Evaluating Two Mysteries: Camden Coalition Findings

Teresa Cutts, PhD, Wake Forest School of Medicine, and Gary Gunderson, MDiv, DMin, Wake Forest School of Medicine and Wake Forest Divinity School

July 27, 2020

Introduction

Two fundamental mysteries characterize the health care sector’s embrace of the concept of population health [1]. Mystery A is why health care systems cannot control the extraordinary concentration of cost among the sickest 1–10 percent of patients. Mystery B is how to understand the relationship between those extraordinary costs of high-tech health care and the reality of neighborhoods near the hospitals, which hold the greatest disparities, most underserved groups, and thus a high proportion of the costs. The answer to both questions leads away from hospitals onto the streets and into the homes of community members with lower incomes and chronic illnesses. One of the most iconic population health programs working to understand why both mysteries persist, the Camden Coalition of Camden, New Jersey, was recently evaluated by Finkelstein, Taubman, and Doyle in the New England Journal of Medicine (NEJM) using a randomized controlled trial (RCT) methodology [2], prompting this perspective.

In Memphis, Tennessee from 2008 to 2013, the population health work conducted by the authors wrapped highly complicated educational programming and referrals around all partners and produced years of encouraging data, now adapted into other social and institutional settings. Our similar, subsequent work statewide in North Carolina, the North Carolina Way, now engages over 500 congregations and hundreds of other community partners [3]. We argue that traditional educational programming, computer referrals, data analytics, and RCTs have not been constructed to accurately measure the effectiveness of these community-based, “messy” approaches to improving population health. That is why the RCT findings of the Camden Coalition are so disappointing—the data cannot and should not be evaluated by methods that can't accurately convey its impact.

The Launch of Hot Spots and the Camden Coalition

Jeffrey Brenner was a young physician who rode with his local police department in 2011 to what police called crime “hot spots” in Camden. Brenner noticed overlap with addresses of patients he knew were incurring high costs of care at the four local hospitals, so he set about trying to understand why this intersection was so clear to him and not many others. Data points in other neighborhoods conveyed the same obvious message—encouraging many hospitals to pay consultants to look at their own data and see if their high-needs patients overlapped with crime hot spots or other patterns.

Brenner’s work converged with the movement to identify the social determinants of health as a critical part of delivering health care services, which was marked globally by a World Health Organization study published in 2013 [4]. The Memphis Model, implemented in Tennessee, also emerged during this period and viewed the needs of lower-income community members through the eyes of community faith partners [5] and health system data. Brenner was one of the keynote speakers at a 2011 White House Conference, “Improving Health Outcomes through Faith-Based and Community Partnerships,” that launched Stakeholder Health—a national voluntary learning collaborative of mission-driven hospitals. The Camden Coalition’s approach involved tending to all of the needs of the most complex patients, including social, mental, and the traditional social determinants of health like steady housing, gainful employment, access to healthy foods, and access to transportation. In American health care, Brenner deserves credit for launching the hot-spotting techniques and movement.

Therefore, many of us were disappointed in the findings of this evaluation, that used a randomized trial
methodology and found that the Camden Coalition approach failed in reducing hospital readmissions within 180 days after discharge, the primary outcome evaluated [2]. The trial focused entirely on Mystery A (cost of “super utilizers”), not Mystery B (relationship between high costs and the social needs of neighborhoods). The authors of this paper explored the complexity of Mystery B in a 2015 NAM Perspectives paper [6], promoting the concept of health systems caring for underserved populations within certain zip codes and census tracts in aggregate (versus panels of high user cohorts). The 2020 evaluation of the Camden Coalition’s work may put a big asterisk next to the claims of all who are using cost-containment arguments to encourage hospitals, Medicaid, and insurance companies to invest in social variables (e.g., housing, food security). Decision makers may use these findings as a reason for backing away from similarly complex efforts in light of the COVID-19 phenomenon, which is creating an avalanche of ever-changing social and mental health needs as hospitals struggle to keep the clinical system afloat.

The Problem with Traditional Evaluation

Finkelstein and colleagues’ study may also add doubt to the less common argument that the proper focus for cost control is not cohorts of high utilizers, but the aggregate cost of the populations in zip codes and census tracts [7]. This geographic approach requires an even more complex, sustained, and multi-partnered strategy. Many in health care may accept the “failure” of the Camden Coalition’s work as permission to go back to the old medicine-in-the-hospital model. Even worse, they may use the findings to justify simply ignoring the lower-income, higher-need people and neighborhoods by retreating into a model of community benefit measured by giving away clinical services instead of engaging Mystery B. We suggest avoiding those paths and, instead, examining the complexity of the Camden experience with humility from multiple perspectives—those of program leaders, funders, evaluators, persons served, and external variables.

