1. Bundled Payment: Effects on Health Care Spending and Quality

Closing the Quality Gap: Revisiting the State of the Science

Executive Summary

Background

This review is part of the Closing the Quality Gap: Revisiting the State of the Science series, which aims to provide critical analysis of the existing literature on quality improvement strategies for a selection of diseases and practices. The review focuses on “bundled payment,” a strategy for health care quality improvement and cost containment. This strategy has been the subject of increasing interest, with the Centers for Medicare & Medicaid Services announcing a large national bundled payment initiative in August 2011. Other reviews in the series will address a range of quality improvement topics arising from portfolios (areas of research) of the Agency for Healthcare Research and Quality (AHRQ).

We define “bundled payment” as a method in which payments to health care providers are related to the predetermined expected costs of a grouping, or “bundle,” of related health care services. In contrast, fee-for-service payment typically involves payments for individual services, while capitation involves a single per capita prospective payment for all services over a fixed period of time, regardless of the number of services or episodes of care provided within that period. Within the bundled payment model, a variety of specific payment methods are possible. Bundles may be defined in different ways, covering varying periods of time and including single or multiple health care providers of different types.

Evidence-based Practice Program

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

The full report and this summary are available at www.effectivehealthcare.ahrq.gov/reports/final.cfm.
types. In addition, given the diversity and complexity of the U.S. health care delivery system, bundled payment programs may be implemented in a variety of contexts that may influence their effects on spending and quality.

Bundled payment should create a financial incentive for providers to reduce the number and cost of services contained in the bundle. Providers are typically given discretion over the allocation of the services used to treat the patient’s episode most effectively. This flexibility may encourage providers to use resources to coordinate care; often, these services are not reimbursed under fee-for-service payment. If the bundle includes services delivered by multiple providers in multiple settings, providers have to create a mechanism for managing the shared payment for a given treatment or condition, which could also foster coordination.

Several types of undesired effects of bundled payment have also been postulated. The most significant potential undesired effects include underuse of effective services within the bundle, avoidance of high-risk patients, and an increase in the number of bundles reimbursed (increasing health spending). Providers under bundled payment may “game” the system by changing coding practices to maximize reimbursement for the bundle (“upcoding”) or by moving services in time or location to qualify for separate reimbursement (“unbundling”).

**Objectives**

This review was designed to address the uncertainties about the effects of bundled payment on spending and quality measures. It should help readers (1) understand what the evidence shows about the effects of bundled payment on health care spending and quality of care, and (2) understand key design and contextual features of bundled payment programs and their association with bundled payment effects.

The review addressed three Key Questions:

1. What does the evidence show on the effects of bundled payment versus usual (predominantly fee-for-service) payment on health care spending and quality measures?

2. Does the evidence show differences in the effects of bundled payment systems by key design features?

3. Does the evidence show differences in the effects of bundled payment systems by key contextual factors?

**Conceptual Framework**

We use the conceptual model in Figure A to understand these Key Questions. This model is based on ones developed by Dudley et al. and Andersen to describe organizations’ response to payment incentives in general and in the context of access to health care specifically. Several key design features define a particular set of incentives and disincentives associated with any specific bundled payment strategy. The impact of these design features is addressed by Key Question 2. The financial and nonfinancial characteristics of these incentives are primary determinants of an organization’s need to change practice in response to the modified payment policy. This response, however, may be mediated by key contextual factors, including both predisposing and enabling factors. Predisposing factors include the general financial environment (such as baseline levels of financial performance and efficiency), other incentives outside of the bundled payment program, market variables, and characteristics of participating provider organizations (such as charter and mission). Enabling factors include the capabilities and goals of participating organizations and the degree to which these organizations are integrated, as well as staff and patient characteristics. The impact of these contextual factors is addressed by Key Question 3.

The center of the model reflects how organizations respond to the incentives created by bundled payment in both desirable and undesirable ways. Key Question 1 addresses how different potential responses affect study outcomes, including health care spending and health care quality.
Methods

Topics for the Closing the Quality Gap: Revisiting the State of the Science series were solicited from the portfolio leads at AHRQ. The nominations included a brief background and context, the importance of and/or rationale for the topic, the focus or population of interest, relevant outcomes, and references to recent or ongoing work. Among the topics that were nominated, the following considerations were made in selection for inclusion in the series: the ability to focus and clarify the topic area appropriately, relevance to quality improvement and a systems approach, applicability to the Evidence-based Practice Center program/amenability to systematic review, potential for duplication and/or overlap with other known or ongoing work, relevance and potential impact in improving care, and fit of the topics as a whole in reflecting the AHRQ portfolios.

