



EXECUTIVE SUMMARY

**Assessing the Effects of
Primary Care Transformation:
Emerging Themes and Practical Strategies
to Strengthen the Evidence**

by Nancy McCall, ScD, and Kristin Geonnotti, PhD

The Milbank Memorial Fund's Multi-State Collaborative (MC) is a working group of 18 states and regions actively engaged in multi-payer primary care transformation through the implementation of patient-centered medical home (PCMH) programs. These innovative efforts include payment reform and enhanced multi-disciplinary support services. Each of the MC programs were early adopters of multi-payer primary care—their leaders made investments of time and resources before knowing what the outcomes would be.

Assessing their own programs has been an important component of the MC programs—and each MC member has been conducting an evaluation of the effectiveness of its PCMH program. As the programs developed, they grew in complexity—as did the methods needed to evaluate them. The challenge was to ensure that promising results were not missed—or impacts overstated.

In order to understand the extent to which MC PCMH programs are improving outcomes for a core set of key health care utilization and spending measures, the Fund asked Mathematica Policy Research to develop this report. It analyzes eight of the 18 MC member evaluations in order to assess the strength of the evidence being reported and to provide a foundation for learning how to strengthen future advanced primary care evaluations.

Key Findings

As part of the report, the authors cataloged the evaluation methods and results of eight participating MC members' most recent evaluation, assessing the strength of each, and identifying common strengths and gaps in methods. Using a structured analysis, the authors assessed four domains of evaluation design in each of the eight projects: comparison group, study design, whether the study is well-powered enough to detect effects, and statistical methods. The authors identified methodological reasons in each of the evaluations to question both favorable and unfavorable findings, with implications about relative accuracy and strength.

Assessment standards for the four domains and a summary of findings:

- **Comparison group is sound.** There is widespread consensus in the research community that a sound comparison group is critical to producing unbiased results. The comparison group should have similar geographic, sociodemographic, and health care system characteristics as the intervention group. Seven of the eight evaluations did that. Six of the eight also allowed for comparison of the pre- and post-intervention periods.
- **The study design is rigorousness.** Evaluation design needs to be rigorous. A variety of evaluation approaches were used in the MC evaluations, all with elements of a rigorous evaluation design.

- **Study is well-powered to detect effects.** None of the MC evaluations studied were well-powered enough—of sufficient sample size for the study subgroups and measures being studied—to detect definitively that there were changes in utilizations or expenditure.
- **The statistical methods are sound.** There was a lack of uniformity in the statistical methods used to estimate program effects among participants. This may be a reflection of limitations in availability of resources or time for adequate statistical analysis.

Finally, the report provides practical strategies for future evaluations of primary care transformation programs:

- **Addressing statistical power and methods.** Making sure that evaluation methods are rigorous and that evaluators are open to alternative methods of analysis, such as Bayesian methods.
- **Making sure that the intervention patient and practice group is as similar as possible to the comparison group.** Evaluators should seek to understand if the comparison group is participating in another health system or payment reform initiative.

Future Directions

Because of the dynamic and complex environment of health care systems today, future evaluations of health care delivery system reform efforts that involve primary care should examine the overall and specific features of the primary care transformation, including all reform efforts and their relationship to one another.

Qualitative research tools will help link program features to program outcomes. Lastly, we need to move away from traditional methods and testing that imply that there is a simple “yes” or “no” assessment.

Strengthening evaluation methods will increase the level of confidence in this work. Positive program evaluations will reflect *real change* and will encourage continued support. Written for evaluators, funders of PCMH programs and evaluations, implementers of PCMH programs, and policymakers, this report aims to increase our understanding of the effectiveness of the PCMH model of primary care.

Primary care transformation and PCMHs themselves continue to evolve in response to service delivery and payment innovations. In a health care environment that is rapidly adopting new service and delivery models—each of them different and with varying effects on quality and cost outcomes—evaluation efforts must balance rigor and adaptation. To learn, improve, and make informed decisions about allocation of resources, all stakeholders must have the most accurate information possible. This report describes some of the fundamental principles that should be considered in this effort as we move ahead.