Nearly a decade after the Patient Protection and Affordable Care Act was passed, we can now evaluate whether its component policies were successful. Rigorous evaluation is a critical part of the life cycle of health policies. However, most policy evaluations have an important limitation. We ask whether a policy has achieved its intended aims—ie, was this a good policy? However, policy makers need to answer whether today's policies should remain in effect. “Is continuing this policy good policy?” is often a very different question. To answer that, the critical question is whether the policy's targeted metrics still have room for improvement and whether improvements can be achieved without unacceptable unintended consequences.

Success Should Not Equal Permanence

For instance, the Hospital Readmissions Reduction Program (HRRP) penalizes hospitals with higher-than-expected readmission rates for various conditions. Readmissions decreased broadly and rapidly after the introduction of these penalties, by 1 to 3 percentage points for a variety of medical and surgical conditions.\(^1,2\) Thus, many experts initially thought the program was successful.\(^1\) However, after the rapid initial reduction in readmissions, further improvements slowed to the baseline trend. As the program has expanded to other conditions, additional improvements in those conditions’ readmissions rates have not followed.\(^2\)

Additionally, researchers and stakeholders have raised the possibility of unintended consequences, such as exacerbating health disparities and even paradoxically increasing mortality rates.\(^3\) The size of the reductions in readmissions has also come under question, given new data showing that increases in coded comorbidities explain at least half of the improvements in medical readmissions.\(^4,5\)

As any clinician can attest, not all readmissions should or could be avoided—a certain minimum rate of readmissions is inevitable and potentially even desirable. Therefore, it makes sense that readmissions rates would eventually reach a floor, ie, a point of diminishing improvements. More generally, as any policy meets its aims, the room for further improvement shrinks. However, the potential for harm remains, given that all policies take attention and resources away from other important priorities. Thus, the risk-benefit ratio may worsen as the policy remains in place. What should policy makers do to make prospective policy decisions in the face of retrospective data?

Design Policy With the End in Mind

Designing any policy must begin with understanding the scope of the problem, and it is not enough to state that the problem exists. It is equally critical to understand variation in the outcome. The degree of variation in readmission rates before HRRP had been described: in surgical conditions, the 50th percentile of risk-adjusted readmission rates was 13%, and the 25th percentile was 10%.\(^6\) Thus, it should have come as no surprise that readmission rates bottomed out after a reduction of 2 to 3 percentage points.\(^2\)

With this in mind, we should better define policies’ improvement goals in advance. According to the design of HRRP, hospitals with readmissions rates greater than their peers are penalized, no matter the absolute level. This continues the pressure to reduce readmissions rates (ie, to achieve

---

**Open Access.** This is an open access article distributed under the terms of the CC-BY License.
lower rates than peer hospitals) even after many hospitals within a peer group have achieved an acceptable level. A policy without an end point implies that improvements should continue indefinitely.

**Consider Randomized Deimplementation**

In many cases, it is challenging to say whether a policy reached its intended outcome without major unintended consequences. This is particularly true of Medicare payment policies, which today are either nearly universal (e.g., HRRP, Hospital Value-Based Purchasing) or voluntary (e.g., episodic bundling initiatives, accountable care organizations). As researchers often point out, both mandatory and voluntary programs pose serious challenges to policy evaluation.

In addition, it can be difficult to anticipate what would happen if a policy were repealed. If its effects were sticky, reversing the policy may not reverse all its benefits: improvements in technology and care processes could persist even without continued penalties. Alternately, the improvements could unwind as soon as the incentives do. How can we determine whether discontinuing a policy would preserve its benefits and limit its harms? If the effect of a policy is unclear or likely negative, policy makers have an opportunity to deimplement it in a thoughtful, controlled way.

This is not a hypothetical concept: randomized policy deimplementation had a pivotal role in evaluating restrictions to resident duty hours. In 2011, regulations were passed to limit all interns’ shifts to a 16-hour maximum to improve patient safety and resident well-being. However, there was concern that the regulations might decrease continuity of care because of more frequent shift changes. Given this equipoise, the Flexibility in Duty Hour Requirements for Surgical Trainees trial deimplemented the 2011 regulations on shift duration for surgical interns through a cluster-randomized design: residencies were randomized into 1 group in which the 2011 reforms were optional or into another in which the reforms remained. After 1 year, no difference in clinical outcomes was observed between both groups. Thus, the 2011 reforms were reversed for surgical programs nationwide. This experience illustrates how a policy that is initially introduced universally can still be deimplemented in a randomized way to better understand its effects.

Policy makers can also apply what is known as the stepped-wedge model for policy deimplementation, gradually rolling back policy elements in a randomized way to determine the optimal size and scope of incentives. For instance, when deimplementing readmissions penalties, the size of penalties could be scaled back gradually, with randomization at the regional or hospital level, to determine the lowest rate of penalties that maintains improvements in readmissions. Alternately, the breadth of targeted conditions could be narrowed until the spillover effects to other nontargeted conditions begin to diminish. Like down-titrating a medication, stepped-wedge deimplementation could facilitate carefully preserving of policy benefits (i.e., any delivery improvements that it encouraged) while scaling back harms.

Although HRRP has created significant controversy, the questions this policy has provoked are bound to repeat themselves without a more forward-looking policy framework. This requires appreciating that even a successful policy may have diminishing returns and that past success does not prevent ongoing and cumulative harm.
Conflicts of Interest Disclosures: Dr Chhabra reported receiving funding from the University of Michigan Institute for Healthcare Policy and Innovation Clinician Scholars Program, the Agency for Healthcare Research and Quality, and the National Institutes of Health Division of Loan Repayment outside the submitted work. Dr Ryan reported receiving grant funding related to this work from the National Institute on Aging outside the submitted work. Dr Dimick reported receiving grant funding from the National Institutes of Health, the Agency for Healthcare Research and Quality, and the BlueCross BlueShield of Michigan Foundation outside the submitted work and being a cofounder of ArborMetrix, Inc, a company that makes software for profiling hospital quality and efficiency.

Additional Information: Lisa Rosenbaum provided thoughtful comments on an earlier version of this article.

REFERENCES