A Technical Expert Panel reviewed and provided input on topic definition, Key Questions, the search strategy, and preliminary search results. A draft report was reviewed by 11 peer reviewers and posted for public comment.

Studies published between January 1, 1985, and January 17, 2011, that address the Key Questions described above were included. The following studies were excluded: (1) studies that did not report any of the outcomes of interest; (2) studies that did not report on a bundled payment intervention as defined above; (3) background articles or...
articles strictly limited to describing theoretical models. Studies of interventions implemented in countries other than the United States were included if they met broad criteria for generalizability to the United States, such as implementation in a health care delivery organization comparable to one found in the United States.

A librarian performed the initial literature search. One trained reviewer, with input on questionable titles from a second trained reviewer, scanned the titles and abstracts of the list generated by the librarian and selected studies for full-text screen. For each of the selected studies, reviewers performed further reference mining by scanning titles listed in the reference section to identify additional articles to be included. Reviewers reconciled their selections and made joint decisions, following the inclusion/exclusion criteria listed above. Given the large and relatively older body of research on the Medicare Inpatient Prospective Payment System, the researchers, in consultation with AHRQ and the Technical Expert Panel, chose to consider a review of review articles for the assessment of this program.

We summarized the evidence for effectiveness and risks of bundled payment in comparison with usual payment methods. We present the results in a narrative synthesis and evidence tables (Appendix A).

We assessed the methodological quality of individual studies and reviews as good (low risk of bias), fair, or poor (high risk of bias). Studies rated “poor” or “good” were also given a brief explanation of the basis for the rating. The rating was based on criteria developed by AHRQ. We also rated the overall strength of the evidence using methods adapted by AHRQ from the GRADE (Grading of Recommendations Assessment, Development and Education) Working Group and assessed the overall applicability of the studies reviewed.

Results

We reviewed 58 studies, excluding studies of the Medicare Inpatient Prospective Payment System (IPPS), for which we reviewed 4 review articles. Among the reviewed studies, 48 employed observational designs, while 9 were descriptive. Only one study was randomized at the provider level, and we identified no studies of bundled payment programs randomized at the patient level. The included studies examined 20 different bundled payment interventions. Most articles examined U.S. public insurance prospective payment systems or international prospective payment systems. All but three of the bundled payment interventions in the included studies included public payers only. Bundled payment interventions may aggregate costs longitudinally (i.e., over time within a single provider), aggregate costs across providers, and/or involve warranties by which the costs of complications are rolled into a single payment. All but 4 of the 20 bundled payment interventions involved bundling of services during a period of time by a single provider, such as a hospital, skilled nursing facility, or home health provider.

Our search identified other bundled payment programs in progress, including the Medicare Acute Care Episode demonstration, PROMETHEUS Payment®, and others. However, we did not identify published evaluations of these programs that met our inclusion criteria. We briefly describe several of these programs in this review.

