
July 18 2011

Presentation at a Meeting sponsored by the Alliance for Health Reform

Marsha Gold

Based on a paper by Gold, Helms, and Guterman for The Commonwealth Fund
Issues Addressed in Paper

- Focusing on change that matters (agenda setting)
- Documenting innovations to support effective learning and spread (formative feedback and process analysis)
- Balancing flexibility and speed with rigor in developing evidence to support policy change (outcomes evaluation)
Focusing on Change That Matters

- $10 billion isn’t that much money
- Is there potential for a large impact on joint “quality/cost metric”?
- Large gains over small populations or small gains over large populations
- Plausibility, “back-of-the-envelope” estimates of potential given numbers, demographics, and range of likely effects
What does the innovation seek to achieve and how?

Logic models, defined measures for “success”

Tracking what was implemented and when versus what was planned

Timely measurement and feedback to innovators—on metrics that matter to them
Documenting and Learning from Innovation—II

- Efficiency: investing in shared metrics and approaches for cross-site learning
  - Characteristics of innovations
  - Characteristics of context
  - Common metrics of success

- Realistic expectations: implementation always takes longer than expected and more so if the context is complex

- Minimize barriers that slow or drain momentum
Evidence to Support Broad Program Change

- HHS has authority, but CMS actuary must certify that change won’t add to program costs (and for the secretary, that it has demonstrated the potential to improve quality)

- What will/should be the standard of evidence?
Historical Approaches to Evaluation

- Careful definition of target population for the purpose of judging success
- One or more comparison groups “otherwise similar” to serve as benchmark
- Metrics typically constructed from centralized data files, existing or new
- Long time frame to distinguish short-term effects from stable long-term effects
Likely Reality of Innovation Testing

- “National” demonstrations across widely divergent organizations
- “Bottom up” innovation with variation in detail across sites
- “Contaminated” comparison groups
- Desire for rapid feedback on “right direction” even though some effects may take time to surface.
Realism on Standards of Evidence

- Trade-offs between “type 1 and type 2 errors”
- Moving too quickly versus too slowly on improvements: how “good” are things now, what risks are there in change?
- Congressional history: Legislators have acted before evaluations are done. They have also failed to act despite evaluation results showing what was or was not successful.
Conclusions

- Demonstrating the feasibility of implementation is an important and potentially powerful outcome.

- Utility of pilots enhanced by high-quality evidence that:
  - Clarifies in advance what is to be tested and why
  - Provides ongoing, consistent, and timely measures of intended and unintended changes
  - Includes appropriate analysis to aid in attribution
  - Gives “enough” information on context and implementation to allow suitable and spread to be assessed by diverse stakeholders
Trade-off between “rigor” and “rigor mortis”

Distinguish useful initiatives to propagate from efforts that mainly preserve the status quo or are actually harmful

Avoid stifling innovation to improve system because “no data are good enough”

Appropriate trade-offs likely to vary across innovations at different stages or with different risks/rewards
For More Information

- Download Gold, Helms, and Guterman paper at www.cmwf.org

- Marsha Gold at 202-484-4227 or mgold@mathematica-mpr.com