The Research that Underpins Quality Improvement

Lloyd Provost, Maxine Power, Mary Dixon-Woods

December 9th, 1:00-4:30

Integration of Research and QI:
The research that underpins quality improvement

An interactive workshop on how to combine qualitative and quantitative methods in quality improvement research.

After a short introduction, participants will work in small groups on a case study exercise.

A plenary discussion will address issues and challenges raised by the case study.
Program theory

Commit to an ethic of learning
Why research? Why not just do?

“One of the most characteristic aspects of the present situation is that specific reforms are advocated as though they are certain to be successful”. Donald T Campbell, Reforms as Experiments, 1969

---

Why research? Why not just do?

- So that everyone can learn from what you did.
  - Ethical commitment to adding to cumulative knowledge
  - Peer-reviewed publication is the standard for scientific credibility
- To find out whether it achieved your goals.
- To find out what the intervention actually was (not what it was intended to be).
- To identify the effects of the intervention (good and bad).
- To produce evidence of costs and value for money.
Why research?

- To figure out how and why the intervention produced the effects it did.
- To determine whether to continue with an intervention or replicate it elsewhere.
- To make modifications and adaptations to improve the intervention.

Some possible research methods for studying improvement

- Trials
  - Randomised at individual level
  - Cluster randomised
  - Stepped wedge
  - Wait-list
- Before and after studies
- Controlled before and after studies
- Mixed methods
- Qualitative only
- Case studies
- etc etc
The role of program theory

Need to choose design carefully, using program theory and *nature of answer wanted* as a guide.

Empower patients to ask doctors to wash their hands

Infection goes down
Empower patients to ask doctors to wash their hands

Then a miracle occurs

Infection goes down

Need for program theory

“Nothing improves research design so much as having a clear idea about what is being investigated.

An important function of theory in research design is to help researchers ensure that they are playing in the right ballpark to begin with—that is, to help them avoid studying the wrong thing” (Lipsey, 1993)
Logic models and program theory

<table>
<thead>
<tr>
<th>LOGIC MODEL</th>
<th>PROGRAM THEORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specifies the components and inputs</td>
<td>Account of how the program works to produce the intended outcomes – and the unintended ones</td>
</tr>
<tr>
<td>Specifies the (desired) outcomes</td>
<td></td>
</tr>
<tr>
<td>Assumed linkages between them</td>
<td></td>
</tr>
</tbody>
</table>

What is program theory?

A plausible account of how the intervention works

- Identifies the mechanisms that link those activities and resources to the desired outcomes
- Identifies conditions likely to be favourable or unfavourable to functioning of the mechanisms
- Identifies possible unwanted outcomes (the side effects)
How to build the program theory

- Off the shelf theory + previous studies provide a point of departure
- Program personnel are working with a theory (even if they don’t know it) – which you need to elicit
  - Interviews
  - Program documents
  - Observations

Challenges of building program theory

- Elicited theory is often incomplete or not fully coherent
- It sometimes reflects an attempt to please or appease various parties
- Often weaknesses are evident from the start
  - Conceptually implausible or fanciful
  - Too ambitious
  - Doesn’t account for different environments and capacities
  - Mundane and logistical issues poorly addressed
  - Timescale too short
Without a decent program theory

- The logic underlying the intervention may be flawed
- The intervention may not be causally linked to desired outcomes
- You’re not sure whether the effects you see are due to the intervention or the implementation
- It’s hard to replicate the program or design new, similar ones

Without knowledge of mechanisms

Tendency towards magical thinking
A remarkably successful improvement program

Presentation title set in header
How did it work?