Despite the heterogeneity of settings, interventions, study designs, and measures used, reviewers noted relatively consistent impacts of bundled payment on spending and quality measures (Key Question 1), which are summarized below. Few studies explicitly included analyses of differential effects by key contextual factors (Key Question 3), and none included analyses of differential effects by key design features (Key Question 2). We did not attempt to address Key Questions 2 and 3 through comparisons across interventions because these analyses would be limited by the heterogeneity of the interventions and evaluation methods. Table A summarizes the studies included in the review, followed by our assessments of the evidence for each Key Question. The majority of studies were of U.S. public insurance prospective payment systems. The remainder of the studies were of a U.S. private-sector single-setting payment system, international bundled payment systems, and U.S. bundled payment systems including multiple providers or sites of care.
<table>
<thead>
<tr>
<th>Payment System</th>
<th>No. of Included Studies</th>
<th>Intervention Date</th>
<th>Payer</th>
<th>Provider Type(s)</th>
<th>Services Included in Bundle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>U.S. Public Insurance Prospective Payment Systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare Inpatient Rehabilitation Facility Prospective Payment System</td>
<td>11,19,30-38</td>
<td>January 2002</td>
<td>Medicare</td>
<td>Inpatient rehabilitation facilities</td>
<td>Per-discharge payment for operating and capital costs for each of 92 case-mix groups.</td>
</tr>
<tr>
<td>Medicare Home Health Prospective Payment System</td>
<td>5,11,19,39-47</td>
<td>October 2000</td>
<td>Medicare</td>
<td>Home health agencies</td>
<td>Payment per 60-day episode for all nursing care, therapy, and aide services for each of 153 Home Health Resource Groups.</td>
</tr>
<tr>
<td>Medicare Long-Term Acute Care Hospital Prospective Payment System</td>
<td>48</td>
<td>October 2002</td>
<td>Medicare</td>
<td>Long-term acute care hospitals</td>
<td>Per-discharge payment for all operating and capital costs for each of 318 Medicare long-term care diagnosis-related groups.</td>
</tr>
<tr>
<td>Department of Veterans Affairs Resource Allocation Methodology</td>
<td>49</td>
<td>1985-88</td>
<td>Department of Veterans Affairs</td>
<td>Inpatient hospitals</td>
<td>Per-admission payment (diagnosis-related groups) for hospital services.</td>
</tr>
<tr>
<td>Maine Medicaid Nursing Home Prospective Payment</td>
<td>50</td>
<td>1982-85</td>
<td>Maine Medicaid</td>
<td>Nursing homes</td>
<td>Per diem payment for nursing home services (patient care, dietary care, lodging, etc.).</td>
</tr>
<tr>
<td>New Hampshire Medicaid Hospital Prospective Payment</td>
<td>51</td>
<td>January 1989</td>
<td>New Hampshire Medicaid</td>
<td>Inpatient hospitals</td>
<td>Payment includes per-discharge prospective component based on diagnosis-related group as well as an actual costs component.</td>
</tr>
<tr>
<td><strong>U.S. Private-Sector Single-Setting Bundled Payment System</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case Rate for Managed Behavioral Health Care</td>
<td>52,53</td>
<td>July 1995</td>
<td>Single U.S. managed behavioral health care organization</td>
<td>Psychiatrists, psychologists, social workers</td>
<td>All outpatient mental health services for 1 year following referral.</td>
</tr>
<tr>
<td><strong>International Bundled Payment Systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium Inpatient Prospective Payment System</td>
<td>54</td>
<td>1995</td>
<td>Belgian public health insurance</td>
<td>Inpatient hospitals</td>
<td>Per-admission payment for nonmedical hospital services (e.g., nursing staff,</td>
</tr>
<tr>
<td>Payment System</td>
<td>No. of Included Studies</td>
<td>Intervention Date</td>
<td>Payer</td>
<td>Provider Type(s)</td>
<td>Services Included in Bundle</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>-------------------</td>
<td>--------------------------------------------</td>
<td>---------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>International Bundled Payment Systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Health Service (England)</td>
<td>155</td>
<td>2003</td>
<td>English National Health Service</td>
<td>Hospitals</td>
<td>Hospital services. (More detail not reported.)</td>
</tr>
<tr>
<td>Italy Inpatient Prospective Payment</td>
<td>156</td>
<td>January 1995</td>
<td>Italian Ministry of Health</td>
<td>Inpatient hospitals</td>
<td>Hospital and physician services. (More detail not reported.)</td>
</tr>
<tr>
<td>Japan Outpatient Hemodialysis Bundled Payment</td>
<td>157</td>
<td>April 2006</td>
<td>Not reported</td>
<td>Dialysis clinics and hospitals</td>
<td>Dialysis drugs. (More detail not reported.)</td>
</tr>
<tr>
<td>Netherlands Inpatient Prospective Payment</td>
<td>158</td>
<td>2005</td>
<td>Not reported</td>
<td>Inpatient hospitals</td>
<td>Hospital and physician services. (More detail not reported.)</td>
</tr>
<tr>
<td>Sweden Inpatient Prospective Payment System</td>
<td>159</td>
<td>1992</td>
<td>Stockholm County Council</td>
<td>Inpatient hospitals</td>
<td>Hospital services. (More detail not reported.)</td>
</tr>
<tr>
<td>Taiwan Hospital Case Payment</td>
<td>360-62</td>
<td>October 1997</td>
<td>Taiwan National Health Insurance Bureau</td>
<td>Hospitals</td>
<td>Per-episode payment for required medical services, “optional” medical services, pharmaceuticals, and evaluations within a 2-week period for surgeries.</td>
</tr>
<tr>
<td><strong>U.S. Bundled Payment Systems Including Multiple Providers/Sites of Care</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geisinger ProvenCareSM</td>
<td>163</td>
<td>February 2006</td>
<td>Geisinger Health Plan, PA</td>
<td>Three demonstration hospitals</td>
<td>Per-discharge payment for evaluation, hospital, and professional fees; routine postdischarge care; management of complications.</td>
</tr>
<tr>
<td>Medicare Participating Heart Bypass Demonstration</td>
<td>164</td>
<td>May 1991-June 1996</td>
<td>Medicare</td>
<td>Seven demonstration hospitals</td>
<td>Per-discharge payment for all inpatient hospital and physician services, including related readmissions.</td>
</tr>
<tr>
<td>Payment System</td>
<td>No. of Included Studies</td>
<td>Intervention Date</td>
<td>Payer</td>
<td>Provider Type(s)</td>
<td>Services Included in Bundle</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------------------</td>
<td>------------------------------------</td>
<td>--------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Medicare Cataract Surgery Alternative Payment Demonstration</td>
<td>165</td>
<td>1991-April 1996</td>
<td>Medicare</td>
<td>Four demonstration sites</td>
<td>Preoperative diagnostics and evaluation, surgical services, postoperative exams to 120 days, management of common complications.</td>
</tr>
<tr>
<td>Michigan Arthroscopic Surgery Bundling Pilot With Warranty</td>
<td>166</td>
<td>April 1987</td>
<td>Blue Care Network (Blue Cross Blue Shield of MI)</td>
<td>Single surgical practice and hospital</td>
<td>Single payment for all care related to surgery, including repeat surgery, repeat hospitalization, or any other related services rendered by the providers for 2 years.</td>
</tr>
</tbody>
</table>

