at McGill (PRAM), Department of Family Medicine, McGill University, Montréal, Canada.
Identifying the theory:

For a realist synthesis of a single case, the underlying logic is understood as ‘program theory.’ Every program has a theory, whether it is obvious or not

- For a realist review synthesizing many cases, the underlying theory is considered “Middle-range”

- **Middle-Range Theory:** not abstract to the point of being disconnected from the actual on-the-ground realities of program planning and implementation, yet, not specific to the point of being relevant to only one type of program.

- Middle-Range Theory According to Merton*:

  “theory involves abstraction, of course, but it is close enough to observed data to be incorporated in propositions that permit empirical testing.”

Initial Conceptual Model
(from prior work on SMAs)

Patient Factors (demographics, adherence, other factors)

Patient with Diabetes

Informed, Activated Patient

Patient at 3 Months

Medication Adjustments

Patient at 12 Months

*Improving Self-Management Activities
*Improving Problem Solving Skills
*Overcoming Barriers to Treatment

Shared Medical Appointment
Peer support
Multidisciplinary Expertise

Series of management-related questions provided by J. Øvetvreit

03/08
• Number of articles by year
  – 2008 or earlier: 38
  – 2009: 10
  – 2010: 9
  – 2011: 14

• Number of articles by U.S. or international
  – U.S.: 56 (78.9%)
  – International: 15 (21.1%)
Characteristics of SMAs

- Educational component – 88.7%
- Multidisciplinary members – 64.8%
- Included a behavioral intervention – 50%
- Included medication adjustment – 55.7%
- Included peer-to-peer support – 87.3%
- Included clinician training – 42.9%
Characteristics of SMAs

**Visit duration**
- <60 minutes: 1.7%
- 60-89 minutes: 9.9%
- 90-120 minutes: 59.2%
- >120 minutes: 12.7%
- Missing: 16.9%
- Missing: 23.9%

**Visit frequency**
- Once: 8.5%
- Weekly: 16.9%
- Every 2 weeks: 2.8%
- Monthly: 29.6%
- Every 2 months or more: 18.3%
Q8: Choose to participate in SMA

Q7: Peer support

Informed, Activated Patient

• Education about:
  • Self-Management
  • Problem-Solving Skills

Multi-disciplinary Expertise

Medication Adjustments

Improving Self-Management Activities

Insulin Inertia

Outcomes Clinical Cost Satisfaction

Q3, 4 and 5

SMA

Patient Factors (demographics, adherence, other factors)

Patient with Chronic Disease

Q1, 2

Q6, 7, 10

Version 2_11-8-11

10/11b
External Context
- Geography/location
- Healthcare system

Patient Context
- Social/Behavioral Context
  - Educational attainment/ Health literacy/ numeracy
  - Social/ Psychological Factors
- Physiological Context
  - Disease and severity
  - Co-morbid conditions
  - Psychological Factors

Desire to participate in SMA
Ability to participate in SMA

SMA
- Education
- Behavioral Intervention, e.g., motivational interviewing
- Peer to Peer Support
- Multi-professional participation
- Multi-professional leadership
- Medication Adjustment
- Intervention Dose, e.g., length, frequency, size

(Updated) Mechanisms
- Self-Management
- Treatment adherence
- Self-efficacy
- Quality of life
- Functional status
- Depression/ anxiety
- Clinical outcomes
- cost/health services utilization
- Process measures of care

Context (high level) Mechanisms

Q3,4,5 Outcomes

Q6

Q7A,8

Q9

Q7B, 10

04/12
Some preliminary (hypothetical) demi-regularities

C₁
Across geographies (urban vs suburban vs rural)

M₁
Travel distances affect participation rate

O₁
Effects similar

C₂
Across patient factors (Socioeconomic status)