Improvement collaboratives

- Number of different models
- IHI Breakthrough Series (BTS) widely used
- Characteristic structure of
  - multiple teams,
  - learning sessions led by faculty
  - Use of improvement methods, including PDSA cycles and measurement, between sessions,
Intervention: BTS Collaborative

Figure 2: Breakthrough Series Model

Select Topic
Recruit Faculty
Develop Framework and Changes

Enroll Participants
Prework

LS1: Learning Session
AP: Action Period
P-D-S-A: Plan-Do-Study-Act

Supports:
- Email • Visits • Phone Conferences • Monthly Team Reports • Assessments

Semantic Congresses and Publications

Taken from the IHI Breakthrough Series Collaborative White Paper. www.ihi.org

Working Group Exercise 1
(5 minutes)

How and why would you expect a BTS collaborative to work?
The Stroke 90:10 collaborative

Improving stroke care across North West England using the IHI Breakthrough Series Collaborative

Maxine Power PhD MPh
National Improvement Advisor,
Department of Health,
England.
Context: NHS North West

- 23 hospital Trusts
- 8 mental health hospitals
- 8 specialist hospitals
- 1 ambulance trust

272,000 staff

National Sentinel Audit of Stroke Care
2000-2008

Evidence Based Policy ++++
Sentinel Audit

<table>
<thead>
<tr>
<th>Indicator</th>
<th>2004 (%)</th>
<th>2006 (%)</th>
<th>2008 (%)</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain Scan (&lt; 24 hrs)</td>
<td>56</td>
<td>41</td>
<td>59</td>
<td>52</td>
</tr>
<tr>
<td>Aspirin or alternative antiplatelet(&lt; 48 hrs)</td>
<td>60</td>
<td>65</td>
<td>85</td>
<td>63</td>
</tr>
<tr>
<td>Screen for swallowing disorders (&lt; 24 hrs)</td>
<td>58</td>
<td>67</td>
<td>72</td>
<td>66</td>
</tr>
<tr>
<td>Patient weighed (once during admission)</td>
<td>43</td>
<td>55</td>
<td>72</td>
<td>57</td>
</tr>
<tr>
<td>Physiotherapy assessment (&lt; 72 hrs)</td>
<td>56</td>
<td>67</td>
<td>84</td>
<td>69</td>
</tr>
<tr>
<td>Occupational therapy assessment (&lt; 7 days)</td>
<td>56</td>
<td>66</td>
<td>66</td>
<td>63</td>
</tr>
<tr>
<td>Patient’s mood assessed (during admission)</td>
<td>43</td>
<td>59</td>
<td>65</td>
<td>56</td>
</tr>
<tr>
<td>Goals Agreed by MDT</td>
<td>77</td>
<td>80</td>
<td>86</td>
<td>81</td>
</tr>
<tr>
<td>Treated within a stroke unit (50%+ LOS)</td>
<td>38</td>
<td>53</td>
<td>58</td>
<td>50</td>
</tr>
<tr>
<td><strong>North West</strong></td>
<td><strong>54</strong></td>
<td><strong>61</strong></td>
<td><strong>72</strong></td>
<td><strong>62</strong></td>
</tr>
</tbody>
</table>

90 by 2010

**Aim**

To determine whether hospitals participating in an improvement collaborative improved more than controls on key process measures from the Sentinel Audit.
Working Group Exercise 2
(15 minutes)

Design the study

What we did
Study Design: Cluster randomised controlled trial with interrupted time series design

- 24 hospitals invited to participate
- Consent & randomisation
  - Intervention (12)
    - Collaborative Jan 2009 for 1 year
  - Control (12)
    - No collaborative
      - 2009 for 1 year

Logic model

Understanding the changes we can make that will result in improvement:

- Immediate recognition & management to minimize loss of brain tissue
- Consider for t-PA & deliver
- Brain imaging
- Swallow screen
- Consider for aspirin & deliver
- Assess nutrition, hydration & weight

- Treated in a stroke unit
- Physiotherapy assessment
- Occupational therapy assessment
- Assessment of Mood
- Patient centered MDT goals

- Home visit prior to D/C
- Caregiver involvement and support
- Continuity with health issues
- Personal care, mobility & nutrition
- Mental health
- Lifestyle change

Improved Stroke Outcomes

- Mortality, complications, functional status & quality of life
- Optimize outcome & prevent complications
- Management of long term care & support for living with stroke
Strategy for Change