*Implementation was delayed in four States.

Source: Authors’ analysis of search results.
**Key Question 1: Impact of Bundled Payment on Health Care Spending and Quality Measures**

The published evidence suggests that transitioning from cost-based or fee-for-service payment to bundled payment resulted in declines in spending and utilization, with small changes in quality measures that were in different directions. The evidence suggests that the transition from a cost-based or fee-for-service reimbursement to bundled payment was generally associated with a decline in spending of 10 percent or less. Bundled payment was associated with a decrease in utilization of services included in the bundle, often measured as reductions in length of stay or utilization of specific services (5-percent to 15-percent reductions in many cases).

The inconsistency in findings on quality measures included both differences in the direction and magnitude of effects on different quality measures within a single study and differences in the direction and magnitude of effects for similar quality measures between studies. For a given bundled payment intervention, either some quality measures improved while others worsened or studies arrived at different conclusions about the effect of bundled payment on related quality measures. Little evidence was reported about other potential negative consequences of bundled payment, although studies of several programs noted that bundled payment programs resulted in shifts of utilization to other settings of care.

Overall, reviewers graded the strength of evidence for this Key Question as “low,” indicating that there is low confidence that the evidence reflects the true effect, and that further research is likely to change our confidence in the estimate of effect and is likely to change the estimate. This rating was based primarily on reviewers’ assessments that the included studies were, as a whole, at high risk of bias and residual confounding, although the expected direction of the effect of residual confounding could not be assessed. The results of included studies were consistent in the direction of the effect for spending and utilization measures but inconsistent for quality measures. Based on the consistency of findings across heterogeneous interventions and evaluations, it is likely that the direction of the observed effects on spending and utilization measures would not change in future studies, although the magnitude of the effect could change in studies employing different methods for addressing bias and confounding.

**Key Question 2: Differential Effects by Key Design Features**

No studies explicitly tested the effect of intervention design features, such as variations in the set of services included in a bundle, on spending or quality measures. We did not perform comparisons of design-feature impacts across studies because of the heterogeneity of the bundled payment programs studied. We do, however, provide some discussion of the potential impact of design features on study outcomes. Reviewers graded the evidence for this Key Question as insufficient to permit an estimation of effects due to the lack of evidence.