Science draws us into trying things that may not prove to be so. William Foege, a scientist who figured out how to eliminate smallpox, said that “you have to believe something before you can see it.” That is, you have to believe in a process before you create it, fund it, and do the work that eventually produces a data stream that can finally be evaluated. Every one of those steps is a projection, not just from the creator’s mind (Brenner’s, ours, or yours). The initial idea, the funding proposal, and the final operational budget for a health improvement effort are three different projections, each of which tends to subtly bend the design toward competing stakeholders’ appetite for Mystery A or B. Most settle for Mystery A (address the high cost) because it seems more doable, which means provable. Stakeholders tacitly agree to do the more important things...later. The work then continues for some years, among a multitude of complex and unpredictable humans, amidst unknown community variables and constant program adaptations. Brenner himself emphasized how the Camden work had adapted similarly [8], but these changes were not evaluated in the RCT; the article evaluated the 2014 version of the Camden Coalition, not the current one.

Traditional RCTs may be ideal for rigorously evaluating pharmaceutical or more linear interventions, but such methodology is not adequate for capturing the full iterative flavor of years-long, community-based work like that in Camden. There are questions about RCTs’ utility in measuring “complicated,” and certainly, more “complex” interventions like those of Brenner’s team. Deaton and Cartwright sum up these arguments nicely:

RCTs would be more useful if there were more realistic expectations of them and if their pitfalls were better recognized. For example, and contrary to many claims in the applied literature, randomization does not equalize everything but the treatment across treatments and controls, it does not automatically deliver a precise estimate of the average treatment effect (ATE), and it does not relieve us of the need to think about (observed or unobserved) confounders. [9,10]

Brenner’s Camden Coalition model certainly falls into the purview of socially complex phenomena with observed and unobserved confounding variables, as aptly described in the pithy summary above. Deaton and Cartwright go on to say that:

RCTs do indeed require minimal assumptions and can operate with little prior knowledge. This is an advantage when persuading distrustful audiences, but it is a disadvantage for cumulative scientific progress, where prior knowledge should be built upon and not discarded. RCTs can play a role in building scientific knowledge and useful predictions but they can only do so as part of a cumulative program, combining with other methods, including conceptual and theoretical develop-
These authors believe evaluating and learning “why things work” applies accurately to what is missing in the RCT of Brenner’s team’s work. RCTs are one way of understanding the causal phenomenon producing community outcomes, but can be dangerous if not approached with a broad understanding of the many factors in play.

RCTs are entirely appropriate for testing highly controllable interventional techniques against a control group; the gold standard for those who think we live in a controllable world, even when we know we don’t. But nothing in human life is ever controllable. It is vexingly difficult to create a controlled situation in the real and messy world.

One of the failures in the Camden story is confusing one evaluation for the end of a story, which, in reality, is still continuing. The Camden Coalition remains scrappy, smart, tenacious, and committed to the people of Camden, while constantly adjusting its work, informed by a variety of stakeholders, including many evaluators. Evaluators have a bigger toolbox of techniques than the RCT, and researchers and program leaders care—or should—about more than the limited outcomes measured as hospital readmissions for the 742 patients in the study.

Data can be cruel to one’s assumptions. In Memphis, in 2008, we hypothesized that the support of congregations would reduce length of stay for their connected, hospitalized patients. This was, of course, unfounded, which the early data quickly showed [11], but the very same data showed a surprising cost-savings anyway. We realized quite early in the process that the data showed that savings seemed to be coming from the fact that patients were presenting at the health system, slightly more likely to be at the right door, at the right time, ready for care, and not alone. The hospital was not doing any of that work, but hundreds of congregations were. A profound and intentional selection bias was at play: the appropriate bias toward people who are poor and vulnerable, by congregations that believe God expects them to love the most. An institutional review board or study design may rigidly exclude some subjects, but congregations would not exclude anyone they think God wants included at any time.

Moving Forward with a Messy Co-Labor Approach

Evaluators are servants of the “arts and crafts” of knowing what matters and choosing appropriate means of testing. It is a sensitive art of high moral purpose and must respect the reality of the work on the streets. What works there, if anything does, is messy co-laboring among all those who, for a variety of reasons, care about the poor and stigmatized in any given geography. That Messy Co-Labor (or MCL; it deserves capital letters) is not at all like the clean abstraction of Collective Impact, much less clinical trials that frequently promise outcomes easily dismissed by junior evaluators. MCL is rarely contained within a computer-selected cohort, drifting across its boundaries as soon as it enters the actual apartment or under the bridge, where it extends its kind heart to others along the way. MCL is more like what Brenner noticed: hot spots that are actually human neighborhoods that are alive and worth the trouble.