<table>
<thead>
<tr>
<th></th>
<th>Brain imaging</th>
<th>Swallow</th>
<th>Aspirin</th>
<th>Weight</th>
<th>Audit score</th>
<th>All or none</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>√</td>
<td>x</td>
<td>x</td>
<td>√</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Patient 2</td>
<td>x</td>
<td>√</td>
<td>x</td>
<td>√</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Patient 3</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Patient 4</td>
<td>√</td>
<td>x</td>
<td>√</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Total</td>
<td>75%</td>
<td>50%</td>
<td>50%</td>
<td>75%</td>
<td>63</td>
<td>25%</td>
</tr>
</tbody>
</table>

Two Bundles

Bundle 1 (acute)
- Brain Scan
- Swallow Screen
- Aspirin
- Weight

Bundle 2 (rehab)
- OT assessment & treatment
- PT assessment & treatment
- MDT goals set & reviewed
- Mood assessment
- More than 50% stay in SU

Random Sample of 20 patients per month per organisation collected by case note review
(July 2008 – December 2010)
Quality Improvement and Measurement

Health Care Quality Improvement

“A broad range of activities of varying degrees of complexity and methodological and statistical rigor through which health care providers develop, implement, and assess small-scale interventions and identify those that work well and implement them more broadly in order to improve clinical practice and outcomes.”

Quality Improvement Research

The design, development and evaluation of complex interventions aimed at the re-design of health care system to produce improved outcomes.
**Characteristics of Health Care Quality Improvement**

- Contextual factors (background variables or confounders in research) are a major focus
- The initial intervention (changes to the system) are adapted and modified as study progresses
- Measurement happens over time (improvement is temporal)
- Graphical analysis and presentation (SPC)
- Involvement of local expertise in conducting project
- Multiple experimental cycles for quick feedback and learning
- Multi-factor experiments to learn from complex systems with non-linear and dynamic cause and effect relationships
- Building reliability of the interventions is often a major part of the improvement effort
- Sustainability is a consideration from the beginning of the project

---

**Rigorous Learning in Complex Systems**

![Diagram](image)

*Adapted from Berwick, 2009*
Minimum Standard for Reporting Data in a QI Project: Annotated Time Series

Beyond Run Charts: Shewhart Charts

The Shewhart chart is a statistical tool used to distinguish between variation in a measure due to common causes and variation due to special causes.

(Most common name is a control chart, more descriptive would be learning charts or system performance charts.)
Experimental Designs for QI Research

*Speroff and O’Connor, Study Designs for PDSA Quality Improvement Research, Q. Manage Health Care, Vol 13, No.1, 2004*

<table>
<thead>
<tr>
<th>Study design</th>
<th>Characteristic</th>
<th>Primary concern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-series (AB)</td>
<td>Continuous, longitudinal data</td>
<td>Historical control</td>
</tr>
<tr>
<td>Equivalent time-series (ABAB)</td>
<td>Replication</td>
<td>Carry over effects</td>
</tr>
<tr>
<td>Multiple baseline</td>
<td>Lagging of interventions</td>
<td>Contamination</td>
</tr>
<tr>
<td>(AAAB, AABB, ABBB)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factorial</td>
<td>Experimental design</td>
<td>Confounding</td>
</tr>
</tbody>
</table>

**Time Series (or AB) Study Design**

(interrupted time series design)

*Speroff and O’Connor, Study Designs for PDSA Quality Improvement Research*
Multiple Interrupted Time Series Design

Speroff and O'Connor, Study Designs for PDSA Quality Improvement Research, 2004

Three Categories of Measures in QI

Outcome Measures
Voice of the customer or patient. How is the system performing? What is the result?

Process Measures
Voice of the workings of the system. Are the parts/steps in the system performing as planned?

Balancing Measures
Looking at a system from different directions/dimensions. What happened to the system as we improved the outcome and process measures? (e.g. unanticipated consequences, other factors influencing outcome)

HC Data Guide, p 36
Combining Process Measures

- Item-by-item measures – report each process, often results in too many measures
- Composite Measures – average performance, process focused.
- “All or none” measures – patient and system focused
- Consider “all or none” when each component is performing >50%

Collaborative Data Analysis – Small Multiples to Support Aggregate Displays

15 Diabetes Clinic teams

% of Patients with HbA1c <7
- Aggregate of Diabetes Teams

Scales for Small Multiple Graphs

J-07 M-07 S-07 J-08 M-08
Small Multiples in QI Research Publication

Small Multiple Run Charts in QI Research Publication

Figure 3: Small multiples chart for monthly percentage of patients with diabetes screened for microalbuminuria, all practices, March 2006–March 2008.
Break – 30 mins

Working Group Exercise 3
(15 mins)

Challenges
What are the challenges with measurement that you would expect to encounter in the Stroke 90:10?