The reviewed studies included a heterogeneous set of bundled payment programs. Reviewed payment systems differed in the degree to which the bundled payment applied to multiple independent providers and/or provider types. Only the Medicare Participating Heart Bypass Demonstration, Medicare Cataract Surgery Alternative Payment Demonstration, Michigan Arthroscopic Surgery Bundling Pilot, and Geisinger ProvenCareSM Program integrated physician payments with hospital payments when these payments were previously separate. Studies of these programs found evidence for reduced spending with inconsistent effects on different quality measures. However, there was limited basis for a comparison of the magnitude of these findings with the magnitude of findings from studies of bundled payment interventions that apply to a single institutional provider. The two Medicare demonstrations reported some difficulty in administration of the bundled payment programs, in part due to challenges of distributing payment among the multiple providers participating in the bundled payment.

Overall, the reviewed payment systems employed a variety of bundle definitions that were tailored to the relevant care setting. Risk adjustment and separate outlier payments were common, with methods varying among programs.

Quality metrics or incentives were rarely integrated into bundled payment systems. Despite the potential for undesired effects of bundled payment on quality of care, programs generally did not include quality as an intrinsic part of the bundled payment mechanism. Among the programs reviewed, only the Geisinger ProvenCareSM Program integrated pay-for-performance components into a bundled payment system, and therefore no differential effects across evaluations can be reported for this design feature. It is possible that inclusion of quality-related incentives as a component of future bundled payment programs will change providers’ response to the program in a way that impacts quality measures, but the effect is unknown at this time. Other programs used quality
measurement in program monitoring and evaluation or as a criterion for program participation but not as a component of the payment method. Many of the bundled payment programs studied were implemented prior to the recent proliferation of pay-for-performance programs. The science of health care quality measurement and quality incentives has developed since many of the programs were implemented. In some cases, the bundled payment programs reviewed will be accompanied in the future by a separate pay-for-performance program that addresses the same providers and services. It is unclear how these pay-for-performance programs will interact with the bundled payment programs studied or the differential impact on quality of including the quality-related incentives as an integrated part of bundled payment or as a stand-alone program. Among more recent bundled payment programs that have been announced or initiated but not yet evaluated, some use quality measurement in a monitoring and evaluation role (e.g., Medicare national bundled payment initiative), while others incorporate pay-for-performance with the bundled payment (e.g., PROMETHEUS Payment).

Reviewed payment systems also differed in their approach to establishing initial bundle reimbursement rates. Historical expenditures were typically used to determine the initial bundled payment rates, but programs differed in whether the bundled payment rate was set at an amount estimated to increase, decrease, or maintain historical expenditure levels.

Several reviewed studies either directly studied the implementation process of one or more bundled payment systems or included implementation-related anecdotes. Most studies providing survey or anecdotal evidence from providers noted that new bundled payment systems faced significant initial resistance from providers.

**Key Question 3: Differential Effects by Key Contextual Factors**

The effects of most key contextual factors were not addressed by any reviewed studies. Several studies compared differential effects of bundled payment on spending among for-profit and not-for-profit providers. In general, for-profit providers experienced larger declines in utilization under bundled payment than their not-for-profit counterparts (including U.S. and non-U.S. hospitals, skilled nursing facilities, inpatient rehabilitation facilities, and home health agencies). The four review articles on the Medicare IPPS reported that hospitals under greater financial pressure had greater reductions in utilization in response to IPPS.

Several studies reported outcomes separately for patients with relatively severe disease or injury and patients with less severe disease or injury, but the results were inconsistent between studies. One study reported different impacts of bundled payment on spending and utilization by geographic area but did not provide an explanation of this differential impact.

For the effects of three specific contextual factors—patient severity, provider for-profit/not-for-profit status, and provider financial pressure—reviewers graded the strength of evidence on one outcome, utilization, as “low.” The primary reasons for this rating were risk of bias, risk of residual confounding, inconsistency of findings, and imprecision of findings. The strength of evidence for other contextual factors and study outcomes was rated insufficient due to lack of evidence.

Several important contextual factors were described in reviewed studies, but their impact was not directly assessed. Some bundled payment interventions were implemented in the context of simultaneous but independent health care spending reduction efforts, including payment reforms other than bundled payment. For example, the Medicare Resource-Based Relative Value Scale physician fee schedule was implemented in 1992, overlapping with the time periods examined in several of the reviewed studies. Capitation and the spread of health maintenance organizations are two other examples of reforms that occurred during the time period of reviewed bundled payment programs. Due to the study designs used, studies were not able to differentiate the effects of bundled payment programs from related but independent interventions.