M₂
SES associated with educational attainment

O₂
Effects similar

C₃
Across chronic diseases

M₃
Self management principles similar across disease

O₃
Effects similar
Some preliminary hypothetical demi-regularities

<table>
<thead>
<tr>
<th>Context</th>
<th>Mechanism</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>$M_1$ Behavioral component</td>
<td>$O_1$ better outcomes if present</td>
</tr>
<tr>
<td>$C_1$</td>
<td>$M_2$ Multiprofessional</td>
<td>$O_1$ better outcomes if present</td>
</tr>
<tr>
<td>$C_1$</td>
<td>$M_3$ Medication Adjustment</td>
<td>$O_1$ better outcomes if present</td>
</tr>
<tr>
<td>$C_1$</td>
<td>$M_4$ Duration of SMA</td>
<td>$O_1$ better outcomes if longer duration (90-120 minutes)</td>
</tr>
<tr>
<td>$C_1$</td>
<td>$M_5$ Peer support</td>
<td>$O_1$ better outcomes if present</td>
</tr>
</tbody>
</table>

Also found that SMAs that used more of these had better outcomes.
Further unpacking a mechanism seeking further support from established theory

- Peer Support is a strategy – a high level mechanism that becomes the context for the next level.
Further unpacking a mechanism seeking further support from established theory

- Multiprofessional/interprofessional team is a strategy – a high level mechanism that becomes the context for the next level.

Beliefs about /attitudes towards self-management

- $M_1$ motivation to comply with important others reinforced by multiplicity of professionals
- $M_2$ trust in professionals others reinforced by multiplicity of professionals
- $M_3$ interprofessional practice is more patient centered.

Outcome (O)

$+++$ Beliefs about /attitudes towards self-management

More effective self-management
How is this theory driven, context dependent evaluation different?

- Iterative development of conceptual model for contexts, mechanisms, outcomes
- Based on principles and not a series of sequential steps
- No particular preference for quantitative or qualitative methods
- Multidisciplinary stakeholders and participants
- Stakeholders are regarded as key sources for eliciting program theory and providing data on how the program works
Lessons Learned

• Generalizability is not “just” a philosophical question. It is core to practice in the real world. Not necessarily a need for further “research” but a need for researchers to report on different information.

• Cannot manufacture new data - can only look at existing data in a different way.

• Current literature has very little information on patient perspective. Need for qualitative data.
<table>
<thead>
<tr>
<th>Traditional ‘Cochrane’ Review</th>
<th>Realist Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identify the review question</td>
<td>1. Clarify scope of review (identify review question; refine purpose of review; Articulate key theories to be explored)</td>
</tr>
<tr>
<td>2. Search for primary studies, using clear predefined inclusion and exclusion criteria</td>
<td>2. Search for relevant evidence, refining inclusion criteria in the light of emerging data</td>
</tr>
<tr>
<td>3. Appraise quality of studies using a predefined and validated critical appraisal checklist, considering relevance to research question and methodological rigour</td>
<td>3. Appraise quality of studies using judgement to supplement formal checklists, and considering relevance and rigour from a ‘fitness for purpose’ perspective</td>
</tr>
<tr>
<td>4. Extract standard items of data from all primary studies using template or matrix</td>
<td>4. Extract different data from different studies using an eclectic and iterative approach</td>
</tr>
<tr>
<td>5. Synthesise data to obtain effect size and confidence interval and/or transferable themes from qualitative studies</td>
<td>5. Synthesise data to achieve refinement of programme theory – that is, to determine what works for whom, how and under what circumstances</td>
</tr>
<tr>
<td>6. Make recommendations, especially with reference to whether findings are definitive or whether further research is needed</td>
<td>6. Make recommendations, especially with reference to contextual issues for particular policymakers at particular times</td>
</tr>
<tr>
<td>7. Disseminate findings and evaluate extent to which practitioners’ behaviour changes in a particular direction</td>
<td>7. Disseminate findings and evaluate extent to which existing programmes are adjusted to take account of elements of programme theory revealed by the review</td>
</tr>
</tbody>
</table>