How do you think learning from the data could be maximised?
Qualitative research

Reseaching an intervention in action

For complex interventions, you typically need multiple methods to figure out

- What were the effects (wanted and unwanted) of the intervention?
- What was the program really (not what intended to be)
- How does it work? Are there alternate pathways?
- Does it operate differently in different contexts?
Qualitative research

- Has different roles in different phases
- Identify areas for research
- Generate hypotheses
- Offer explanations for findings
- Identify mechanisms
- Assess feasibility, implementability and sustainability
- Identify relevant features of contexts

Programme in action may look very different from programme as intended.

- Key activities may not happen or look different from how expected
- The mechanisms may not function in the way intended
- External and internal contexts may powerfully modify what is possible
- Programme designers may improvise, innovate and adapt
- So part of the job of the research is working out what really happens, and updating the theory
Ethnography

- Ethno – human society/culture; Graphy – to describe
- Is the systematic study of a particular cultural group or phenomenon
- The ethnographer is the data collection instrument
  
  “grounded in a commitment to the first hand experience and exploration of a particular social and cultural setting” (Atkinson, 2001)

- Ethnographers use multiple data sources and methods of data collection
- Data collection involves observation, chats, interviews, focus groups, documentary analysis

What are you looking for?

- What is the setting like? What structures and resources are available? Sketch as well as note; you may also be able to take photos

- Culture: the way we do things around here
  - Observable patterns of behaviour, customs, ways of doing things, norms, assumptions
  - Networks, key people
Triangulation

- Trying to get different perspectives on the same issue
  - E.g. Do the doctors and nurses think the problem is the same?
- Getting different types of data
- Need multiple sources, collected by diverse methods

Interviews

- You will be doing lots of chats as you hang out
- Often these are supplemented with tape-recorded semi-structured interviews
- These allow you to test out and explore in more detail things you have observed
- Allow you to get more of a handle on how people in the field understand things
Documents

- You can collect documents and other artefacts and subject them to analysis

- Things like notices, posters, minutes of meetings, information sheets, instructions, handbooks etc are very important

Analysis

- As patterns emerge, follow them across the fieldnotes etc

- Try to collect the clues and dots; look for things that don’t fit or test your ideas

- Needs to be an authentic, plausible story that fits with the data

- Mid-range theory = explaining generic features of specific social phenomena
Task for you

- Design in outline a qualitative study to find the mechanisms of change for Stroke 90:10
- What challenges would you expect to encounter? Why?

Working Group Exercise 4
(10 mins)

Design a qualitative study to investigate the program theory and mechanisms of action in Stroke 90:10
Results of Stroke
90:10
Effects of Change: Bundle 1

Analysis of ‘Difference in Differences’ of the slope of the line:
no relative benefit

OR 1.07 (95% CI 1.01, 1.15; p=0.032) control
OR 1.00 (95% CI 0.90, 1.11; p=0.975) intervention.

Effects of Change: Bundle 2

Analysis of ‘Difference in Differences’ of the slope of the line:
benefit 20% inc

OR 1.22 (95% CI 1.14, 1.31; p<0.001) control
plus an additional 1.20 (95% CI 1.07, 1.35; p=0.002) for intervention.
Qualitative data

- Small multiples
- Semi-structured in-depth with 34 individuals from 11 (out of 26) participating trusts that demonstrated varying levels of performance. Between 1 and 10 interviewees from each organisation.
- Two interviews with the collaborative faculty. 4 interviewees in total

Challenges of the commons

- Though many showed a remarkable commitment to the values of the collaborative, not all did
- Values of honesty, shared purpose, learning, humility and commitment to sharing information and ideas were critical to getting the collaborative to work
- But some people took more than they gave, and not all demonstrated full honesty or were deluded
Managerial and clinical leadership