Numerous research and evaluation methods help us test a wider array of strategies and techniques, more suited to MCL. Mixed methods of individual or cohort-based case studies, coupled with qualitative pre-post data points can tell both the narrative and numbers “story” that speaks the language of diverse stakeholders of community members, donors, financial health system specialists, CEOs, and academics. Time series internal benchmarking of programs and outcomes as a version of program evaluation share deep narrative details of the iterative program changes (such as those that Brenner’s team adaptively made from 2011 to 2018), aligned with external variables. External variables impacting patient outcomes include the inevitable end of fiscal year shortages, noted by Camden Coalition staff: they could navigate patients to community resources, but often, those resources were not actually available from their partners. Natural experiments are another potentially valuable evaluation option, such as the multilevel interventions that decreased Latino infant mortality rates in a few years’ time, as shown by Dr. Sunny Anand and his team in Memphis [12].

Randomized community trials may be useful, representing randomization by group, although they are seen as less efficient statistically than randomization by individual. As we know from our work with faith communities, it is the context and culture of the community that informs how to best develop an intervention and adds to (or detracts from) its ultimate success in impacting health outcomes [11]. Another nuanced aspect of evaluation for any community-based intervention is the pre-work of religious or community health asset mapping, a methodology that the Memphis and North Carolina teams adapted to allow a deeper understanding of sites, history, and potential past traumas [13].

Lastly, we would call on the MIT researchers (and the
population health field as a whole) to engage social scientists to review the role that trust plays as a variable affecting the apparent success of the treatment, particularly in certain patient subsets. As pointed out by our colleague, Lucy Gilson of the University of Cape Town, whether a person is mandated to involuntarily “trust” or use the resources of a given provider within health systems affects the context and quality of that patient-provider relationship and can greatly inform health outcomes [14]. This is especially important when the patient group has historical reasons to not trust the privileged providers. Randomization by individual does not allow patients to choose which group or providers they see.

There is no going back inside the clinical box of the hospital. The insurance business and government simply will not pay for expensive disconnected interventions related to helping people live their lives, go to jobs, and be neighbors and citizens. People will pay for health, but increasingly resent paying for health care. In North Carolina, we are slowly moving toward a fundamentally different idea of how Medicaid should manage, and improve, the lives of its 1.8 million members. The key idea is to pay for social determinants, meaning social goods and services—rides, housing, food, and services related to interpersonal trauma.

The strategy is solid—invest in the social factors driving the conditions, which drive the cost. The problem is that the social factors actually are social, not commodities that are easy to deal out to those who need them. While it is good to offer Medicaid members an expanded list of goods, deliverables, and titrated services beyond the strictly biomedical, the outcomes will not be dramatically altered without a change in the social realities. We may need more systematic attention to non-individualized MCL to create social fabric where it is in shreds, to create webs of trust where there are currently only transactional, instrumental, and anonymous relationships now. The hard part is not in doing clinical things in new places such as really tough hot spots. The art is in becoming a part of the complex human system called community, which is exactly what models like the Camden Coalition have tried to do, and why the complexity of their approaches sometimes eludes traditional quantification and impact tracking.

Conclusion

The menu of possible creative actions represents many things being tried and adapted continuously. Some work, some do not, and some work in surprising ways, so every kind of evaluation is valued. Just because community-based approaches do not meet more traditional means of measuring success (e.g., RCTs) does not mean they should be discarded—a complex problem is best solved by a complex approach, which can be difficult to measure accurately and comprehensively. The Camden Coalition team looked through the eyes of computers, while the Memphis and North Carolina teams looked through the eyes of people of faith. All teams saw the same rich complexity (including needs and assets) of our most vulnerable and marginalized persons. That vision inspires us to tenaciously and thoughtfully do and evaluate our work of integrating medical and community-based care, so that the field moves forward, and, more importantly, improves the health and well-being of people in need, within an equity and justice framework.

References

Evaluating Two Mysteries: Camden Coalition Findings


**DOI**

https://doi.org/10.31478/201510b

**Suggested Citation**


**Author Information**

Teresa Cutts, PhD is a research assistant professor at Wake Forest School of Medicine, with her primary appointment in the Division of Public Health Science, Dept. of Social Sciences and Health Policy and is an affiliate in the Maya Angelou Center for Health Equity. Gary Gunderson, MDiv, DMin, is the Vice President of the FaithHealth Division at Wake Forest Baptist Health, has faculty appointments at both Wake Forest School of Medicine and Divinity School, and is a member of the Roundtable on Population Health Improvement.

**Acknowledgments**

The authors wish to thank Alina Baciu for her cogent editorial work and clarifying language that greatly improved this paper.

**Conflict-of-Interest Disclosures**

None to disclose.

**Correspondence**

Questions or comments should be directed to Teresa Cutts at cutts02@gmail.com.

**Disclaimer**

The views expressed in this paper are those of the author and not necessarily of the author’s organizations, the National Academy of Medicine (NAM), or the National Academies of Sciences, Engineering, and Medicine (the National Academies). The paper is intended to help inform and stimulate discussion. It is not a report of the NAM or the National Academies. Copyright by the National Academy of Sciences. All rights reserved.