None of the reviewed studies provides insight into differential results by the degree of integration between health care delivery organizations and payers or between various health care delivery organizations. As noted above, most reviewed bundled payment interventions applied to a specific outpatient care setting or to inpatient environments where hospital and physician payments are not integrated. None of the reviewed studies reported on the effects of payer or provider competitive environments.

**Applicability**

The main intended audience for this report is policymakers, payers, and providers in the United States that are considering implementation of a bundled payment program. The findings of this review are likely to be applicable most directly to those considering a bundled payment program targeting single providers. The majority of bundled payment programs in the included
studies focused on single institutional providers, such as inpatient hospitals, skilled nursing facilities, or inpatient rehabilitation facilities.

Recent interest in bundled payment has focused mostly on programs that bundle payments for multiple providers and/or provider types. The findings of this review are less applicable to these types of programs. There were several reviewed studies of bundled payment across multiple settings, but these included a small number of participating provider organizations that are not representative of the U.S. delivery system. More evidence is likely to be available in the future as evaluations of ongoing programs, such as the Medicare Acute Care Episode Demonstration, are published.

The interventions studied were typically specific to a single payer, most often Medicare or various public insurance systems outside the United States. The applicability of findings in studies involving one payer (e.g., Medicare) to other payer contexts is limited due to differences in beneficiary characteristics, provider bargaining power, and competitive pressures.

Interventions implemented in countries other than the United States may have limited applicability to the U.S. context due to differences in health system organization, financing, and delivery. Although non-U.S. studies were screened for comparability with a U.S. delivery setting, in practice this criterion was difficult to apply and no studies were excluded on the basis of nationality.

All reviewed studies assessed the impact of bundled payment relative to either fee-for-service or cost-based payment. The magnitude and direction of effects relative to fee-for-service or cost-based payment may differ from absolute effects. For example, bundled payment might slow an increase in absolute spending relative to usual payment. Transitions to bundled payment from other payment methods (e.g., salary or capitation) may have other effects.

Finally, evidence on bundled payment applies specifically to cases in which reimbursement based on episodes of care is both reasonable and feasible. Bundled payment may be less feasible or effective when applied to health care related to conditions without clearly defined treatment regimens, conditions with multiple treatment approaches, or rare conditions.

**Recommendations for Further Research**

There are several ways that future evaluations could produce a stronger evidence base. Policymakers and evaluation researchers must recognize the tension between producing timely practical evidence and conducting rigorous evaluations. The most rigorous study designs are usually feasible only when policymakers plan for an evaluation experiment in the course of implementation. We focus our recommendations on improving retrospective quasi-experimental studies, which formed the bulk of research reviewed for this report.

**Use stronger evaluation study designs.** Most reviewed studies used a pre-post design with no comparison group and a relatively short post period. Use of stronger study designs, such as difference-in-differences analyses with randomized control and intervention groups, would reduce risk of bias. The benefits to validity from including a comparison group likely outweigh the associated increase in evaluation time and cost.

Most studies considered a brief time horizon (less than 2 years) after the implementation of bundled payment systems. Given the challenges in implementation and redesign of care processes reported in studies with qualitative components, a longer time horizon is likely necessary to observe many important impacts. However, the benefits of a longer time horizon must be balanced against the need for rapid information on program effects and the risk of attributing changes due to secular variation to the payment intervention. Future evaluations could balance these needs by reporting at several points in time after intervention.

Practical data, time, and funding constraints often dictate the choice of evaluation study design. The limitations intrinsic to retrospective observational studies prevent the reviewed studies from approaching the “gold standard” equivalent of a randomized controlled trial. However, the two studies rated good and several studies rated fair outlined natural experiments or quasi-experimental strategies to identify the effect of bundled payment on spending and quality outcomes. Future evaluations should consider these and related methods that could improve evaluation validity with little effect on the timeliness of results.