Agreement between the participants and the Faculty Team

Well we think certainly that people who had the leadership support did better or senior medical staff supported much better. (Faculty Team)

R: … clinical leadership - that’s needed in order to make the improvements. (Clinical Team)

Behavioural variables

Agreement between the participants and the Faculty Team

We were really committed, it was a good core team, you know that worked together. (Clinical team)

They just never sat back and said right, that’s good enough. (Faculty Team)
Exceptional individuals

Agreement between the participants and the Faculty Team

Oh God, Person4 was dynamite. She was...
002: She had a whip.
003: Unbelievable. Where she didn’t get in that organisation wasn’t worth going. So she collected data on all patients and she was the one that really signalled in the first instance, how poor the coding was, wasn’t she? (Faculty Team)

(person 4) will tell you about it. Because literally she collected it all herself. She had no help from anyone […] She managed to do it all.

Teamwork

Agreement between the participants and the Faculty Team

So the difference there was, they had the most cohesive and coherent team in the whole programme (Faculty Team)

We felt the whole team had to be involved and that’s what we did, we involved the team. We have a stroke operations group meeting once a month, where we update everybody, keep them all fully involved in what was going on, publishing the results of all the key performance indicators every month in a big coloured laminated sheet. (Clinical team)
Geography

More emphasis placed by the Faculty Team

The geography was just overwhelming though. They just didn’t see that they were ever going to be able to succeed (Faculty Team)

Staffing and resources

Disagreement between the participants and the Faculty Team

003: <Person 23> turned up probably three times.
001: Yeah I can remember her being some of the worst you know, we need more money, we need more resource...

But I didn’t input probably as much as I might have wanted to do. But that’s down to resources and freeing up enough staff. (Clinical team)

We tried to cope with it as best we could in that the logistics of it was that you couldn’t even send the same people to the learning session twice. (Clinical team)
Learning infrastructure and improvement methodology

More emphasis placed by the Faculty Team

And didn't really have enough of a grasp of the science, the method to be able to recover. Which sort of speaks to the fact that they didn't have reliability in their systems. So they were probably dependent on quite a lot of hard work and vigilance. (Faculty Team)

I think they would have admitted that they were keen and enthusiastic and had a vision, but didn't have much of a method. So it’s been the old fashioned school of, we’ll tell you what to change and you’ll damn well change it. You know, which goes some way I think to making improvement, but I think you come up against a few bumps I think. (Faculty team)

“Priority thickets”

More emphasis placed by the Faculty Team

However the organisation was not in the best of places. So the organisation was sort of landlocked into a sort of a battleground for services, for all sorts of things. And they were under the cosh because their activity was high and therefore they weren’t achieving some of the basics. They were also on an estate that was really in trouble, like they kept getting fire notices of shutting them down. So there’s a lot of other stuff to distract them from what they should be doing. (Faculty Team)
Qualitative findings: Management support

- Stroke 90:10 helped raise profile and status of stroke care
- Persuaded management to release or dedicate resources
  - *But what we felt was the only way that we could actually get patients in directly and quicker was to have some open assessment beds. So it actually took us speaking with our senior management, actually the chief exec and the chief of operations really. We went to the top. We wouldn't, I don't think we'd have been able to do that without almost the 90:10 backup.* (008)
- But some trusts did not provide managerial support and teams were left floundering

Qualitative findings: Data collection

- One of the most challenging aspects
- Stroke 90:10 is “data hungry, full stop” “a nightmare” “is it really necessary?”
- Some identified a specialist data collector, others distributed across teams
- Stroke 90:10 faculty support was welcomed, but still had to arrange for all the notes to be available
- Irritation at “changes in goal posts”
- Frustration at different requests for data from different sources
- Some still using data collection forms after end of the project
Conclusions

- BTS collaborative may have additional benefit in creating focus and empowering change in frontline teams.

- Policy and system changes also may have impact, in particular where changes are comparatively narrow in focus.

Qualitative and quantitative research in improvement science

- QI research needs to integrate the rigor of research with the flexibility of quality improvement.

- Time series designs are practical for QI research since improvement in temporal.

- QI research needs both quantitative and qualitative methods.