**Use standardized measures of impact on costs and quality.** The measures used varied across studies. While different measures are likely more relevant to different interventions and implementations, increased consistency in the use of measures could increase the comparability of future evaluations of the impact of bundled payment. Collaboratives of evaluators, such as the Patient-Centered Medical Home Evaluators’ Collaborative, are one potential mechanism for identifying priority measurement areas and measurement approaches. Evaluation sponsors could also encourage evaluators to use standard measurement approaches or to collaborate with evaluators of similar interventions to harmonize measurement approaches.
Use stratification to understand the impact of bundled payment on specific patient groups. Few studies reported results for patient subgroups, but several that did so found significant differences in effects between subgroups. Future research on the impact of payment system changes should focus on specific patient populations (e.g., relatively seriously ill patients) and types of service.

Incorporate quantitative and qualitative measures of program design and contextual factors. This type of evidence will be critical as Medicare and private payers experiment with various payment and delivery reforms, including not only bundled payment but also reforms that include global payment or shared savings for accountable care organizations, and as they seek to identify which aspects of these pilots have the potential to be scaled widely. Important design features to be addressed include the definition of the bundle (How many providers are included? What is the length of time? Which services are included and excluded from the bundle?); methods for limiting financial risk, including risk adjustment and outlier payments; use of quality measurement; and methods for distributing payment among participating providers. Important contextual factors to be addressed include whether bundled payment is more effective in more highly integrated delivery settings, the role of financial pressure and the general financial environment on responses to bundled payment, and differential effects between subgroups of patients.

Incorporate measurement of ancillary or spillover effects. Only a handful of studies explored broader consequences of bundled payment beyond the setting of care or patient group targeted by the intervention. Several reviewed studies demonstrated that bundled payment programs had effects on other settings and patient groups. Future evaluations should be designed to detect these effects.

Conclusion

In summary, the introduction of bundled payment was associated with (1) reductions in health care spending and utilization, and (2) inconsistent and generally small effects on quality measures. These findings were consistent across different bundled payment programs and settings, but the strength of the body of evidence was rated as low, due mainly to concerns about bias and residual confounding. These findings were subject to several important limitations. Most of the bundled payment interventions studied in reviewed articles (16/20) were limited to payments to single institutional providers (e.g., hospitals, skilled nursing facilities) and so have limited generalizability to newer programs that include multiple providers and/or multiple provider types. Exclusion criteria and the search strategy we used may have omitted some relevant studies from the results. The review is limited by the quality of the underlying studies. The interventions studied were often incompletely described in the reviewed articles.

For policymakers considering implementation of bundled payment programs, this evidence provides some support that the programs are likely to be an effective strategy for reducing health care spending. While the effects on health care quality are less certain, the available evidence does not support concerns about the worst potential adverse effects of bundled payment.

Policymakers considering bundled payment programs should also consider several caveats, however. First, future bundled payment programs are likely to differ in important ways from those reviewed in this study. Most of the programs reviewed were single-setting prospective payment systems that replaced fee-for-service reimbursement systems. In contrast, recent and forthcoming bundled payment programs, such as the Medicare national bundled payment initiative, focus on bundling services provided by different providers over the course of an episode defined by a condition, diagnosis, or procedure. The few completed studies of programs involving payment for multiple providers found evidence for reduced costs and inconsistent impacts on quality, although there were some reports of implementation difficulty. These programs are likely to be more complex than most of the reviewed programs and therefore may have different effects. Nevertheless, multiple-setting programs present a logical next step building on the largely positive effects of previous single-setting bundled payment programs.

Second, although evidence of effects on quality was inconsistent, bundled payment has the potential to either adversely affect quality or be used as part of a quality improvement strategy. Future bundled payment programs should incorporate a robust quality monitoring and improvement component, potentially as an integrated part of the payment system.

Third, the strength of evidence on bundled payment effects was low, reflecting the difficulty in evaluating large-scale policy interventions that occur in a rapidly changing health care system. Better information from evaluations could improve the impact of bundled payment programs, in particular by illustrating how the programs could be adapted for adoption in the variety of health care delivery contexts found in the United States.
References


Full Report


For More Copies

For more copies of Bundled Payment: Effects on Health Care Spending and Quality. Closing the Quality Gap: Revisiting the State of the Science: Evidence Report/Technology Assessment Executive Summary No. 208 (AHRQ Publication No. 12-E007-1), please call the AHRQ Publications Clearinghouse at 800–358–9295 or email ahrqpubs@ahrq.gov.